SIGNAL RECOVERY FROM INCONSISTENT NONLINEAR OBSERVATIONS

Patrick L. Combettes and Zev C. Woodstock

North Carolina State University, Department of Mathematics, Raleigh, NC 27695-8205, USA

ABSTRACT

We show that many nonlinear observation models in signal recovery can be represented using firmly nonexpansive operators. To address problems with inaccurate measurements, we propose solving a variational inequality relaxation which is guaranteed to possess solutions under mild conditions and which coincides with the original problem if it happens to be consistent. We then present an efficient algorithm for its solution, as well as numerical applications in signal and image recovery, including an experimental operator-theoretic method of promoting sparsity.

***Index Terms*—** Firmly nonexpansive operator, inconsistent nonlinear observations, signal recovery, variational inequality.

1. INTRODUCTION

Data formation models and prior knowledge in signal recovery often come in the form of equations which link an ideal solution $\overline{x}$ to a prescribed value p, say $W\overline{x} = p$, where W is an operator between Euclidean spaces $\mathcal{H}$ and $\mathcal{G}$. This equation can model an observation of $\overline{x}$ obtained by a sensing device W, or a known property of $\overline{x}$. For instance, in the classical work [21], W is a projector onto a vector subspace D of $\mathcal{H}$, and the goal is to find a point x in a vector subspace C of $\mathcal{H}$ given $p = \operatorname{proj}_D x$. We generalize this linear framework as follows.

Problem 1 Let I be a nonempty finite set and let C be a nonempty closed convex subset of a Euclidean space $\mathcal{H}$. For every $i \in I$, let $\mathcal{G}_i$ be a Euclidean space, let $p_i \in \mathcal{G}_i$, let $L_i \colon \mathcal{H} \to \mathcal{G}_i$ be a nonzero linear operator, and let $F_i \colon \mathcal{G}_i \to \mathcal{G}_i$ be *firmly nonexpansive*, i.e.,

$$(\forall (y, z) \in \mathcal{G}_i^2) \quad \langle y - z \mid F_i y - F_i z \rangle \geqslant \|F_i y - F_i z\|^2. \tag{1}$$

The task is to

$$\text{find } x \in C \text{ such that } (\forall i \in I)\ F_i(L_i x) = p_i. \tag{2}$$

Adopting firmly nonexpansive operators in our model allows for the enforcement of complex nonlinear equations such as the following (further instances of firmly nonexpansive operators are provided in Section 2).

Example 2 For every $j \in \{1, \ldots, m\}$, let $\rho_j > 0$ and let G_j be a Euclidean space. The shrinkage operator on $\mathcal{G} = \bigtimes_{j \in J} \mathsf{G}_j$, given by

$$F \colon (\mathsf{y}_j)_{1 \leqslant j \leqslant m} \mapsto \left(\left(1 - \frac{\rho_j}{\max\{\|\mathsf{y}_j\|, \rho_j\}} \right) \mathsf{y}_j \right)_{1 \leqslant j \leqslant m}, \tag{3}$$

is used to sparsify signals across groups of indices [22]. Since F is firmly nonexpansive, a signal $x \in \mathcal{H}$ can be recovered from its sparsified version $p \in \mathcal{G}$, i.e., $F(Lx) = p$, where $L \colon \mathcal{H} \to \mathcal{G}$ is a decomposition operator.

The algorithms of [9, 10] are guaranteed to yield a solution to Problem 1, provided an exact solution exists. However, in many situations, the operators $(F_i, L_i)_{i \in I}$ may not be known perfectly or the prescribed values $(p_i)_{i \in I}$ may be corrupted, meaning that a solution to Problem 1 may not exist, in which case the algorithms from [9, 10] are known to diverge. To handle this situation, we propose a tractable relaxed problem which is guaranteed to possess solutions under mild conditions.

