An Iterative Method with General Convex Fidelity Term for Image Restoration*

Miyoun Jung, Elena Resmerita, and Luminita Vese

Department of Mathematics, University of California, Los Angeles, U.S.A.
Industrial Mathematics Institute, Johannes Kepler University, Linz, Austria

Abstract. We propose a convergent iterative regularization procedure based on the square of a dual norm for image restoration with general (quadratic or non-quadratic) convex fidelity terms. Convergent iterative regularization methods have been employed for image deblurring or denoising in the presence of Gaussian noise, which use L^2 [1] and L^1 [2] fidelity terms. Iusem-Resmerita [3] proposed a proximal point method using inexact Bregman distance for minimizing a general convex function defined on a general non-reflexive Banach space which is the dual of a separable Banach space. Based on this, we investigate several approaches for image restoration (denoising-deblurring) with different types of noise. We test the behavior of proposed algorithms on synthetic and real images. We compare the results with other state-of-the-art iterative procedures as well as the corresponding existing one-step gradient descent implementations. The numerical experiments indicate that the iterative procedure yields high quality reconstructions and superior results to those obtained by one-step gradient descent and similar with other iterative methods.

Keywords: proximal point method, iterative regularization procedure, image restoration, bounded variation, Gaussian noise, Laplacian noise.

1 Introduction

Proximal point methods have been employed to stabilize ill-posed problems in infinite dimensional settings, using L^2 [1] and L^1 data-fitting terms [2], respectively. Recently, Iusem-Resmerita [3] proposed a proximal point method for minimizing a general convex function defined on a general non-reflexive Banach space which is the dual of a separable Banach space. Our aim here is to investigate and propose, based on that method, several iterative approaches for image restoration.

In Tadmor et al [4], an iterative procedure for computing hierarchical (BV, L^2) decompositions has been proposed for image denoising, and this was extended to more general cases for image restoration and segmentation in [5]. Osher et al [1]

* This research has been supported in part by a UC Dissertation Year Fellowship, by Austrian Science Fund Elise Richter Scholarship V82-N18 FWF, and by NSF-DMS grant 0714945.

K. Daniilidis, P. Maragos, N. Paragios (Eds.): ECCV 2010, Part I, LNCS 6311, pp. 185–199, 2010.

proposed another iterative procedure for approximating minimizers of quadratic objective functions, with the aim of image denoising or deblurring, providing significant improvements over the standard model introduced by Rudin, Osher, Fatemi (ROF) [6]. This turned out to be equivalent to proximal point algorithm on a nonreflexive Banach space as well as to an augmented Lagrangian method for a convex minimization problem subject to linear constraints. In addition, He et al [2] generalized the Bregman distance based iterative algorithm [1] to L^1 fidelity term by using a suitable sequence of penalty parameters, and proved the well-definedness and the convergence of the algorithm with L^1 fidelity term, which is an iterative version of L^1-TV considered by Chan and Esedoglu [7], and presented denoising results in the presence of Gaussian noise. Benning and Burger [8] derived basic error estimates in the symmetric Bregman distance between the exact solution and the estimated solution satisfying an optimality condition, for general convex variational regularization methods. Furthermore, the authors of [8] investigated specific error estimates for data fidelity terms corresponding to noise models from imaging, such as Gaussian, Laplacian, Poisson, and multiplicative noise.

Recently, Iusem and Resmerita [3] combine the idea of [1] with a surjectivity result, shown in [9] and [10], in order to obtain a proximal point method for minimizing more general convex functions, with interesting convergence properties. For the optimization case where the objective function is not necessarily quadratic, they use a positive multiple of an inexact Bregman distance associated with the square of the norm as regularizing term; a solution is approached by a sequence of approximate minimizers of an auxiliary problem. Regarding the condition of being the dual of a Banach space, we recall that nonreflexive Banach spaces which are duals of other spaces include the cases of l_∞ and L^∞, l_1 and BV (the space of functions of bounded variation) which appear quite frequently in a large range of applications.

Here we apply the proximal point method introduced in [3] to general ill-posed operator equations and we propose several algorithms for image restoration problems, such as image deblurring in the presence of noise (for Gaussian or Laplacian noise with corresponding convex fidelity terms). Finally, numerical results are given for each image restoration model. Comparisons with other methods of similar spirit or one-step gradient descent models are also presented.

2 Proposed Iterative Method for Solving Ill-Posed Operator Equations

Our proposed iterative method is based on the proximal point method and convergence results of Iusem-Resmerita from [3]. We first recall the necessary definitions and terminology. Let X be a nonreflexive Banach space and X^* its topological dual. For $u^* \in X^*$ and $u \in X$, we denote by $\langle u^*, u\rangle = u^*(u)$ the duality pairing. Denote by $h(u) = \frac{1}{2}\|u\|^2$, for $u \in X$.

For $\varepsilon > 0$, the ε-subdifferential of h at a point $u \in X$ is [11]

$$\partial_\varepsilon h(u) = \{u^* \in X^* : h(v) - h(u) - \langle u^*, v - u\rangle \geq -\varepsilon, \forall v \in X\}.$$

The normalized ε-duality mapping of X, introduced by Gossez [9], extends the notion of duality mapping as follows

$$J_\varepsilon(u) = \{u^* \in X^* : \langle u^*, u\rangle + \varepsilon \geq \frac{1}{2}\|u^*\|^2 + \frac{1}{2}\|u\|^2\}, \tag{1}$$

and an equivalent definition for the ε-duality mapping is $J_\varepsilon(u) = \partial_\varepsilon\left(\frac{1}{2}\|u\|^2\right)$.

The inexact Bregman distances with respect to the convex function h and to an ε-subgradient ξ of h were defined in [3] as follows:

$$D^\varepsilon(v, u) = h(v) - h(u) - \langle \xi, v - u\rangle + \varepsilon. \tag{2}$$

Note that $D^\varepsilon(v, u) \geq 0$ for any $u, v \in X$ and $D^\varepsilon(u, u) = \varepsilon > 0$ for all $u \in X$.

