mathematics

Article

On an Extension of a Spare Regularization Model

Abdellatif Moudafi

CNRS-L.I.S UMR 7296, Aix Marseille Université, Campus Universitaire de Saint-Jérôme, 13397 Marseille, France; abdellatif.moudafi@univ-amu.fr

Abstract: In this paper, we would first like to promote an interesting idea for identifying the local minimizer of a non-convex optimization problem with the global minimizer of a convex optimization one. Secondly, to give an extension of their sparse regularization model for inverting incomplete Fourier transforms introduced. Thirdly, following the same lines, to develop convergence guaranteed efficient iteration algorithm for solving the resulting nonsmooth and nonconvex optimization problem but here using applied nonlinear analysis tools. These both lead to a simplification of the proofs and to make a connection with classical works in this filed through a startling comment.

Keywords: minimization; feasibility; l_0-norm; Moreau envelope; fixed-point algorithm; forward–backward iterations; proximity mapping

MSC: 49J53; 65K10; 49M37; 90C25; 49J53

check for updates

Citation: Moudafi, A. On an Extension of a Spare Regularization Model. *Mathematics* **2023**, *11*, 4285. https://doi.org/10.3390/math11204285

Academic Editor: Salvatore Sessa

Received: 10 September 2023
Revised: 9 October 2023
Accepted: 12 October 2023
Published: 14 October 2023

Copyright: © 2023 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Compressed sensing (see, for example, [1–7]), was used to invert incomplete Fourier transforms in the context of sparse signal/image processing, and the l_1-norm was applied as a regularization for reconstructing an object from randomly selected incomplete frequency samples. Both the sparse regularization method and the compressed sensing method use the l_1-norm as a regularization to impose sparsity for the reconstructed signal under certain transforms. Because the models based on the l_1-norm are convex, they can be solved efficiently by available algorithms. Recently, the application of non-convex metrics as alternative approaches to l_1 norm has been favored, see for example, [8–11]. The main goal of this paper is to suggest the employ of the Moreau envelope associated with the l_0-norm as a regularization. Note that the sparsity of a vector is originally measured by the l_0-norm of the vector, i.e., the number of its nonzero components. However, the l_0-norm is discontinuous at the origin, which is not appropriate from a computational point of view. The envelope of the l_0-norm is a Lipschitz surrogate of the l_0-norm, which is nonconvex. Through [7], a local minimizer of a function that is the sum of a convex function and the l_0-norm can be identified with a global minimizer of a convex function which permits algorithmic development of convex optimization problems. For inverting incomplete Fourier transforms, the use of the l_0-norm allows to formulate a sparsity regularization model that can reduce artifacts and outliers in the reconstructed signal. It also allow us to design an efficient algorithm for the resulting nonconvex and nonsmooth optimization problem by means of a fixed-point formulation. Moreover, the link of this minimization problem with the related convex minimization problem will permit to prove convergence of our proposed algorithm. Furthermore, a connection with proximal/projection gradient methods is also provided by appealing to two key formulas.

2. A Sparse Regularization Model

In order to obtain to the essential information to share, we took the same paper outline as in [6] and we assume the reader has some basic knowledge of monotone operator theory and convex analysis as can be found, for example, in [12–15].

In what follows, we propose an extension of a sparse regularization model based on the Moreau envelope of the l_0-norm for inverting incomplete Fourier transforms considered in [6]. Likewise, relying on properties of the Moreau envelope of the l_0-norm, we obtain an equivalent formulation favorable for algorithmic development.

Given two Euclidean spaces of dimensions N and d, a nonempty, closed and convex subset $Q \subset \mathbb{R}^d$ and a matrix $K : \mathbb{R}^N \to \mathbb{R}^d$, we are interested in this work regarding the following problem:

$$\text{Find } y \in \mathbb{R}^N \text{ such that } Ky \in Q, \tag{1}$$

This formalism is also at the heart of the modeling of many inverse problems posed by phase recovery problems and other real-world problems, see [16] and references therein.

Our job is to describe the sparse regularization model for Equation (1) in order to obtain a sparse vector y. The l_0-norm, which counts the number of nonzero components of a vector $x \in \mathbb{R}^N$, is naturally used to measure its sparsity and is defined by

$$\|x\|_0 = \sum_{i=1}^{N} |x_i|_0,$$

with $|x_i|_0 = 1$ if $x_i \neq 0$ and $|x_i|_0 = 0$ if $x_i = 0$.

Now, let P_Q be the projection from $\mathbb{R}^N$ onto the set Q. Since the constraint is equivalent to the fact that $Ky - P_Q(Ky) = 0$, we derive the following equivalent Lagrangian formulation

$$\min_{y \in \mathbb{R}^N} \frac{1}{2}\|(I - P_Q)Ky\|^2 + \gamma\|y\|_0, \tag{2}$$

with $\gamma > 0$ a Lagrangian multiplier.

Both non-convexity and discontinuity of the l_0-norm at the origin lead to computational difficulties. To overcome these problems, we use a Lipschitz regularization of the l_0-norm by its Moreau envelope. According to [14,17], for a positive number λ, the Moreau envelope of $\|\cdot\|_0$ with index λ at $x \in \mathbb{R}^N$ is defined by

$$env_{\lambda\|\cdot\|_0}(x) = \min_{z \in \mathbb{R}^N} \left(\|z\|_0 + \frac{1}{2\lambda}\|x - z\|^2\right). \tag{3}$$

$env_{\lambda\|\cdot\|_0}$ is continuous and locally convex near the origin. Moreover, as $\lim_{\lambda \to 0} env_{\lambda\|\cdot\|_0} = \|\cdot\|_0$, $env_{\lambda\|\cdot\|_0}$ is a good approximation of $\|\cdot\|_0$ when λ is small enough. Therefore, with an appropriate choice of the parameter λ, $env_{\lambda\|\cdot\|_0}$ can be used as a measure of sparsity and allows to avoid drawbacks $\|\cdot\|_0$. For a fixed $Q \subset \mathbb{R}^d$ and for $y \in \mathbb{R}^N$, we let

$$H(y) = \frac{1}{2}\|(I - P_Q)Ky\|^2 + \gamma env_{\lambda\|\cdot\|_0}(y), \tag{4}$$

where γ is a positive parameter.