In general, there is no tractable relaxation of Problem 1 as a convex minimization problem. For instance, consider the least-squares relaxation

$$\underset{x \in C}{\text{minimize}} \ f(x), \text{ where } f \colon x \mapsto \textstyle\sum_{i \in I} \|F_i(L_i x) - p_i\|^2. \tag{4}$$

In our setting, f is typically nonconvex and nondifferentiable [2, 13], which makes it impossible to guarantee the construction of solutions. Another plausible approach would be, for every $i \in I$, to introduce the closed convex set $D_i = \big\{ y \in \mathcal{G}_i \mid F_i y_i = p_i \big\}$ and solve

$$\underset{x \in C}{\text{minimize}} \ g(x), \quad \text{where} \quad g \colon x \mapsto \textstyle\sum_{i \in I} d_{D_i}(L_i x)^2, \tag{5}$$

where d_{D_i} is the distance function to D_i. However, (5) is intractable because evaluating the operators $(\operatorname{proj}_{D_i})_{i \in I}$ is either impossible or computationally expensive, and therefore we cannot evaluate g or its gradient. We circumvent these issues with the following variational inequality relaxation.

Problem 3 Consider the setting of Problem 1 and let $\{\omega_i\}_{i \in I} \subset]0, 1]$ satisfy $\sum_{i \in I} \omega_i = 1$. The task is to

$$\text{find } x \in C \text{ such that } (\forall y \in C)\ \textstyle\sum_{i \in I} \omega_i \langle L_i(y - x) \mid F_i(L_i x) - p_i \rangle \geqslant 0. \tag{6}$$

In Section 2, we discuss background and illustrate the flexibility and the breadth the proposed firmly nonexpansive model. Section 3 establishes that Problem 3 is an appropriate relaxed formulation of Problem 1, provides guarantees for when Problem 3 possesses solutions, and presents a block-iterative algorithm to solve Problem 3. Numerical demonstrations, including an experiment for promoting sparsity within Problem 3, are in Section 4.

2. FIRMLY NONEXPANSIVE MODELING

Notation 4 Let C be a nonempty closed convex subset of $\mathcal{H}$. We denote the *projection operator onto* C by proj_C and its *normal cone*

The work of P. L. Combettes was supported by the National Science Foundation under grant CCF-1715671 and the work of Z. C. Woodstock by the National Science Foundation under grant DGE-1746939.

operator by N_C. We denote the *range* of an operator A by $\operatorname{ran} A$ and the *adjoint* of a linear operator L by L^*. We denote the discrete Fourier transform of $x \in \mathbb{R}^N$ via $\widehat{x}$ and its inverse via $x^\vee$.

We first observe that many nonlinearities appearing in signal processing can be modeled using firmly nonexpansive operators (see [10] for details and complements).

Example 5 Let C be a nonempty closed convex subset of $\mathcal{H}$. The projection operator proj_C is firmly nonexpansive, and operators in this class model hard clipping [1, 13, 19], distortion [17, Section 10.4.1], isotonic regression [3], and downsampling [15].

Example 6 Let $f\colon \mathcal{H} \to]-\infty,+\infty]$ be proper, lower semicontinuous, and convex. The proximity operator $\operatorname{prox}_f\colon x \mapsto \operatorname{argmin}_{y\in\mathcal{H}} f(x) + \|y-x\|^2/2$ is a firmly nonexpansive operator which models soft clipping sensors [2, 14], activation functions in neural networks [8], and shrinkage operators [22].

Example 7 Let $\varnothing \neq D \subset \mathcal{G}$ be closed and convex, set $p = 0 \in \mathcal{G}$, and set $F = \operatorname{Id} - \operatorname{proj}_D$. For every $y \in \mathcal{G}$, $Fy = p$ if and only if $y \in D$. Since F is firmly nonexpansive, convex set constraints [7] are readily enforced in our framework.

Next, we demonstrate that nonlinear equations involving discontinuous operators may be equivalently represented using a firmly nonexpansive operator.