Given $\varepsilon \geq 0$ and a function $g : X \to \mathbb{R} \cup \{+\infty\}$, we say that $\bar{u} \in dom\, g = \{u \in X : g(u) < \infty\}$ is an ε-minimizer of g when

$$g(\bar{u}) \leq g(u) + \varepsilon \tag{3}$$

for all $u \in dom\, g$.

The following proximal point algorithm is proposed in [3] by Iusem and Resmerita.

Initialization
Take $u_0 \in dom\, g$ and $\xi_0 \in J_{\varepsilon_0}(u_0)$.

Iterative step
Let $k \in \mathbb{N}$. Assume that $u_k \in dom\, g$ and $\xi_k \in J_{\varepsilon_k}(u_k)$ are given. We proceed to define u_{k+1}, ξ_{k+1}. Define $D^{\varepsilon_k}(u, u_k) = h(u) - h(u_k) - \langle \xi_k, u - u_k\rangle + \varepsilon_k$ and $\bar{\varepsilon}_k = \lambda_k \varepsilon_{k+1}$.

Determine $u_{k+1} \in dom\, g$ as an $\bar{\varepsilon}_k - minimizer$ of the function $g_k(u)$ defined as

$$g_k(u) = g(u) + \lambda_k D^{\varepsilon_k}(u, u_k), \tag{4}$$

that is to say, in view of (3),

$$g(u_{k+1}) + \lambda_k D^{\varepsilon_k}(u_{k+1}, u_k) \leq g(u) + \lambda_k D^{\varepsilon_k}(u, u_k) + \bar{\varepsilon}_k \tag{5}$$

for all $u \in dom\, g$.

Let $\eta_{k+1} \in \partial g(u_{k+1})$ and $\xi_{k+1} \in J_{\varepsilon_{k+1}}(u_{k+1})$ such that

$$\eta_{k+1} + \lambda_k(\xi_{k+1} - \xi_k) = 0, \tag{6}$$

using two exogenous sequences $\{\varepsilon_k\}$ (summable) and $\{\lambda_k\}$ (bounded above) of positive numbers. By comparison to other iterative methods, Iusem-Resmerita method has several advantages: it allows very general function g and very general regularization; at each step, it is theoretically sufficient to compute only an $\bar{\varepsilon}_k$-minimizer, thus some error is allowed. On the other hand, Iusem-Resmerita method requires the use of the full norm on X (and not only a semi-norm), thus the method may be in practice computationally more expensive.

We now apply this general Iusem-Resmerita algorithm [3] to linear ill-posed inverse problems. Large classes of inverse problems can be formulated as operator equations $Ku = y$.

We define the residual $g(u) = S(y, Ku)$ for any $u \in X$, where S is a similarity measure (see, e.g., [12], [8]). The iterative method introduced in [3] can be applied to this exact data case setting and provides weakly* approximations for the solutions of the equation, provided that at least one solution exists.

Usually, the above equations $Ku = y$ are ill posed, in the sense that the operator K may not be continuously invertible which means that small perturbations in the data y lead to high oscillations in the solutions.

Consider that only noisy data y^δ are given, such that

$$S(y^\delta, y) \le r(\delta), \quad \delta > 0, \tag{7}$$

where $r = r(\delta)$ is a function of δ with

$$\lim_{\delta \to 0_+} r(\delta) = 0. \tag{8}$$

Denote

$$g^\delta(u) = S(y^\delta, Ku).$$

We show now that the general iterative method presented above yields a regularization method for such problems. We will use the following

Assumptions (A)

- The operator $K : X \to Y$ is linear and bounded, and yields an ill-posed problem.
- X and Y are Banach spaces. In addition, X is the topological dual of a separable Banach space.
- The similarity measure S is such that
  1. The function $g^\delta(u) = S(y^\delta, Ku)$ is convex and weakly* lower semicontinuous.
  2. $$\lim_{\delta \to 0_+} g^\delta(u_\delta) = 0 \quad \Rightarrow \quad \lim_{\delta \to 0_+} Ku_\delta = y, \tag{9}$$

     whenever $\{u_\delta\}_{\delta > 0}$ is a net in X, the last limit being understood with respect to the norm of Y.

We consider a positive constant parameter c. The method reads as follows:

Algorithm 1. *Take $u_0 \in dom\, g^\delta$ and $\xi_0 \in J_{\varepsilon_0}(u_0)$.*

Iterative step

Let $k \in \mathbb{N}$. Assume that $u_k \in dom\, g^\delta$ and $\xi_k \in J_{\varepsilon_k}(u_k)$ are given. We proceed to define u_{k+1}, ξ_{k+1}. Define $D^{\varepsilon_k}(u, u_k) = h(u) - h(u_k) - \langle \xi_k, u - u_k \rangle + \varepsilon_k$ and $\bar{\varepsilon}_k = c\varepsilon_{k+1}$.

Determine $u_{k+1} \in dom\, g^\delta$ *as an* $\bar{\varepsilon}_k$*-minimizer of the function* $g_k^\delta(u)$ *defined as*

$$g_k^\delta(u) = g^\delta(u) + cD^{\varepsilon_k}(u, u_k),$$

that is to say,

$$g^\delta(u_{k+1}) + cD^{\varepsilon_k}(u_{k+1}, u_k) \le g^\delta(u) + cD^{\varepsilon_k}(u, u_k) + \bar{\varepsilon}_k$$

for all $u \in dom\, g^\delta$.

Let $\eta_{k+1} \in \partial g^\delta(u_{k+1})$ *and* $\xi_{k+1} \in J_{\varepsilon_{k+1}}(u_{k+1})$ *such that*

$$\eta_{k+1} + c(\xi_{k+1} - \xi_k) = 0.$$

A posteriori strategy. We choose the stopping index based on a discrepancy type principle, similarly to the one in [1]:

$$k_* = \max\{k \in \mathbb{N} : g^\delta(u_k) \ge \tau r(\delta)\}, \tag{10}$$

for some $\tau > 1$.