To recover a sparse vector y from (1), we now propose the sparse regularization model based on the Moreau envelope of the l_0-norm

$$\bar{y} = argmin_{y \in \mathbb{R}^N} H(y). \tag{5}$$

Since $env_{\lambda\|\cdot\|_0}$ is an approximation of $\|\cdot\|_0$, we expect that the proposed model enjoys nice properties and can be solved by efficient iteration algorithms.

As was pointed out earlier, $env_{\lambda\|\cdot\|_0}$ is an excellent sparsity promoting function. Therfore, we adopt $v = env_{\lambda\|\cdot\|_0}(y)$ in this paper.

We reformulate problem (5) to obtain a problem that is well suited and favorable for computation. Relying on definition (3) of $env_{\lambda\|\cdot\|_0}$ in problem (5) and r being fixed, we introduce the following function

$$F(x,y) = \frac{1}{2}\|(I-P_Q)Ky\|^2 + \frac{\gamma}{2\lambda}\|x-y\|^2 + \gamma\|x\|_0. \tag{6}$$

The non-convex function $F(x,y)$ is a special case of those considered in [7]. We then consider the problem

$$(\bar{x},\bar{y}) = argmin_{(x,y)\in \mathbb{R}^N\times\mathbb{R}^N} F(x,y). \tag{7}$$

Next, we prove that problems (5) and (7) are essentially equivalent. A global minimizer of any of these problems will also be called a solution of the problem. We first present a relation between $H(y)$ and $F(x,y)$. Remember that for $\lambda > 0$, the proximity operator of $\|\cdot\|_0$ at $z \in \mathbb{R}^N$ is defined by

$$prox_{\lambda\|\cdot\|_0}(z) = argmin_{x\in\mathbb{R}^N}\{\|x\|_0 + \frac{1}{2\lambda}\|x-z\|^2\}. \tag{8}$$

Clearly, if $x \in prox_{\lambda\|\cdot\|_0}(z)$, then we have that

$$env_{\lambda\|\cdot\|_0}(z) = \|x\|_0 + \frac{1}{2\lambda}\|x-z\|^2. \tag{9}$$

By relation (9), we obtain

$$H(y) = F(x,y),\ \forall\, x \in prox_{\lambda\|\cdot\|_0}(y) \text{ and } \forall\, y \in \mathbb{R}^N. \tag{10}$$

We now give a direct proof of [6], Proposition 1.

Proposition 1. *Let $\lambda > 0$ and $\gamma > 0$. A pair $(\bar{x},\bar{y})$ solves problem (7) if, and only if, $\bar{y}$ solves problem (5) with $\bar{x}$, verifying the following relation*

$$\bar{x} \in prox_{\lambda\|\cdot\|_0}(\bar{y}).$$

Proof. This follows directly from the following successive equalities.

$$\begin{aligned}
\inf_{(x,y)\in\mathbb{R}^N\times\mathbb{R}^N} F(x,y) &= \inf_{(x,y)\in\mathbb{R}^N\times\mathbb{R}^N}\left(\frac{1}{2}\|(I-P_Q)Ky\|^2 + \frac{\gamma}{2\lambda}\|x-y\|^2 + \gamma\|x\|_0\right) \\
&= \inf_{y\in\mathbb{R}^N}\left(\frac{1}{2}\|(I-P_Q)Ky\|^2 + \gamma \inf_{x\in\mathbb{R}^N}\left(\frac{1}{2\lambda}\|x-y\|^2 + \|x\|_0\right)\right). \\
&= \inf_{y\in\mathbb{R}^N}\left(\frac{1}{2}\|(I-P_Q)Ky\|^2 + \gamma env_{\lambda\|\cdot\|_0}(y)\right).
\end{aligned}$$

□

Based on the fact that problems (5) and (7) are essentially equivalent, it suffices to establish that a local minimizer of the nonconvex problem (7) is a minimizer of a convex problem on a subdomain. To that end, we first present a convex optimization problem on a proper subdomain of $\mathbb{R}^N \times \mathbb{R}^N$ related to problem (7) and recall the notion of the support of a vector $x \in \mathbb{R}^N$, denoted by $N(x)$, namely the index set on which the components of x is nonzero, that is $N(x) = \{i : x_i \neq 0\}$. Note that when the support of x in problem (7) is specified, the non-convex problem (7) reduces to a convex one. Based on this observation, we introduce a convex function by

$$G(x,y) = \frac{1}{2}\|(I-P_Q)Ky\|^2 + \frac{\gamma}{2\lambda}\|x-y\|^2,\ (x,y) \in \mathbb{R}^N \times \mathbb{R}^N. \tag{11}$$