Definition 8 Let $Q\colon \mathcal{G} \to \mathcal{G}$ and let $q \in \operatorname{ran} Q$. Then (Q, q) is *proxifiable* if there exists a firmly nonexpansive operator $F\colon \mathcal{G} \to \mathcal{G}$ and $p \in \operatorname{ran} F$ such that $(\forall y \in \mathcal{G})\ Qy = q \Leftrightarrow Fy = p$. In this case (F, p) is a *proxification* of (Q, q).

Example 9 The *hard thresholder*

$$\operatorname{hard}_\gamma : \eta \mapsto [\,\eta,\ \text{if}\ |\eta| > \gamma; 0,\ \text{if}\ |\eta| \leqslant \gamma\,], \tag{7}$$

is discontinuous and used to model sensing devices and compression schemes [4, 12, 18]. Letting $q \in \operatorname{ran} \operatorname{hard}_\gamma$, we acquire a proxification of $(\operatorname{hard}_\gamma, q)$ by use of $\mathsf{S}\colon \eta \mapsto \eta - \gamma \operatorname{sign}(\eta)$. Since (we use the convention $\operatorname{sign}(0) = 0$)

$$\mathsf{S} \circ \operatorname{hard}_\gamma = \operatorname{soft}_\gamma : \eta \mapsto \operatorname{sign}(\eta)\max\{|\eta| - \gamma, 0\} \tag{8}$$

is the firmly nonexpansive *soft thresholder* on $[-\gamma, \gamma]$, $(F, p) = (\operatorname{soft}_\gamma, \mathsf{S}q)$ is a proxification of $(\operatorname{hard}_\gamma, q)$.

Extensions of Example 9 such as group hard thresholding are proxifiable using similar strategies [11].

Example 10 ([11]) Let $\mathcal{G} = \mathbb{R}^{N\times M}$, set $s = \min\{N, M\}$, and let us denote the singular value decomposition of $y \in \mathcal{G}$ by $y = U_y \operatorname{diag}(\sigma_1(y), \ldots, \sigma_s(y)) V_y^\top$. Let $\rho \in]0, +\infty[$, let hard_ρ be as in (7), set $\mathsf{S}\colon \eta \mapsto \eta - \rho \operatorname{sign}(\eta)$, and set

$$\begin{cases} Q\colon y \mapsto U_y \operatorname{diag}\Big(\operatorname{hard}_\rho\big(\sigma_1(y)\big), \ldots, \operatorname{hard}_\rho\big(\sigma_s(y)\big)\Big) V_y^\top \\ S\colon y \mapsto U_y \operatorname{diag}\Big(\mathsf{S}\big(\sigma_1(y)\big), \ldots, \mathsf{S}\big(\sigma_s(y)\big)\Big) V_y^\top. \end{cases} \tag{9}$$

Let $q \in \operatorname{ran} Q$, and set $F = S \circ Q$ and $p = Sq$. Then (F, p) is a proxification of (Q, q). The operator Q is used in image compression to produce low rank approximations, and the associated firmly nonexpansive operator F soft-thresholds singular values at level ρ.

Remark 11 In the setting of Example 10, consider the compression technique performed by the nonconvex projection operator [5] $R\colon y \mapsto U_y \operatorname{diag}\big(\sigma_1(y), \ldots, \sigma_r(y), 0, \ldots, 0\big) V_y^\top$, which truncates singular values at a given rank $r \in \{1, \ldots, s-1\}$. Let $y \in \mathcal{G}$ and set $q = Ry$. Then, for every $\rho \in]\sigma_{r+1}(y), \sigma_r(y)[$, $Qy = q$. Therefore, knowledge of the low rank approximation q to y can be exploited in our framework by proxifying (Q, q) using Example 10. Note that ρ can be estimated from q since one has access to $\sigma_r(q) = \sigma_r(y)$.

3. ANALYSIS AND ALGORITHM

We first show that Problem 3 is a valid relaxation of Problem 1.