We show that the stopping index is finite and that Algorithm 1 together with the stopping rule stably approximate solutions of the equation (proof included in a longer version of this work [13]).

Proposition 1. *Let* $\tilde{u} \in X$ *verify* $K\tilde{u} = y$, *assume that inequality (7) is satisfied, assumptions (A) hold and that the sequence* $\{\varepsilon_k\}$ *is such that*

$$\sum_{k=1}^{\infty} k\varepsilon_k < \infty. \tag{11}$$

Moreover, let the stopping index k_* *be chosen according to (10). Then* k_* *is finite, the sequence* $\{\|u_{k_*}\|\}_\delta$ *is bounded and hence, as* $\delta \to 0$, *there exists a weakly**-*convergent subsequence* $\{u_{k(\delta_n)}\}_n$ *in* X. *If the following conditions hold, then the limit of each weakly** *convergent subsequence is a solution of* $Ku = y$:

i) $\{k_*\}_{\delta>0}$ *is unbounded;*

*ii) Weak**-*convergence of* $\{u_{k(\delta_n)}\}_n$ *to some* $u \in X$ *implies convergence of* $\{Ku_{k(\delta_n)}\}_n$ *to* Ku, *as* $n \to \infty$ *with respect to the norm topology of* Y.

A priori strategy. One could stop Algorithm 1 by using a stopping index which depends on the noise level only, by contrast to the previously chosen k_* which depends also on the noisy data y^δ. More precisely, one chooses

$$k(\delta) \sim \frac{1}{r(\delta)}. \tag{12}$$

One can also show that the sequence $\{u_{k(\delta)}\}_{\delta>0}$ converges weakly* to solutions of the equation as $\delta \to 0$.

Proposition 2. *Let* $\tilde{u} \in X$ *verify* $K\tilde{u} = y$, *assume that inequality (7) is satisfied, assumptions (A) hold and that the sequence* $\{\varepsilon_k\}$ *obeys (11). Moreover, let* *the stopping index $k(\delta)$ be chosen according to (12). Then the sequence $\{\|u_{k(\delta)}\|\}_\delta$ is bounded and hence, as $\delta \to 0$, there exists a weakly*- convergent subsequence $\{u_{k(\delta_n)}\}_n$ in X. If the following condition holds, then the limit of each weak* convergent subsequence is a solution of $Ku = y$: weak*-convergence of $\{u_{k(\delta_n)}\}_n$ to some $u \in X$ implies convergence of $\{Ku_{k(\delta_n)}\}_n$ to Ku, as $n \to \infty$ with respect to the norm topology of Y.*

3 Several Proximal Point Based Approaches for Image Restoration

We present a few image restoration settings which fit the theoretical framework investigated in the previous section. We assume that noisy blurry data f corresponding to y^δ is given, defined on a open and bounded domain Ω of $\mathbb{R}^N$. First, we briefly mention prior relevant work in image processing.

In Tadmor et al [4], an iterative procedure for computing hierarchical (BV, L^2) decompositions has been proposed for image denoising, and this was extended to more general cases for image restoration and segmentation in [5]. For image deblurring in the presence of Gaussian noise, assuming the degradation model $f = Ku + n$, the iterative method from [5] computes a sequence u_k, such that each u_{k+1} is the minimizer of $\lambda_0 2^k \|v_k - Ku_{k+1}\|_2^2 + \int_\Omega |Du_{k+1}|$, where $v_{-1} = f$, $k = 0, 1, ...$ and $v_k = Ku_{k+1} + v_{k+1}$. The partial sum $\sum_{j=0}^k u_j$ is a denoised-deblurred version of f, and converges to f as $k \to \infty$.

Osher et al [1] proposed an iterative algorithm with quadratic fidelity term S and a convex regularizing functional h (e.g. TV-regularizer $h(u) = \int_\Omega |Du|$): starting with u_0, u_{k+1} is a minimizer of the functional

$$g_k(u) = S(f, Ku) + D(u, u_k) = \frac{\lambda}{2}\|f - Ku\|_2^2 + [h(u) - h(u_k) - \langle p_k, u - u_k\rangle], \quad (13)$$

where $p_k = p_{k-1} + \lambda K^*(f - Ku_k) \in \partial h(u_k)$ and $\lambda > 0$ is a parameter. The authors of [1] proved the well-definedness and the convergence of iterates u_k, and presented some applications to denoising or deblurring in the presence of Gaussian noise, obtaining significant improvement over the standard Rudin et al. model [6,14], which is

$$\min_u \left\{ \frac{\lambda}{2}\|f - Ku\|_2^2 + \int_\Omega |Du| \right\}. \quad (14)$$

We also refer to [15] where convergence rates for the iterative method (13) are established.

He et al [2] modified the above iterative algorithm [1] by using the varying parameter $\frac{1}{2^k\lambda}$ with $\lambda > 0$ instead of fixed parameter $\lambda > 0$, inspired by [16] and [4]:

$$g_k(u) = S(f, u) + D(u, u_k) = S(f, u) + \frac{1}{2^k\lambda}[h(u) - h(u_k) - \langle p_k, u - u_k\rangle], \quad (15)$$

where $S(f,u) = s(f-u)$ with s being a nonnegative, convex, and positively homogeneous functional, which is continuous with respect to weak* convergence in BV, e.g. $s(f-u) = \|f-u\|_2^2$ or $s(f-u) = \|f-u\|_1$. Thus, the authors proved the well-definedness and the convergence of the algorithm with L^1 fidelity term, which is also the iterative version of the L^1-TV model considered by Chan and Esedoglu [7], and presented denoising results in the presence of Gaussian noise.

Below we set the general iterative algorithm for image deblurring in the presence of Gaussian and Laplacian noise, with the corresponding (convex) fidelity terms.

3.1 Image Deblurring in the Presence of Noise

Let X, Y be Banach spaces, $X \subset Y$, where X is the dual of a separable Banach space. We consider the standard deblurring-denoising model given by $f = Ku+n$ where $f \in Y$ is the observed noisy data, $K : Y \to Y$ is a convolution operator with blurring kernel K (i.e. $Ku := K * u$), $u \in X$ is the ideal image we want to recover, and n is noise.