Clearly, $F(x,y) = G(x,y) + \gamma\|x\|_0$ and $G(x,y)$ is convex and differentiable on $\mathbb{R}^N \times \mathbb{R}^N$. We define now, for a given index set N, a subspace of $\mathbb{R}^N$ by setting

$$B_N = \{x \in \mathbb{R}^N,\ N(x) \subset N\}. \tag{12}$$

B_N is convex and closed (see [6]), and we consider then the minimization problem on $B_N \times \mathbb{R}^N$ defined by

$$argmin\{G(x,y),\ (x,y) \in (x,y) \in B_N \times \mathbb{R}^N\}. \tag{13}$$

Problem (13) is convex, thanks to the convexity of both the function G and the set $B_N \times \mathbb{R}^N$. Next, we will show the equivalence between the non-convex problem (7) and the convex problem (13) with an appropriate choose of the index set N. To this end, we investigate properties of the support set of certain sequences in $\mathbb{R}^N$ and for a given index set N, we define an operator $P_{B_N} : \mathbb{R}^N \to B_N$ by

$$P_{B_N}(y) = y_i \text{ if } i \in N \text{ and } 0 \text{ otherwise}. \tag{14}$$

This operator is indeed the orthogonal projection from $\mathbb{R}^N$ onto N, (see [6] Lemma 3). A convenient identification of the proximity operator of the l_0-norm with the projection P_{B_N} and some properties of the sequence generated by $prox_{\lambda\|\cdot\|_0}$, with respect to the existence of an integer which will denote $\bar{\kappa}$, were developed in (Lemmas 4–7 together with Proposition 2 [6]), and which are still valid in our context.

Recall also the closed form formula of the proximity of l_0. For all $z \in \mathbb{R}^N$,

$$prox_{\lambda|\cdot|_0}(z) = \begin{cases} \{z_i\} & \text{if } |z_i| > \sqrt{2\lambda}; \\ \{z_i, 0\} & \text{if } |z_i| = \sqrt{2\lambda}; \\ \{0\} & \text{otherwise} \end{cases} \tag{15}$$

A connection between problems (7) and (13) is given by the following result.

Theorem 1. *$\lambda, \gamma > 0$, and $(\bar{x}, \bar{y}) \in \mathbb{R}^N \times \mathbb{R}^N$ be given. The pair $(\bar{x}, \bar{y})$ is a local minimizer of the non-convex problem (7) if, and only if, $(\bar{x}, \bar{y})$ is a minimizer of the convex problem (13) with $N := N(\bar{x})$.*

Proof. Follows directly by using ([6] Corollary 4.9), with $\phi(y) := \frac{1}{2}\|(I - P_Q)Ky\|^2, \mu := \frac{\gamma}{2\lambda}$ and $D := I$. □

Following the same lines of ([6] Propositions 1 and 3), we can identify and connect global and local minimizers of (7) with those of (5).

3. A Fixed Point Approach

We will propose an iterative method for finding a local minimizer of (7) relying on a fixed-point formulation. For all the facts we will use, we refer to [14].

Let us begin with a characterization of the convex problem (13).

Proposition 2. *Suppose $\lambda, \gamma > 0$. If $C \subset \{1, 2, \cdots, N\}$, then the problem (13) with $N := C$ has a solution and a pair $(\bar{x}, \bar{y}) \in \mathbb{R}^N \times \mathbb{R}^N$ solves (13) with $N := C$ if, and only if,*

$$\bar{x} = P_{N(\bar{x})}(\bar{y}) \quad and \quad \bar{y} = \bar{x} - \frac{\lambda}{\gamma} K^*(I - P_Q)K\bar{y}. \tag{16}$$

Proof. The existence of solutions follows by the fact that B_N is compact together with the coercivity of G with respect to the second variable. On the other hand, the optimality condition of the minimization problem

$$\min_{(x,y)\in B_N\times \mathbb{R}^N} \big(\frac{1}{2}\|(I-P_Q)Ky\|^2+\frac{\gamma}{2\lambda}\|x-y\|^2\big),$$

reads as

$$(0,0)\in \big(K^*(I-P_Q)K\bar{y}-\frac{\gamma}{\lambda}(\bar{x}-\bar{y}),\frac{\gamma}{\lambda}(\bar{x}-\bar{y})+N_{B_N}(\bar{x})\big),$$

or, equivalently,

$$\bar{y}=\bar{x}-\frac{\lambda}{\gamma}K^*(I-P_Q)K\bar{y} \text{ and } \bar{y}\in \bar{x}+\frac{\lambda}{\gamma}N_{B_N}(\bar{x}) \Leftrightarrow \bar{x}=(I+\frac{\lambda}{\gamma}N_{B_N})^{-1}(\bar{y})=P_{B_N}(\bar{y}).$$

□

Application of both Theorem 1 and Proposition 2 leads to the following characterization of a local minimizer of the problem (7).

Theorem 2. *Let $\lambda,\gamma>0$ be fixed. A pair $(\bar{x},\bar{y})\in \mathbb{R}^N\times\mathbb{R}^N$ is a local minimizer of (7) if, and only if, $(\bar{x},\bar{y})$ verifies (16).*

Let us now give a characterization of a global minimizer of (16).

Theorem 3. *Let $\lambda,\gamma>0$ be fixed. If a pair $(\bar{x},\bar{y})\in \mathbb{R}^N\times\mathbb{R}^N$ is a local minimizer of (7), then $(\bar{x},\bar{y})$ satisfies the relations*

$$\bar{x}=prox_{\lambda\|\cdot\|_0}(\bar{y}) \quad \text{and} \quad \bar{y}=\bar{x}-\frac{\lambda}{\gamma}K^*(I-P_Q)K\bar{y}. \tag{17}$$

Conversely, if a pair $(\bar{x},\bar{y})$ verifies (17), then If a pair $(\bar{x},\bar{y})$ is a local minimizer of (7).