Proposition 12 ([11]) *Suppose that the set of solutions to Problem 1 is nonempty. Then it coincides with the set of solutions to Problem 3.*

Proposition 13 ([11]) *Problem 3 admits a solution in each of the following instances:*

(i) $\sum_{i\in I} \omega_i L_i^* p_i \in \operatorname{ran}(N_C + \sum_{i\in I} \omega_i L_i^* \circ F_i \circ L_i)$.

(ii) *C is bounded.*

(iii) $\operatorname{ran} N_C + \sum_{i\in I} \omega_i L_i^*(\operatorname{ran} F_i) = \mathcal{H}$.

(iv) *For some $i \in I$, L_i^* is surjective and one of the following holds:*

  (a) *F_i is surjective.*

  (b) *$\|F_i(y)\| \to +\infty$ as $\|y\| \to +\infty$.*

  (c) *$\operatorname{ran}(\operatorname{Id} - F_i)$ is bounded.*

  (d) *$F_i = \operatorname{prox}_{g_i}$, for some convex function $g_i\colon \mathcal{G}_i \to \mathbb{R}$.*

Proposition 14 ([11]) *Consider the setting of Problem 3 under the assumption that it has a solution. Let K be a strictly positive integer and let $(I_n)_{n\in\mathbb{N}}$ be a sequence of nonempty subsets of I such that*

$$(\forall n \in \mathbb{N}) \quad \bigcup_{k=0}^{K-1} I_{n+k} = I. \tag{10}$$

Let $x_0 \in \mathcal{H}$, let $\gamma \in]0, 2[$, and, for every $i \in I$, let $t_{i,-1} \in \mathcal{H}$ and set $\gamma_i = \gamma/\|L_i\|^2$. Iterate

$$\begin{array}{l} \text{for } n = 0, 1, \ldots \\ \quad \text{for every } i \in I_n \\ \quad\quad t_{i,n} = x_n - \gamma_i L_i^*\big(F_i(L_i x_n) - p_i\big) \\ \quad \text{for every } i \in I \smallsetminus I_n \\ \quad\quad t_{i,n} = t_{i,n-1} \\ \quad x_{n+1} = \operatorname{proj}_C\big(\sum_{i=1}^m \omega_i t_{i,n}\big). \end{array} \tag{11}$$

Then $(x_n)_{n\in\mathbb{N}}$ converges to a solution to Problem 3.

The algorithm (11) is block-iterative, which allows the user to save time-per-iteration and leverage parallelism by appropriate selection of the blocks $(I_n)_{n\in\mathbb{N}}$. The condition (10) is mild, since it imposes that every element of I be activated at least once every K iterations. Note that in large-scale scenarios in which all data cannot be simultaneously loaded into memory, block-activation strategies make algorithm (11) implementable. Also, algorithm 11 is more tractable than solving (4) or (5), since it converges from any initial point and it does not require evaluating the operators $(\operatorname{proj}_{D_i})_{i\in I}$.