Here, we present two noise models in infinite dimension prompted by the corresponding finite dimensional models based on the conditional probability $p(f|Ku)$: the Gaussian model and the Laplace model. In finite dimensional spaces, the conditional probability $p(f|Ku)$ of the data f with given image Ku is the component of the Bayesian model that is influenced by the type of distribution of the noise (and hence the noisy data f).

Assuming $X = BV(\Omega)$ and $Y = L^p(\Omega)$ with $p = 1$ or 2, we have

$$h(u) = \frac{1}{2}\|u\|_{BV}^2 = \frac{1}{2}\left(\int_\Omega |u|dx + \int_\Omega |Du|\right)^2.$$

Here Ω is a bounded and open subset of $\mathbb{R}^N$.

In addition, we consider convex functions of the form $g(u) = S(f, Ku)$ for any $u \in X$, where S is convex with respect to u for a fixed f. Then, we propose the following general iterative algorithm to recover u:

Algorithm 4.1. Let $u_0 = 0$, $\xi_0 = 0$, $\varepsilon_0 = 0$ and iterate for $k \in \mathbb{Z}$, $k \geq 0$.

- Given (u_k, ξ_k), define $\bar{\varepsilon}_k = c\varepsilon_{k+1}$, and compute u_{k+1} as a $\bar{\varepsilon}_k - minimizer$ of the functional $g_k(u) = S(f, Ku) + c[h(u) - h(u_k) - \langle \xi_k, u - u_k\rangle + \varepsilon_k]$.
- Determine $\eta_{k+1} \in \partial_u S(f, Ku_{k+1})$ and $\xi_{k+1} \in J_{\varepsilon_{k+1}}(u_{k+1})$ such that $\eta_{k+1} + c(\xi_{k+1} - \xi_k) = 0$.

Note that we use the gradient descent method to minimize $g_k(u)$. In what follows, in practice, we assume that we work with functions $u \in W^{1,1}(\Omega) \subset BV(\Omega)$. Also, we make the functional $h(u)$ differentiable by substituting it with $h(u) \approx \frac{1}{2}\left(\int_\Omega \sqrt{\epsilon^2 + u^2}dx + \int_\Omega \sqrt{\epsilon^2 + |\nabla u|^2}dx\right)^2$, $\epsilon > 0$ small. The subgradient in this case becomes

$$\partial h(u) \approx \left(\int_\Omega \sqrt{\epsilon^2 + u^2} + \sqrt{\epsilon^2 + |\nabla u|^2}dx\right)\left[\frac{u}{\sqrt{\epsilon^2 + u^2}} - \nabla \cdot \frac{\nabla u}{\sqrt{\epsilon^2 + u^2}}\right].$$

Also, we refer to [17, Section 3.4.1] for the relation between Gateaux differentiability and $\bar{\varepsilon}_k$-minimizers. If u is an $\bar{\varepsilon}_k$-minimizer of the Gateaux-differentiable function $g_k(u)$, then we must have $\|\partial g_k(u)\| \leq \bar{\varepsilon}_k$. In practice, we use time-dependent gradient descent to approximate an $\bar{\varepsilon}_k$-minimizer u by solving $\frac{\partial u}{\partial t} = -\partial g_k(u) + \bar{\varepsilon}_k$ to steady state.

Remark 1. We can start with $u_0 = 0$, $\xi_0 = 0$, $\varepsilon_0 = 0$. Although our theory considers positive parameters ε_k in order to ensure existence of the iterates u_k, one could still initialize the algorithm with $u_0 = 0$, $\xi_0 = 0$, $\varepsilon_0 = 0$ in many situations, including the particular ones investigated below. In such cases, existence of u_1 and ξ_1 is not based on the surjectivity result employed in [3], but rather on direct analysis of the function $S(f, Ku) + ch(u)$ to be minimized.

Gaussian noise. If the data is $f = Ku + n \in Y = L^2(\Omega)$ with Gaussian distributed noise and with the expectation Ku, the conditional probability $p(f|Ku)$ is described as $p(f|Ku) \sim e^{-\frac{\|f-Ku\|_2^2}{2\sigma^2}}$, where σ^2 is the variance of the noise n. Maximizing $p(f|Ku)$ with respect to u, is equivalent to minimizing $-ln$ $p(f|Ku)$, thus we obtain a convex fidelity term to be minimized for $u \in BV(\Omega)$, $S(f, Ku) = \frac{1}{2}\|f - Ku\|_2^2$. The function $g(u) = S(f, Ku)$ satisfies the conditions enforced in Assumptions (A) in dimension one and two. Moreover, let $r(\delta) = \delta^2/2$ (see (7)).

Since such a quadratic S is Gâteaux-differentiable, its subgradient is given by $\partial_u S(f, Ku) = K^*(Ku - f)$ which leads to $\xi_{k+1} = \xi_k - \frac{1}{c}K^*(Ku_{k+1} - f)$. We propose the following numerical algorithm:

Numerical Algorithm.

I. Let $u_0 = 0$, $\xi_0 = 0$, $\varepsilon_0 = 0$ and iterate for $k \in \mathbb{Z}$, $k \geq 0$ until $\|f - Ku_{k+1}\|_2 \leq \sigma$:

- For $u = u_{k+1}$, use $\frac{\partial u}{\partial t} = K^*(f - Ku) - c[\partial h(u) - \xi_k] + c\varepsilon_{k+1}$
- For ξ_{k+1}, use $\xi_{k+1} = \xi_k + \frac{1}{c}K^*(f - Ku_{k+1})$.

In addition, following [1], we let $\xi_k = \frac{K^* v_k}{c}$ so that we have $v_{k+1} = v_k + (f - Ku_{k+1})$.