Proof. The optimality condition of the minimization problem

$$\min_{(x,y)\in \mathbb{R}^N\times \mathbb{R}^N} \big(\frac{1}{2}\|(I-P_Q)Ky\|^2+\frac{\gamma}{2\lambda}\|x-y\|^2+\gamma\|x\|_0\big),$$

reads as

$$(0,0)\in \big(K^*(I-P_Q)K\bar{y}-\frac{\gamma}{\lambda}(\bar{x}-\bar{y}),\frac{\gamma}{\lambda}(\bar{x}-\bar{y})+\gamma\partial\|\cdot\|_0(\bar{x})\big),$$

or, equivalently,

$$\bar{y}=\bar{x}-\frac{\lambda}{\gamma}K^*(I-P_Q)K\bar{y} \text{ and } \bar{y}\in\bar{x}+\lambda\partial\|\cdot\|_0(\bar{x}) \Leftrightarrow \bar{x}\in(I+\lambda\partial\|\cdot\|_0)^{-1}(\bar{y})=prox_{\lambda\|\cdot\|_0}(\bar{y}).$$

□

In view of Theorem 3, we propose the following explicit–implicit Algorithm for solving problem (7)

$$x_{k+1}\in prox_{\lambda\|\cdot\|_0}(y_k) \quad \text{and} \quad y_{k+1}=x_{k+1}-\frac{\lambda}{\gamma}K^*(I-P_Q)Ky_{k+1} \tag{18}$$

When the projection can be computed efficiently, updates of both variables x and y in Algorithm (18) at each iteration can be efficiently implemented.

Proposition 3. *If $\lambda, \gamma > 0$, then the operator $I + \frac{\lambda}{\gamma} K^*(I - P_Q)K$ is invertible. Hence, the second part of (18) reads as*

$$y_{k+1} := J_{\frac{\lambda}{\gamma}}^{K^*(I-P_Q)K}(x_{k+1}) = (I + \frac{\lambda}{\gamma} K^*(I - P_Q)K)^{-1}(x_{k+1}). \tag{19}$$

Proof. It is well known that $K^*(I - P_Q)K$ is a maximal monotone operator (more precisely, it is an inverse strongly monotone operator, see the beginning of the proof of Theorem 4), hence $I + \frac{\lambda}{\gamma} K^*(I - P_Q)K$ is invertible by Minty Theorem and its inverse, which is the so-called resolvent operator, is single valued and firmly nonexpansive. □

4. Convergence Analysis

In this section, we investigate the convergence behavior of Algorithm (18). As in [6], after a finite number of iterations, the support of the sparse variable x_k defined by Algorithm (18) will remain unchanged, and hence solving the non-convex optimization problem (7) by algorithm (18) reduces to solving a convex optimization problem on the support.

First, we consider a function E, which is closely related to both functions F and G, and we define

$$E : \mathbb{R}^N \to \mathbb{R} \text{ at } y \in \mathbb{R}^N \text{ by} \quad E(y) := \frac{L}{2}\|(I - P_Q)Ky\|^2, \tag{20}$$

where $L := 1 + \frac{\lambda}{\gamma}$; $E(y)$ will be denoted for short by $E(y)$.

Now, we prove a convergence result of Algorithm (18).

Theorem 4. *Let $(x_k, y_k)_{k\in\mathbb{N}}$ be a sequence generated by Algorithm (18) with an initial $(x_0, y_0) \in \mathbb{R}^N \times \mathbb{R}^N$ for problem (7). If $\lambda, \gamma > 0$ are positive numbers, then we have the following properties*

1. *$F(x_{k+1}, y_{k+1}) \le F(x_k, y_k)$ for all $k \ge 0$ and the sequence $\big(F(x_k, y_k)\big)_{k\in\mathbb{N}}$ is convergent;*
2. *The sequence $(x_k, y_k)_{k\in\mathbb{N}}$ has a finite length, namely*

$$\sum_{k=0}^{+\infty} \|x_{k+1} - x_k\|^2 < +\infty, \quad \textit{and} \quad \sum_{k=0}^{+\infty} \|y_{k+1} - y_k\|^2 < +\infty, \tag{21}$$

3. *The sequence $(x_k, y_k)_{k\in\mathbb{N}}$ is asymptotically regular, that is*

$$\lim_{k\to+\infty} \|x_{k+1} - x_k\| = 0 \quad \textit{and} \quad \lim_{k\to+\infty} \|y_{k+1} - y_k\|^2 = 0.$$

Proof. The function E is differentiable with a 1-Lipschitz continuous gradient. Indeed,

$$\begin{aligned} \langle K^*(I-P_Q)K(x) - K^*(I-P_Q)K(y), x - y\rangle &= \langle (I-P_Q)K(x) - (I-P_Q)K(y), Kx - Ky\rangle \\ &\ge \|(I-P_Q)K(x) - (I-P_Q)K(y)\|^2 \\ &\ge \|K^*(I-P_Q)K(x) - K^*(I-P_Q)K(y)\|^2. \end{aligned}$$

This ensures that ∇E is 1-Lipschitz continuous.

On the other hand, since $y_{k+1} = x_{k+1} - \frac{\lambda}{\gamma} K^*(I - P_Q)Ky_{k+1}$, we can write

$$\frac{\gamma}{2\lambda}\|x_{k+1} - y_{k+1}\|^2 = \frac{\gamma}{2\lambda}\|\frac{\lambda}{\gamma} K^*(I - P_Q)Ky_{k+1}\|^2.$$

Taking into account definition of the $\|\cdot\|$ and the fact that $KK^* = I$, we obtain that

$$\frac{\gamma}{2\lambda}\|x_{k+1} - y_{k+1}\|^2 = \frac{\lambda}{2\gamma}\|(I - P_Q)Ky_{k+1}\|^2.$$

Hence,

$$F(x_{k+1}, y_{k+1}) = \frac{1}{2}(1 + \frac{\lambda}{\gamma})\|(I - P_Q)Ky_{k+1}\|^2 + \gamma\|x_{k+1}\|_0 = E(y_{k+1}) + \gamma\|x_{k+1}\|_0.$$