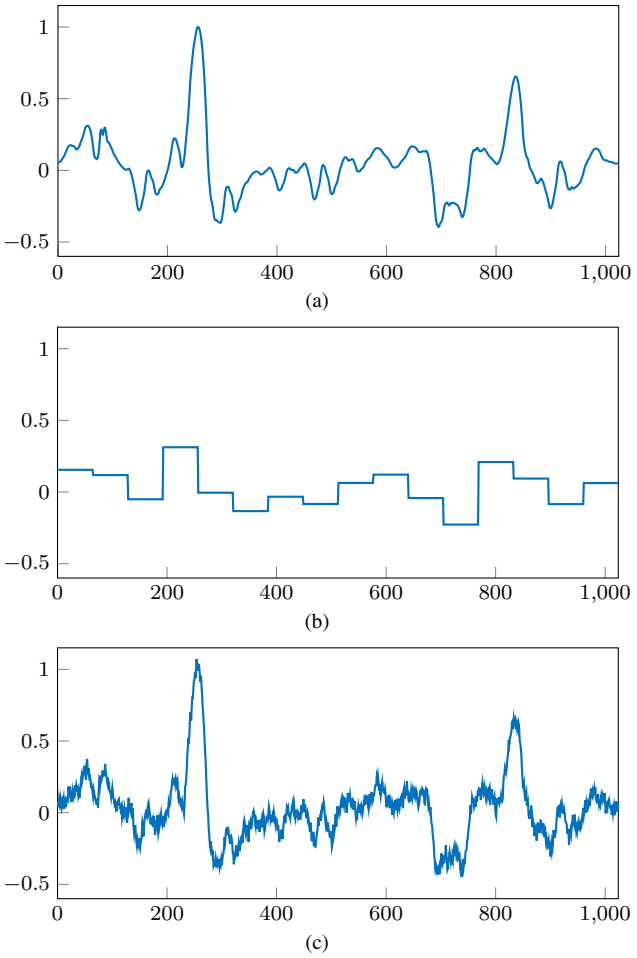

Fig. 1. Experiment of Section 4.1: (a): Original signal $\overline{x}$. (b): Piecewise constant approximation p_1. (c): Recovered signal.

4. NUMERICAL EXPERIMENTS

Remark 15 In minimization-based signal recovery, it is customary to add a function g to the objective to promote desirable properties in the solutions. A prominent example is the promotion of sparsity through the addition of a penalty such as the ℓ^1 norm in $\mathbb{R}^N$ [6, 20]. In the more general setting of Problem 3, this can be mimicked by adding the prescription $Fy = 0$, where $F = \mathrm{Id} - \mathrm{prox}_g$. Note that exact satisfaction of the equality $Fy = 0$ means that y minimizes g. In general, when incorporated to Problem 3, the pair $(F, p) = (\mathrm{Id} - \mathrm{prox}_g, 0)$ is intended to promote the properties g would have promoted in a standard minimization problem. We investigate in Section 4.3 this technique to encourage sparsity in $\mathbb{R}^N$ through the use of the operator $F = \mathrm{proj}_{B_\infty(0;\rho)} = \mathrm{Id} - \mathrm{prox}_{\rho\|\cdot\|_1}$, where $B_\infty(0;\rho)$ is the ℓ^∞ ball with radius $\rho \in \left]0, +\infty\right[$. The simulations are run in GNU Octave on a computer running Linux Ubuntu version 20.04 with a 2.60GHz dual-core processor and 8GB of RAM.

4.1. Signal Recovery

The goal is to recover the original signal $\overline{x} \in \mathcal{H} = C = \mathbb{R}^N$ ($N = 1024$) of Figure 1(a) from the following.

- A piecewise constant approximation p_1 of $\overline{x}$, given by $p_1 = \mathrm{proj}_{D_1}(\overline{x} + w_1)$, where $w_1 \in \mathcal{G}_1 = \mathcal{H}$ represents noise and D_1 is the subspace of signals in $\mathcal{G}_1$ which are constant by blocks along each of the 16 sets of 64 consecutive indices in $\{1, \ldots, N\}$ (see Figure 1(b)). The signal-to-noise ratio is $20\log_{10}(\|\overline{x}\|/\|w_1\|) = -2.3$ dB. We model this observation by setting $L_1 = \mathrm{Id}$ and $F_1 = \mathrm{proj}_{D_1}$.
- A bound $\rho_2 = 0.025$ on the magnitude of the finite differences of $\overline{x}$. To enforce this, following Example 7, we set $\mathcal{G}_2 = \mathbb{R}^{N-1}$, $L_2 \colon \mathcal{H} \to \mathcal{G}_2 \colon (\xi_i)_{1\leqslant i\leqslant N} \mapsto (\xi_{i+1} - \xi_i)_{1\leqslant i\leqslant N-1}$, $p_2 = 0$, and $F_2 = \mathrm{Id} - \mathrm{proj}_{D_2}$, where $D_2 = \big\{y \in \mathcal{G}_2 \mid \|y\|_\infty \leqslant \rho_2\big\}$.
- A collection of $m = 1200$ noisy thresholded scalar observations $r_3 = (\chi_j)_{j\in J} \in \mathbb{R}^m$ of $\overline{x}$, where $J = \{3, \ldots, m+2\}$. The true data formation model is