With $v_0 = 0$, since $c\xi_0 = 0 = K^*0 = K^*v_0$, we may conclude inductively that $c\xi_k \in Range(K^*)$, and hence there exists $v_k \in Y^* = L^2(\Omega)$ such that $c\xi_k = K^*v_k$. Hence, we can have the following alternative numerical algorithm:

II. Let $u_0 = 0$, $v_0 = 0$, $\varepsilon_0 = 0$ and iterate for $k \in \mathbb{Z}$, $k \geq 0$ until $\|f - Ku_{k+1}\|_2 \leq \sigma$:

- For $u = u_{k+1}$, use $\frac{\partial u}{\partial t} = K^*(f + v_k - Ku) - c\partial h(u) + c\varepsilon_{k+1}$
- For v_{k+1}, use $v_{k+1} = v_k + (f - Ku_{k+1})$.

Laplacian noise. If the data is $f = Ku + n \in Y = L^1(\Omega)$ with n being a Laplacian distributed random variable with mean zero and variance $2\sigma^2$, we

Table 1. Results using different ε_k (u_*: original image)

Gaussian noise, Shape image, $\sigma = \|f - K * u_*\|_2 = 15$ (Fig. 1)

ε_k $(k > 0)$	$a = \|f - K * u_k\|_2$, b=RMSE vs $k = 1, 2, 3, 4, 5$ with $\lambda = 0.1$					
0	a	28.8376	16.1033	**14.9578**	14.3887	13.9292
	b	38.3389	27.7436	**22.0875**	20.1510	19.4987
$\frac{1}{2^k}$	a	28.8054	16.1075	**14.9587**	14.3877	13.9277
	b	38.3168	27.7422	**22.0875**	20.1473	19.4980

Laplacian noise, Rectangles image, $\sigma = \|f - K * u_*\|_1 = 10$ (Fig. 5)

ε_k $(k > 0)$	$a = \|f - K * u_k\|_1$, b=RMSE vs $k = 1, 2, 3, 4, 5$ with $\lambda = 0.05$					
0	a	16.2208	10.3479	**9.9939**	9.9768	9.9615
	b	25.2332	6.1003	**2.4687**	2.2553	2.4489
$\frac{1}{2^k}$	a	15.8850	10.3339	**9.9935**	9.9768	9.9617
	b	24.3540	6.0411	**2.4938**	2.2898	2.4733

Table 2. Stopping criteria and comparisons, $\sigma^2 \sim$ noise variance

Noise Model	Stopping criteria	Comparison with
Gaussian	$\|f - K * u_k\|_2 \leq \sigma$	iterative algorithm using TV (13) or RO (14)
Laplacian	$\|f - K * u_k\|_1 \leq \sigma$	iterative (15) or one-step L^1-TV deblurring model

have $p(f|Ku) \sim e^{-\frac{\|f-Ku\|_1}{\sigma}}$. Then, similarly, we minimize with respect to u the quantity $-ln\ p(f|Ku)$, thus we are led to consider the convex fidelity term

$$S(f, Ku) = \int_\Omega |f - Ku|dx.$$

Moreover, let $r(\delta) = \delta$. Again, the function $g(u) = S(f, Ku)$ satisfies the conditions in Assumptions (A) in dimension one and two.

Unless $Ku \equiv f$, one can think of $\partial_u S(f, Ku) = K^* sign(Ku - f)$ almost everywhere, and moreover we have

$$\xi_{k+1} = \xi_k - \frac{1}{c} K^* sign(Ku_{k+1} - f) \quad a.e.$$

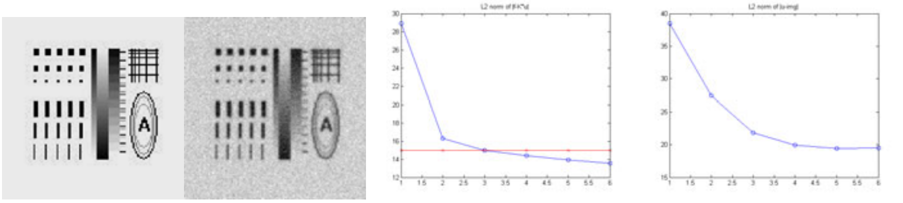
We propose the following numerical algorithm:

Numerical algorithm.

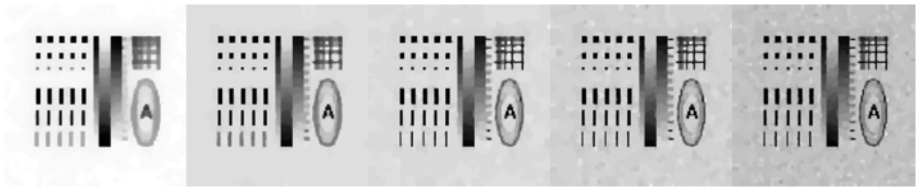
I. Let $u_0 = 0$, $\xi_0 = 0$, $\varepsilon_0 = 0$ and iterate for $k \in \mathbb{Z}$, $k \geq 0$ until $\|f - Ku_{k+1}\|_1 \leq \sigma$:

- For $u = u_{k+1}$, use $\frac{\partial u}{\partial t} = K^* sign(f - Ku) - c[\partial h(u) - \xi_k] + c\varepsilon_{k+1}$
- For ξ_{k+1}, use $\xi_{k+1} = \xi_k + \frac{1}{c} K^* sign(f - Ku_{k+1})$.