Using the celebrate descent Lemma, see for example [12], we can write

$$F(x_{k+1}, y_{k+1}) \leq E(y_k) + \langle \nabla E(y_k), y_{k+1} - y_k \rangle + \frac{L}{2}\|y_{k+1} - y_k\|^2 + \gamma\|x_{k+1}\|_0. \tag{22}$$

Now, we have

$$\begin{aligned}
\frac{\gamma}{\lambda + \gamma}\langle \nabla E(y_k), x_{k+1} - x_k \rangle &= \langle K^*(I - P_Q)Ky_k, y_{k+1} - y_k + \frac{\lambda}{\gamma}(K^*(I - P_Q)Ky_{k+1} - K^*(I - P_Q)Ky_k)\rangle \\
&= \langle K^*(I - P_Q)Ky_k, y_{k+1} - y_k \rangle + \frac{\lambda}{\gamma}\langle K^*(I - P_Q)Ky_k, y_{k+1} - y_k \rangle \\
&- \frac{\lambda}{\gamma}\langle (I - P_Q)Ky_k, P_QKy_{k+1} - P_QKy_k \rangle \\
&\geq (1 + \frac{\lambda}{\gamma})\langle K^*(I - P_Q)Ky_k, y_{k+1} - y_k \rangle = \langle \nabla E(y_k), y_{k+1} - y_k \rangle.
\end{aligned}$$

The Characterization of the orthogonal projection, namely

$$\langle (y - P_Q y, z - P_Q y \rangle \leq 0 \quad \forall z \in Q$$

assures that

$$\langle (I - P_Q)Ky_k, P_QKy_{k+1} - P_QKy_k \rangle \leq 0,$$

and thus

$$\langle \nabla E(y_k), y_{k+1} - y_k \rangle \leq \frac{\gamma}{\lambda + \gamma}\langle \nabla E(y_k), x_{k+1} - x_k \rangle. \tag{23}$$

Now, by using the second equation of (18) and by taking into account the fact that $I - P_Q$ is firmly nonexpansive and that $KK^* = I$, we can write

$$\begin{aligned}
\|x_{k+1} - x_k\|^2 &= \|y_{k+1} - y_k + \frac{\lambda}{\gamma}(K^*(I - P_Q)Ky_{k+1} - K^*(I - P_Q)Ky_k)\|^2 \\
&= \|y_{k+1} - y_k\|^2 + \frac{2\lambda}{\gamma}\langle K^*(I - P_Q)Ky_{k+1} - K^*(I - P_Q)Ky_k, y_{k+1} - y_k \rangle \\
&+ \frac{\lambda^2}{\gamma^2}\|K^*(I - P_Q)Ky_{k+1} - K^*(I - P_Q)Ky_k\|^2 \\
&\geq \|y_{k+1} - y_k\|^2 + \frac{2\lambda}{\gamma}\|(I - P_Q)Ky_{k+1} - (I - P_Q)Ky_k\|^2 \\
&+ \frac{\lambda^2}{\gamma^2}\|(I - P_Q)Ky_{k+1} - (I - P_Q)Ky_k\|^2.
\end{aligned}$$

This yields

$$\|x_{k+1} - x_k\|^2 \leq \|y_{k+1} - y_k\|^2 + \frac{\lambda}{\gamma}(2 + \frac{\lambda}{\gamma})\|(I - P_Q)Ky_{k+1} - (I - P_Q)Ky_k\|^2, \tag{24}$$

hence

$$\|y_{k+1} - y_k\|^2 \leq \|x_{k+1} - x_k\|^2.$$

Taking into account the fact that $0 < \frac{\lambda}{\gamma} < \frac{-1+\sqrt{5}}{2}$, we have $L = 1 + \frac{\lambda}{\gamma} < \frac{\gamma}{\lambda}$ which, combined with the last inequality, ensures that

$$L\|y_{k+1} - y_k\|^2 \le \frac{\gamma}{\lambda}\|x_{k+1} - x_k\|^2. \tag{25}$$

Combining (22), (23) and (25) yields

$$F(x_{k+1}, y_{k+1}) \le E(y_k) + \frac{\gamma}{\gamma+\lambda}\langle \nabla E(y_k), x_{k+1} - x_k\rangle + \frac{\gamma}{2\lambda}\|x_{k+1} - x_k\|^2 + \gamma\|x_{k+1}\|_0. \tag{26}$$

To prove the no-increasing of the sequence $\big(F(x_k, y_k)\big)_{k\in \boldsymbol{N}}$, we first notice that the second part of (18) can be reads as $y_k = x_k - \frac{\lambda}{\lambda+\gamma}\nabla E(y_k)$. Now, by applying definition of the proximal operator of $\lambda\|\cdot\|_0$ at y_k, we have

$$x_{k+1} \in arg\min_{x\in \boldsymbol{R}^N}\Big(\|x\|_0 + \frac{1}{2\lambda}\|x - x_k + \frac{\lambda}{\lambda+\gamma}\nabla E(y_k)\|^2\Big)$$

or, equivalently,

$$x_{k+1} \in arg\min_{x\in \boldsymbol{R}^N}\Big(\|x\|_0 + \frac{1}{2\lambda}\|x - x_k\|^2 + \frac{\lambda}{\lambda+\gamma}\langle \nabla E(y_k), x - x_k\rangle\Big),$$

which ensures that

$$\|x_{k+1}\|_0 + \frac{1}{2\lambda}\|x_{k+1} - x_k\|^2 + \frac{1}{\lambda+\gamma}\langle \nabla E(y_k), x_{k+1} - x_k\rangle \le \|x_k\|_0. \tag{27}$$