$$(\forall j \in J) \quad \chi_j = R(\langle \overline{x} \mid e_j \rangle) + \nu_j, \tag{12}$$

where $(e_j)_{j\in J}$ is a dictionary of random vectors in $\mathbb{R}^N$ with zero-mean i.i.d. entries, the noise vector $w_3 = (\nu_j)_{j\in J}$ yields a signal-to-noise ratio of $20\log_{10}(\|r_3\|/\|w_3\|) = 17.8$ dB, and R is the thresholding operator of the type found in [16] ($\rho = 0.05$), namely

$$R \colon \eta \mapsto \begin{cases} \mathrm{sign}(\eta)\sqrt[4]{\eta^4 - \rho^4}, & \text{if } |\eta| > \rho; \\ 0, & \text{if } |\eta| \leqslant \rho. \end{cases} \tag{13}$$

We assume that R is misspecified as

$$Q \colon \eta \mapsto \begin{cases} \mathrm{sign}(\eta)\sqrt{\eta^2 - \rho^2}, & \text{if } |\eta| > \rho; \\ 0, & \text{if } |\eta| \leqslant \rho, \end{cases} \tag{14}$$

and that the presence of noise is unknown, so that (12) is incorrectly modeled as $(\forall j \in J)\ \chi_j = Q(\langle \overline{x} \mid e_j \rangle)$. While Q is not Lipschitzian, with $S \colon \eta \mapsto \mathrm{sign}(\eta)(\sqrt{\eta^2 + \rho^2} - \rho)$, it is straightforward to verify that $S \circ Q = \mathrm{soft}_\rho$ and that, for every $j \in J$, $(F_j, p_j) = (\mathrm{soft}_\rho, S\chi_j)$ is a proxification of (Q, χ_j). Also, for every $j \in J$, we set $\mathcal{G}_j = \mathbb{R}$ and $L_j = \langle \cdot \mid e_j \rangle$.

We thus consider the instantiation of Problem 3 in which $I = \{1, 2\} \cup J$ and, for every $i \in I$, $\omega_i = 1/(\mathrm{card}\ I)$. Problem 3 has a solution by Proposition 13(iii) [11]. Algorithm (11) produces the signal shown in Figure 1(c) with $\gamma = 1.9$ and the following activation strategy. At every iteration, F_1 and F_2 are activated, while we partition J into four blocks of 300 elements and cyclically activate one block per iteration, i.e.,

$$(\forall n \in \mathbb{N})(\forall j \in \{0, 1, 2, 3\}) \quad I_{4n+j} = \{1, 2, 3 + 300j, \ldots, 2 + 300(j+1)\}, \tag{15}$$

which satisfies condition (10) with $K = 4$. The execution time savings resulting from the use of (15) compared to the full activation strategy (i.e., $I_n = I$ for every $n \in \mathbb{N}$) are displayed in Figure 3(a). The results show that, even when the data is noisy and poorly modeled, Problem 3 produces quite robust recoveries.

4.2. Image Recovery from Phase

The goal is to recover the original image $\overline{x} \in \mathcal{H} = \mathbb{R}^N$ ($N = 256^2$) shown in Figure 2(1a) from the following.