Now again letting $\xi_k = \frac{K^* v_k}{c}$, we can have $v_{k+1} = v_k + sign(f - Ku_{k+1}) \quad a.e.$



(a) original u_* (b) $f = K * u_* + n$ (c) $\|f - K * u_k\|_2$ vs k (d) $\|u_* - u_k\|_2$ vs k



(a) u_1 (b) u_2 (c) u_3 (d) u_4 (e) u_5

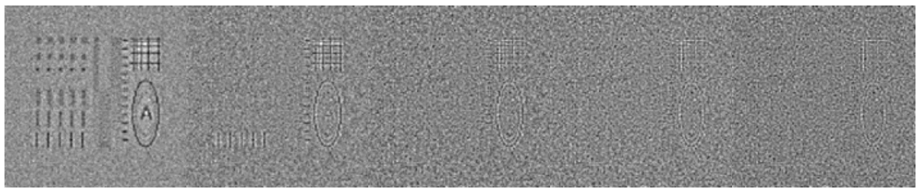


Fig. 1. Results for the Gaussian noise model obtained by the proposed iterative method. 2nd and 3rd row: recovered images u_k and the corresponding residuals $f - K * u_k$. Data: Gaussian blur kernel K with the standard deviation $\sigma_b = 0.7$, and Gaussian noise with $\sigma_n = 15$, $\lambda = 0.1$. $\|f - K * u_3\|_2 = 14.9658$. u_3 is the best recovered image (RMSE=21.8608).

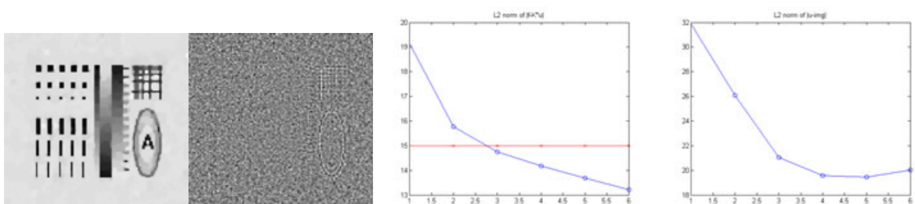
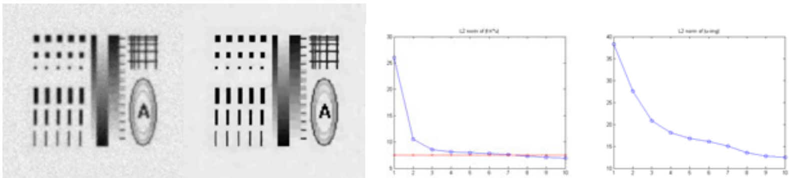


Fig. 2. Results of the iterative algorithm (13) proposed by Osher et al with the same data in Fig. 1. The best recovered image u_3 ($\|f - K * u_3\|_2 = 14.7594$, RMSE=21.0500), residual $f - K * u_3$, and energies $\|f - K * u_k\|_2$, $\|u_* - u_k\|_2$ vs k.

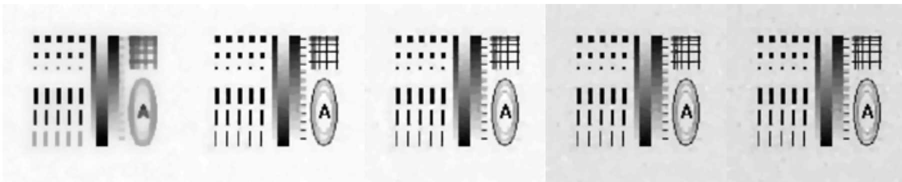
With $v_0 = 0$, since $c\xi_0 = 0 = K^*0 = K^*v_0$, we may conclude inductively that $c\xi_k \in Range(K^*)$, and hence there exists $v_k \in Y^* = L^\infty(\Omega)$ such that $c\xi_k = K^*v_k$. Hence, we have the alternative numerical algorithm:

II. Let $u_0 = 0$, $v_0 = 0$, $\varepsilon_0 = 0$ and iterate for $k \in \mathbb{Z}$, $k \geq 0$ until $\|f - Ku_{k+1}\|_1 \leq \sigma$:

– For $u = u_{k+1}$, use $\frac{\partial u}{\partial t} = K^*[sign(f - Ku) + v_k] - c\partial h(u) + c\varepsilon_{k+1}$
– For v_{k+1}, use $v_{k+1} = v_k + sign(f - Ku_{k+1})$.

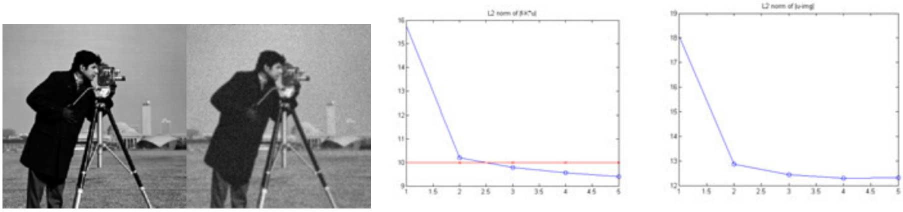


(a) $f = K * u_* + n$ (b) ROF model (c) $\|f - K * u_k\|_2$ vs k (d) $\|u_* - u_k\|_2$ vs k



(a) u_1 (b) u_2 (c) u_3 (d) u_4 (e) u_5

Fig. 3. Stopping index $k(\delta) \sim \delta^{-1}$ and comparison with RO model (RMSE=16.5007). Data: same blur kernel K and parameter $c = 0.1$ with Fig. 1, but different Gaussian noise with $\sigma_n = 7.5$. u_8 is the best recovered image (RMSE=13.5407).



(a) original u_* (b) $f = K * u_* + n$ (c) $\|f - K * u_k\|_2$ vs k (d) $\|u_* - u_k\|_2$ vs k

(a) u_1 (b) u_2 (c) u_3 (d) u_4 (e) u_5

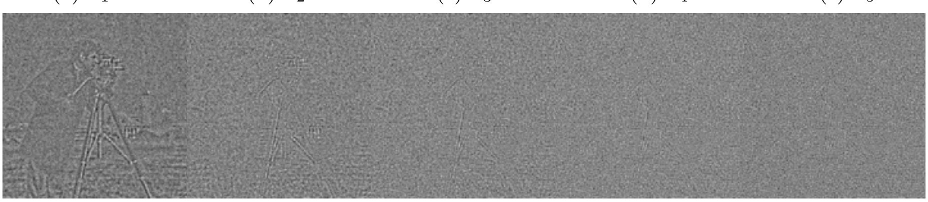
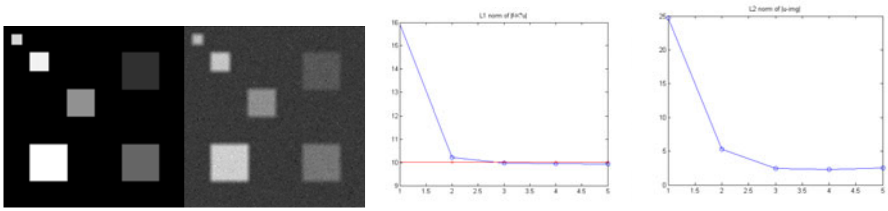


Fig. 4. Results for the Gaussian noise model obtained by the proposed method. Data: Gaussian blur kernel K with $\sigma_b = 1$, and Gaussian noise with $\sigma_n = 10$. Parameters: $c = 0.1$. u_3 is the best recovered image (RMSE=12.2217).