Finally, from (26) and (28), we deduce that

$$F(x_{k+1}, y_{k+1}) \le E(y_k) + \gamma\|x_k\|_0 = F(x_k, y_k).$$

It follows from (27) that

$$\gamma\|x_{k+1}\|_0 + \frac{\gamma}{2\lambda}\|x_{k+1} - x_k\|^2 + \frac{\gamma}{\lambda+\gamma}\langle \nabla E(y_k), x_{k+1} - x_k\rangle \le F(x_k, y_k).$$

In addition form (22), we obtain that

$$-E(y_k) - \langle \nabla E(y_k), y_{k+1} - y_k\rangle + \frac{L}{2}\|y_{k+1} - y_k\|^2 - \gamma\|x_{k+1}\|_0 \le -F(x_{k+1}, y_{k+1}).$$

Summing the above two inequalities and using (23), we obtain

$$\frac{\gamma}{2\lambda}\|x_{k+1} - x_k\|^2 - \frac{L}{2}\|y_{k+1} - y_k\|^2 \le F(x_k, y_k) - F(x_{k+1}, y_{k+1}).$$

This, combined with (24), yields

$$\frac{1}{2}(\frac{\gamma}{2\lambda} - L)\|y_{k+1} - y_k\|^2 + \frac{\lambda}{\gamma}(2 + \frac{\lambda}{\gamma})\|(I - P_Q)y_{k+1} - (I - P_Q)y_k\|^2 \le F(x_k, y_k) - F(x_{k+1}, y_{k+1}). \tag{28}$$

By summing the last inequality and by taking into account the fact the convergence of the sequence $(F(x_k, x_y))_{k\in \boldsymbol{N}}$ together with the fact that $\frac{\gamma}{2\lambda} - L > 0$, we deduce first that

$$\sum_{k=0}^{\infty}\|y_{k+1} - y_k\|^2 < +\infty \text{ and } \sum_{k=0}^{\infty}\|(I - P_Q)y_{k+1} - (I - P_Q)y_k\|^2 < +\infty.$$

The property $\sum_{k=0}^{\infty}\|x_{k+1} - x_k\|^2 < +\infty$ follows then from relation (24). The latter properties ensures clearly the asymptotic regularity of the sequence $(x_k, y_k)_{k\in \boldsymbol{N}}$. $\square$

As in ([6] Lemma 12), in our setting, we also have that the invariant support set of the sequence defined by Algorithm (18) exists for the nonconvex problem (7). Now, let us prove, more directly than in [6], the convergence of the sequence $(x_k, y_k)_{k\in\mathbb{N}}$ generated by (18) relying on averaged operators and Krasnoselskii–Mann Theorem. Averaged mappings are convenient in studying the convergence sequences generated by iterative algorithms for fixed-point problems thanks to the following celebrate theorem, see for example [12,13].

Theorem 5 (Krasnoselskii–Mann Theorem)**.** *Let $M : \mathbb{R}^N \to \mathbb{R}^N$ be averaged and assume $FixM \neq \emptyset$. Then, for any starting point x_0, the sequence $\{M^k x_0\}$ converges weakly to a fixed-point of M.*

Recall also the definitions of nonexpansive and averaged operators, which appear naturally when using iterative algorithms for solving fixed-point problems and which are commonly encountered in the literature; see, for instance, [13]. A mapping $T : \mathbb{R}^N \to \mathbb{R}^N$ is said to be nonexpansive if, for all $x, y \in \mathbb{R}^N$, $\|Tx - Ty\| \leq \|x - y\|$, firmly nonexpansive if $2T - I$ is nonexpansive, or equivalently $\langle Tx - Ty, x - y\rangle \geq \|Tx - Ty\|^2$, for all $x, y \in \mathbb{R}^N$. It is well known that T is firmly nonexpansive if and only if T can be written as $T = \frac{1}{2}(I + S)$, where $S : \mathbb{R}^N \to \mathbb{R}^N$ is nonexpansive. Recall also that mapping $T : \mathbb{R}^N \to \mathbb{R}^N$ is said to be averaged if it can be expressed as $T = (1 - \alpha)I + \alpha S$ with $S : \mathbb{R}^N \to \mathbb{R}^N$ a nonexpansive mapping and $\alpha \in [0, 1]$. Thus, firmly nonexpansive mappings (e.g., projections on convex convex and nonempty subsets and resolvent of maximal monotone operators) are averaged mappings.

Mimicking the analysis in [6], $(\bar{x}, \bar{y})$ is a solution of (13) if, and if $(\bar{x}, \bar{y})$ satisfies (16) and thus $\bar{x}$ is verified as

$$\bar{x} = P_{B_N} \circ J_{\frac{\lambda}{\gamma}}^{K^*(I-P_Q)K}(\bar{x}). \tag{29}$$

Similarly, using the same arguments, we drive that (x_k, y_k) generated by (18) leads to

$$x_{k+1} = P_{B_N}(y_k) \quad \forall k \geq \bar{\kappa}$$

and thus for all $k \geq \bar{\kappa}$, it satisfies

$$x_{k+1} = P_{B_N}(y_k) \quad \text{and} \quad y_{k+1} = x_{k+1} - \frac{\lambda}{\gamma} K^*(I - P_Q)K y_{k+1}. \tag{30}$$

Consequently,

$$\bar{x} = P_{B_N} \circ J_{\frac{\lambda}{\gamma}}^{K^*(I-P_Q)K}(\bar{x}), \tag{31}$$

and

$$x_{k+1} = P_{B_N} \circ J_{\frac{\lambda}{\gamma}}^{K^*(I-P_Q)K}(x_k). \tag{32}$$