- Bounds on pixel values: $\overline{x} \in C = [0, 255]^N$.
- The degraded image $p_1 \in \mathcal{G}_1 = \mathcal{H}$ shown in Figure 2(1b), which is modeled as follows. The image $\overline{x}$ is blurred by $L_1 \colon \mathcal{H} \to \mathcal{G}_1$, which performs discrete convolution with a 15×15 Gaussian kernel with standard deviation of 3.5, then corrupted by additive noise $w_1 \in$

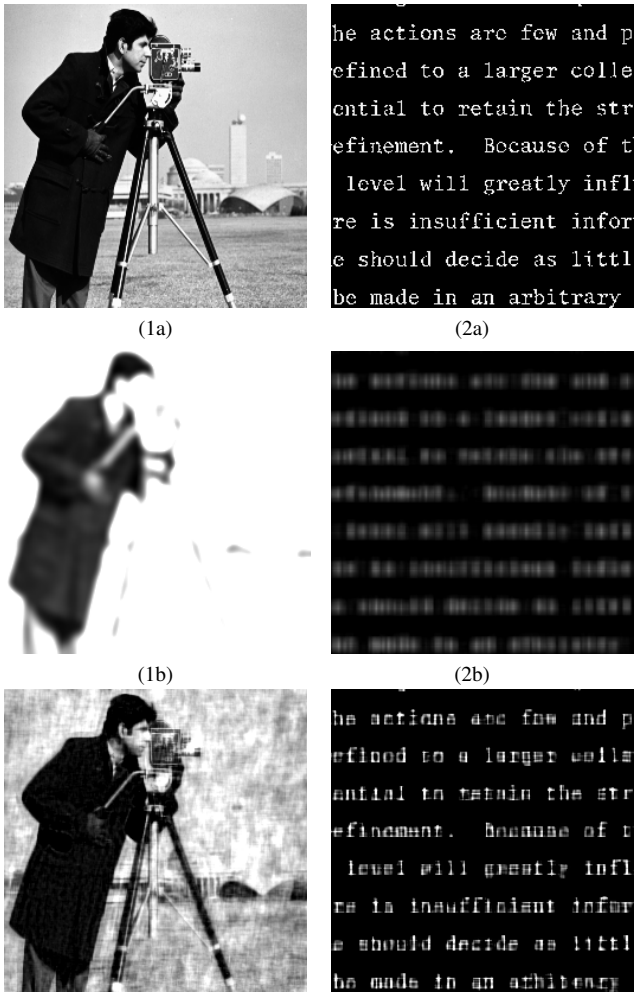

Fig. 2. Experiments of Sections 4.2 (Column 1) and 4.3 (Column 2). Top to bottom: original image $\overline{x}$, degraded image p_1, recovered image.

$\mathcal{G}_1$. The blurred image-to-noise ratio is $20\log_{10}(\|L_1\overline{x}\|/\|w_1\|) = 24.0$ dB. Pixel values beyond 60 are then clipped. Altogether, $p_1 = \mathrm{proj}_{D_1}(L_1\overline{x} + w_1)$, where $D_1 = [0, 60]^N$. This process models a low-quality image acquired by a device which saturates at photon counts beyond a threshold. Hence, we use $F_1 = \mathrm{proj}_{D_1}$ in (6).

• An approximation $\rho_2 = 138$ of the mean pixel value of $\overline{x}$. To enforce this, following Example 7, we set $\mathcal{G}_2 = \mathcal{H}$, $L_2 = \mathrm{Id}$, $p_2 = 0$, and $F_2 : (\eta_k)_{1\leqslant k\leqslant N} \mapsto \Big((1/N)\sum_{k=1}^{N}\eta_k - \rho_2\Big)\mathbf{1}$.

• The phase $\theta \in [-\pi, \pi]^N$ of the 2-D discrete Fourier transform of a noise-corrupted version of $\overline{x}$, i.e., $\theta = \angle(\widehat{\overline{x} + w_3})$, where $w_3 \in \mathcal{H}$ yields an image-to-noise ratio $20\log_{10}(\|\overline{x}\|/\|w_3\|) = 49.0$ dB. To model this information, we set $\mathcal{G}_3 = \mathcal{H}$, $L_3 = \mathrm{Id}$, $p_3 = 0$, and $F_3 : y \mapsto y - (|\widehat{y}|\max\{\cos(\angle(\widehat{y}) - \theta), 0\}e^{i\theta})^{\vee}$.