(a) original u_* (b) $f = K * u_* + n$ (c) $\|f - K * u_k\|_2$ vs k (d) $\|u_* - u_k\|_2$ vs k

(a) u_1 (b) u_2 (c) u_3 (d) u_4 (e) u_5

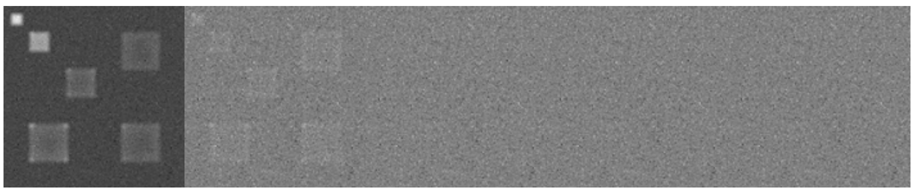
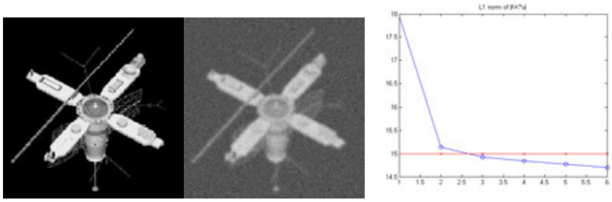


Fig. 5. Results for the Laplacian noise model obtained by the proposed method. Data: Gaussian blur kernel K with $\sigma_b = 3$, and Laplacian noise with $\sigma_n = 10$. Parameters: $\lambda = 0.05$. $\|f - K * u_3\|_1 = 9.9629$. u_3 is the best recovered image (RMSE=2.4417).

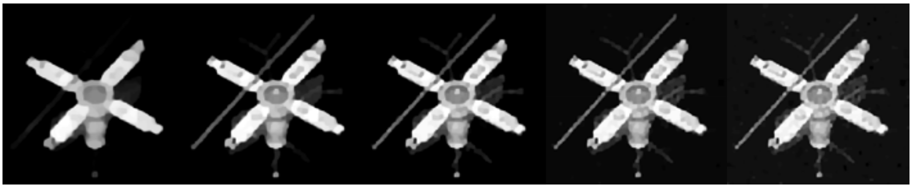
4 Numerical Results

We assume $|\Omega| = 1$. First, with fix the parameter λ,and then we test each model with different ε_k, either $\varepsilon_k = \frac{1}{2^k}$ or $\varepsilon_k = 0$. These different values of ϵ_k produce almost the same results according to the measured values in Table 1 as well as visually. Thus, in all the other examples, we numerically set $\varepsilon_k = 0$. Since our experimental results are done on artificial tests, we can compare the restored images u_k with the true image u_*.

First, we consider the residual $S(f, Ku_k)$ and the L^2 distance between iterates u_k and the original image u_*, $\|u_* - u_k\|_2$ (or root mean square error, denoted RMSE). As k increases, the image u_k recovers more details and fine scales, and eventually gets noise back. Thus, in practice, the residual $g(u_k) = S(f, Ku_k)$ keeps decreasing even when $\epsilon_k \neq 0$ (see Table 1), while $\|u_* - u_k\|_2$ has a minimum value at some k'. But, note that k' does not correspond to the optimal $k_* = \min\{k : g(u_k) = S(f, Ku_k) \leq \sigma^2\}$, i.e $k' > k_*$, which is not surprising because in the presence of blur and noise, $u_{k'}$ can have lower RMSE since $u_{k'}$ may become sharper than u_{k_*} even though $u_{k'}$ becomes noisier than u_{k_*}. However, the visual quality is also best at the optimal k_*. For example, in Fig. 1 with Gaussian noise,



(a) original u_* (b) $f = K * u_* + n$ (c) $\|f - K * u_k\|_2$ vs k (d) $\|u_* - u_k\|_2$ vs k



(a) u_1 (b) u_2 (c) u_3 (d) u_4 (e) u_5

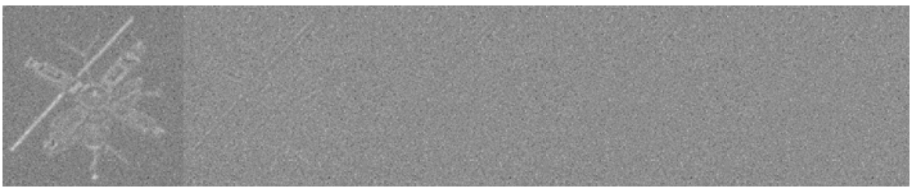
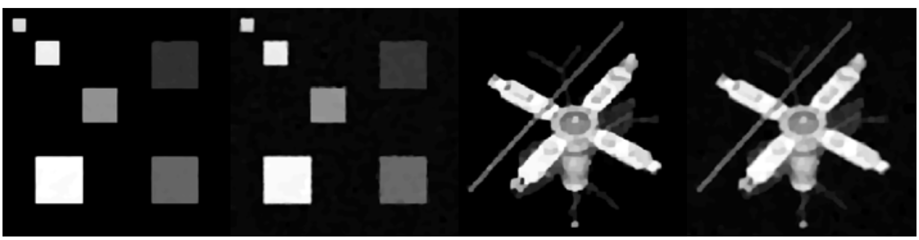


Fig. 6. Results for the Laplacian noise model obtained by the proposed method. Data: Gaussian blur kernel K with $\sigma_b = 2$, and Laplacian noise with $\sigma_n = 15$. Parameters: $\lambda = 0.02$. $\|f - K * u_3\|_1 = 14.9234$. u_3 is the best recovered image (RMSE=17.3498).