It is well-known that firmly nonexpansive mappings (including orthogonal projections on closed convex nonempty subsets and resolvent mappings of maximal monotone operators) are averaged operators. In view of the fact that the composite of finitely many averaged mappings is averaged, see for instance [12], and by applying Krasnoselskii–Mann Theorem, we deduce the convergence of the subsequence $(x_k)_{k\geq\bar{\kappa}}$ to a solution $\bar{x}$ of (29). $\bar{x} \in B_N$, since $(x_k)_{k\geq\bar{\kappa}} \subset B_N$, which is closed, and we also have $N = N(x_{\bar{\kappa}}) = N(\bar{x})$ in view of ([6] Lemma 7). In addition, since the resolvent of a maximal monotone operator is nonexpansive, for all $m > n > \bar{\kappa}$, we can write that

$$\|y_m - y_n\| = \|J_{\frac{\lambda}{\gamma}}^{K^*(I-P_Q)K}(x_m) - J_{\frac{\lambda}{\gamma}}^{K^*(I-P_Q)K}(x_n)\| \leq \|x_m - x_n\|.$$

$(x_k)_{k\geq\bar{\kappa}}$ being convergent, it is a Cauchy sequence and thus it is also the case of the subsequence $(y_k)_{k\geq\bar{\kappa}}$ which, in turn, converges to some limit $\bar{y}$. Now, by passing to the limit in

$y_{k+1} = J_{\frac{\lambda}{\gamma}}^{K^*(I-P_Q)K}(x_k)$ and taking into account the continuity of the resolvent, we obtain $\bar{y} = J_{\frac{\lambda}{\gamma}}^{K^*(I-P_Q)K}(\bar{x})$, where $\bar{x}$ is a solution of (16) which ensures that $(\bar{x}, \bar{y})$ is a solution of (13) with $N = N(\bar{x})$.

Based on the above, these lead to the following Theorem.

Theorem 6. *Let $(x_k, y_k) \in \mathbb{R}^N \times \mathbb{R}^N$ a sequence generated by (18) form an initial point (x_0, y_0). If $\lambda, \gamma > 0$ are chosen such that $0 < \frac{\lambda}{\gamma} < \frac{1+\sqrt{5}}{2}$, then (x_k, y_k) converges to a local minimizer $(\bar{x}, \bar{y})$ of (7). Moreover, $\big(F(x_k, y_k)\big)_{k\in\mathbb{N}}$ is a convergent sequence and if, in addition $|\bar{y}_j| \neq \sqrt{2\beta}$ for all $j \in N(\bar{y})$, then $\bar{y}$ is a local minimizer of (5).*

Finally, let us point out that, since $KK^* = I$, the fixed point iteration in (18) turns into

$$y_{k+1} = x_{k+1} - \frac{\lambda}{2\gamma} K^*(I - P_Q)(Kx_{k+1}). \tag{33}$$

In particular, when $Q = \{r\}$, this reduces to

$$y_{k+1} = x_{k+1} - \frac{\lambda}{2\gamma} K^*(Kx_{k+1} - r).$$

Indeed, as $KK^* = I$,

$$\begin{aligned} J_{\frac{\lambda}{\gamma}}^{K^*(I-P_Q)K}(x_{k+1}) &= x_{k+1} - \frac{\lambda}{\gamma} K^*(I - P_Q)_1(Kx_{k+1}) = x_{k+1} - \frac{\lambda}{\gamma} K^*((\partial i_Q)_1)_1(Kx_{k+1}) \\ &= x_{k+1} - \frac{\lambda}{\gamma} K^*(\partial i_Q)_2(Kx_{k+1}) = x_{k+1} - \frac{\lambda}{2\gamma} K^*(I - P_Q)(Kx_{k+1}). \end{aligned}$$

We used both the fact that for any maximal monotone operator A and $\nu > 0$, we have

$$J_\nu^{K^*AK}(x) = x - \nu K^* A_1(Kx),$$

$A_1 = I - J_1^A$ being the Yosida operator of A with parameter 1, and that for all $\nu, \mu > 0$, we have

$$(A_\nu)_\mu = A_{\nu+\mu}.$$

We used also the fact that the Yosida operator of the subdifferential of the indicator function of Q with parameter 2 (the latter is nothing but the the normal cone to Q) is exactly $\frac{I-P_Q}{2}$.

Therefore, the proposed algorithms are nothing else than a Proximal Gradient and a Projection Gradient Algorithms. Nevertheless, the crucial idea (namely, the support of the sparse main variable generated by the Algorithm remains unchanged after a finite number of iterations) that permits to locate a local minimizer of the nonconvex optimization problem with a global minimizer of a convex optimization one deserves a great interest.

Clearly, the analysis developed here can be extended to split feasibility problems, namely

$$\text{Find } y \in C \text{ such that } Ky \in Q, \tag{34}$$

with $C \subset \mathbb{R}^N$, $Q \subset \mathbb{R}^d$ being two closed, convex subsets and $K : \mathbb{R}^N \to \mathbb{R}^d$ a given matrix. Since the sum of a maximal monotone operator (the normal cone to C) and the monotone Lipschitz one (the operator $\frac{\lambda}{2\gamma}K^*(I - P_Q)K$) is still maximal monotone [14], this can be naturally extended to the following general minimization problem by means of its regularized version, i.e.,

$$\min_{y\in\mathbb{R}^N} \big(f(y) + g(Ky) + \gamma\|y\|_0\big) \quad \textit{through} \quad \min_{y\in\mathbb{R}^N} \big(f(y) + g_\nu(Ky) + \gamma(\|y\|_0)_\lambda\big),$$

with f, g being two proper, convex, lower semicontinuous functions defined on $\mathbb{R}^N, \mathbb{R}^d$, respectively, and $\nu, \lambda > 0$. The proximity map of g and the subdifferential of f will act as the projection on the set Q and the normal cone to C, respectively, since they share both the same properties.