(a) u_3: 2.4417 (b) L^1-TV: 7.8411 (c) u_3: 17.3498 (d) L^1-TV: 17.8202

Fig. 7. Comparison with one-step L^1-TV [7]. (a), (c): our iterative method. (b), (d): one-step L^1-TV ($\|f - K * u\|_1$: (b) 9.8649 , (d) 14.9650). Recovered images u and RMSE values.

u_3 ($k_* = 3$) recovers the details well enough leading to the best visual quality while $\|u_* - u_k\|_2$ is still decreasing, and u_k for $k > 3$ becomes noisier. Thus the optimal k_* is a reasonable choice for Gaussian and Laplace noise models.

In Figures 1-4, we test the Gaussian noise model using L^2 fidelity term, and moreover we compare our result with the iterative algorithm (13) proposed by Osher et al. In Figures 1 and 4, u_3 recovers texture parts or details better than in the previous iterates, with less noise, while the next iterate u_4 becomes noisier.

In addition, from Figures 1 and 2, we observe that our iterative algorithm and the one from (13) proposed by Osher et al provide similar best recovered images and similar behavior. Fig. 3 verifies the a-priori property for the stopping index (12); with less noise ($\sigma_n = 7.5$), the stopping index $k_* = 8$ is twice larger than the one ($k_* = 3$) with $\sigma_n = 15$. Moreover, Fig. 3 shows that our iterative scheme provides superior result to the Rudin-Osher model [?] by recovering details or texture parts better, leading to much better RMSE.

In Figures 5-7, we show the recovered images u_k in the presence of Laplacian noise with L^1 fidelity term, and we compare our results with the iterative algorithm (15) proposed by He et al and the one-step L^1-TV model (analyzed by Chan and Esedoglu [7] when $K = I$). In Figures 5 and 6, u_k restores fine scales and becomes sharper until the optimal $k_* = 3, 2$ respectively, and u_{k_*} gives cleaner images than u_k for $k > k_*$. We have also compared with the iterative algorithm (15), which produces slightly worse result than ours. In Fig. 7, we observe that our iterative method gives cleaner and sharper images and moreover smaller RMSE than by the one-step L^1-TV model with blur.

5 Conclusion

We introduced a generalized iterative regularization method based on the norm square for image restoration models with general convex fidelity terms. We applied the proximal point method [3] using inexact Bregman distance to several ill-posed problems in image processing (image deblurring in the presence of noise). The numerical experiments indicate that for deblurring in the presence of noise, the iterative procedure yields high quality reconstructions and superior results to the one-step gradient-descent models and similar with existing iterative models. For an extended version of this work, we refer the reader to [13], where the details of the proofs are given, the data fidelity term arising from Poisson noise distribution is also considered, together with the deblurring problem using cartoon + texture representation for better texture preservation in the restoration problem.

References

1. Osher, S., Burger, M., Goldfarb, D., Xu, J., Yin, W.: An iterative regularization method for total variation based image restoration. Multiscale Modelling and Simulation 4, 460–489 (2005)
2. He, L., Osher, S., Burger, M.: Iterative total variation regularization with non-quadratic fidelity. J. Math. Imag. Vision 26, 167–184 (2005)
3. Iusem, A.N., Resmerita, E.: A proximal point method in nonreflexive banach spaces. Set-Valued and Variational Analysis 18, 109–120 (2010)
4. Tadmor, E., Nezzar, S., Vese, L.: A multiscale image representation using hierarchical (BV,L^2) decompositions. Multiscale Model. Simul. 2, 554–579 (2004)
5. Tadmor, E., Nezzar, S., Vese, L.: Multiscale hierarchical decomposition of images with applications to deblurring, denoising and segmentation. Commun. Math. Sci. 6, 281–307 (2008)

6. Rudin, L., Osher, S.J., Fatemi, E.: Nonlinear total variation based noise removal algorithms. Phys. D 60, 259–268 (1992)
7. Chan, T.F., Esedoglu, S.: Aspects of total variation regularized l1 function approximation. SIAM J. Appl. Math. 65, 1817–1837 (2005)
8. Benning, M., Burger, M.: Error estimates for variational models with non-gaussian noise. UCLA CAM Report 09-40 (2009)
9. Gossez, J.P.: Opérateurs monotones nonlineaires dans les espaces de banach nonreflexifs. J. Math. Anal. Appl. 34, 371–395 (1971)
10. Marques, A.M., Svaiter, B.: On the surjectivity properties of perturbations of maximal monotone operators in non-reflexive banach spaces (to be published)
11. Ekeland, I., Temam, R.: Convex analysis and variational problems. SIAM, Philadelphia (1999)
12. Pöschl, C.: Regularization with a similarity measure. PhD Thesis, University of Innsbruc (2008)
13. Jung, M., Resmerita, E., Vese, L.: Dual norm based iterative methods for image restoration. UCLA C.A.M. Report 09-88 (2009)
14. Rudin, L., Osher, S.: Total variation based image restoration with free local constraints. IEEE ICIP, 31–35 (1994)
15. Burger, M., Resmerita, E., He, L.: Error estimation for bregman iterations and inverse scale space methods in image restoration. Computing (81)
16. Scherzer, O., Groetsch, C.: Inverse scale space theory for inverse problems. In: Kerckhove, M. (ed.) Scale-Space 2001. LNCS, vol. 2106, pp. 317–325. Springer, Heidelberg (2001)
17. Attouch, H., Buttazzo, G., Michaille, G.: Variational analysis in Sobolev and BV spaces (2006)