5. Conclusions

Based on an interesting idea developed in [6], which leads to identifying a local minimizer of a nonconvex minimization problem with a global optimizer of a convex optimization one, we provide an extension of the sparse regularization model for inverting incomplete Fourier transforms. Next, we propose an efficient convergence guaranteed iteration algorithm for solving the resulting non-convex and non-smooth optimization problem. The fixed-point approach is preferred, as it enables us to develop efficient algorithms with guaranteed convergence. Combined with applied nonlinear analysis tools, this leads both to a simplification of the proofs and to make a connection with classical works as split convex feasibility problems. With this generalization, the proposed approach may be applicable to other real world applications such as inverse problem of intensity-modulated radiation therapy (IMRT) treatment planning [18]. It can be applied equally to cooperative wireless sensor network positioning [19] or adaptive image denoising [20]. Thus, the proposed method is expected to work efficiently for problems that can be reformulated as sparse optimization and convex feasibility problems. We will consider this as a future project for numerical applications as well as other potential extensions, for example, in a non-convex framework. These in turn will pave the way for other applications in the real world, which is the case, for example, in [21].

Funding: This research received no external funding.

Data Availability Statement: Data sharing not applicable. No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Acknowledgments: We would like to thank the anonymous reviewers for their meticulous reading of the paper and for their pertinent comments.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Byrne, C. A unified treatment of some iterative algorithms in signal processing and image reconstruction. *Inverse Probl.* **2004**, *20*, 103–120. [CrossRef]
2. Candes, E.J.; Romberg, J.; Tao, T. Stable signal recovery from incomplete and inaccurate measurements. *Commun. Pure Appl. Math.* **2006**, *59*, 1207–1223. [CrossRef]
3. Donoho, D.L. Compressed sensing. *IEEE Trans. Inf. Theory* **2006**, *52*, 1289–1306. [CrossRef]
4. Shen, F.C.L.; Xu, Y.; Zeng, X. The Moreau envelope approach for the L1/TV image denoising model. *Inverse Probl. Imaging* **2014**, *8*, 53–77.
5. Wu, T.; Shen, L.; Xu, Y. Fixed-point proximity algorithms solving an incomplete Fourier transform model for seismic wavefield modeling. *J. Comput. Appl. Math.* **2021**, *385*, 113208. [CrossRef]
6. Wu, T.; Xu, Y. Inverting Incomplete Fourier Transforms by a Sparse Regularization Model and Applications in Seismic Wavefileld Modeling. *J. Sci. Comput.* **2022**, *92*, 48. [CrossRef]
7. Xu, Y. Sparse regularization with the l_0-norm. *Anal. Appl.* **2023**, *21*, 901–929. [CrossRef]
8. Cao, W.; Xu, H.K. $l_1 - l_p$ DC regularization method for compressed sensing. *J. Nonlinear Convex Anal.* **2020**, *9*, 1889–1901.
9. Lou, Y.; Yan, M. Fast $l_1 - l_2$ Minimization via a proximal operator. *J. Sci. Comput.* **2017**, *74*, 767–785. [CrossRef]
10. Moudafi, A.; Gibali, A. $l_1 - l_2$ regularization of split feasibility problems. *Numer. Algorithms* **2018**, *78*, 739–757. [CrossRef]
11. Moudafi, A.; Xu, H.-K. A DC Regularization of Split Minimization Problems. *Appl. Anal. Optim.* **2018**, *2*, 285–297.
12. Bauschke, H.H.; Combettes, P.L. *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, 2nd ed.; Springer: Berlin/Heidelberg, Germany, 2017.
13. Cegielski, A. Iterative Methods for Fixed Point Problems in Hilbert Spaces. In *Lecture Notes in Mathematics*; LNM; Springer: Berlin/Heidelberg, Germany, 2013; Volume 2057.
14. Rockafellar, R.T.; Wets, R.J-B. *Variational Analysis*; Springer: Berlin/Heidelberg, Germany, 1998.
15. Yamada, I.; Yukawa, M.; Yamagishi, M. Minimizing the Moreau Envelope of Nonsmooth Convex Functions over the Fixed Point Set of Certain Quasi-Nonexpansive Mappings. In *Fixed-Point Algorithms for Inverse Problems in Science and Engineering*; Springer: New York, NY, USA, 2011; pp. 345–390.

16. Censor, Y.; Bortfeld, T.; Martin, B.; Trofimov, A. A unified approach for inversion problems in intensity-modulated radiation therapy. *Phys. Med. Biol.* **2006**, *51*, 2353–2365. [CrossRef] [PubMed]
17. Moreau, J.-J. Fonctions convexes duales et points proximaux dans un espace hilbertien. *Comptes Rendus Acad. Sci. Paris Ser. A Math.* **1962**, *255*, 1897–2899.
18. Censor, Y.; Elfving, T.; Kopf, N.; Bortfeld, T. The multiple-sets split feasibility problem and its applications for inverse problems. *Inverse Probl.* **2005**, *21*, 2071–2084. [CrossRef]
19. Gholami, M.R.; Tetruashvili, L.; Strom, E.G.; Censor, Y. Cooperative Wireless Sensor Network Positioning via Implicit Convex Feasibility. *IEEE Trans. Signal Process.* **2013**, *61*, 5830–5840. [CrossRef]
20. Censor, Y.; Gibali, A.; Lenzen, F.; Schnorr, C. The Implicit Convex Feasibility Problem and Its Application to Adaptive Image Denoising. *J. Comp. Math.* **2016**, *34*, 610–625. [CrossRef]
21. Brooke, M.; Censor, Y.; Gibali, A. Dynamic string-averaging CQ-methods for the split feasibility problem with percentage violation constraints arising in radiation therapy treatment planning. *Int. Trans. Oper. Res.* **2020**, *30*, 181–205. [CrossRef] [PubMed]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.