Due to the noise present in p_1 and θ, and the inexact estimation of ρ_2, this instance of Problem 1 ($I = \{1, 2, 3\}$) is inconsistent. We thus arrive at the relaxed Problem 3 by setting $\omega_1 = \omega_2 = \omega_3 = 1/3$. By Proposition 13(ii), since C is bounded, Problem 3 is guaranteed to possess a solution. The solution shown in Figure 2(1c) is

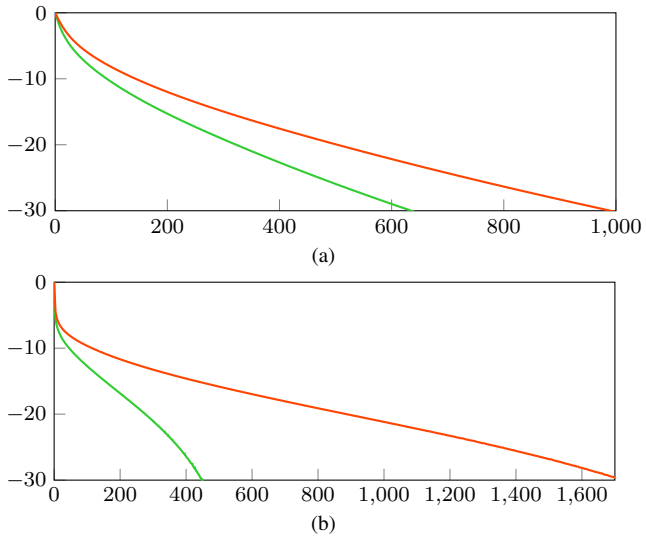

Fig. 3. Relative error $20\log_{10}(\|x_n - x_\infty\|/\|x_0 - x_\infty\|)$ (dB) versus execution time (seconds) for full activation (red) and block activation (green). (a): Section 4.1 and (15). (b): Section 4.3 and (17).

computed using algorithm (11) with $\gamma = 1.9$ and $(\forall n \in \mathbb{N})\ I_n = I$. This experiment illustrates a nonlinear recovery scenario with inconsistent measurements which nonetheless produces realistic solutions obtained by exploiting all available information.

4.3. Sparse Image Recovery

The goal is to recover the original image $\overline{x} \in \mathcal{H} = \mathbb{R}^N$ ($N = 256^2$) shown in Figure 2(2a) from the following.

• Bounds on pixel values: $x \in C = [0, 255]^N$.

• The low rank approximation $q_1 \in \mathcal{G}_1 = \mathcal{H}$ displayed in Figure 2(2b) of a blurred noisy version of $\overline{x}$ modeled as follows. The blurring operator $L_1 : \mathcal{H} \to \mathcal{G}_1$ applies a discrete convolution with a uniform 7×7 kernel, and the operators Q and S are as in Example 10, with threshold $\rho = 500$. Then

$$q_1 = Q(L_1\overline{x} + w_1) \tag{16}$$

is a rank-85 compression, where $w_1 \in \mathcal{G}_1$ induces a blurred image-to-noise ratio of $20\log_{10}(\|L_1\overline{x}\|/\|w_1\|) = 17.6$ dB. By Example 10, we obtain a proxification of (Q, q_1) with $(F_1, p_1) = (S \circ Q, Sq_1)$.

• $\overline{x}$ is sparse. To promote this property in the solutions to (6), following Remark 15, we set $\mathcal{G}_2 = \mathcal{H}$, $L_2 = \mathrm{Id}$, $p_2 = 0$, $\rho_2 = 1.5$, and $F_2 = \mathrm{proj}_{B_\infty(0;\rho_2)}$.

We therefore arrive at an instance of Problem 3 with $I = \{1, 2\}$ and $\omega_1 = \omega_2 = 1/2$. Since C is bounded, Proposition 13(ii) asserts that a solution exists. Algorithm (11) with $\gamma = 1$ yields the recovery in Figure 2(2c). Due to the singular value decomposition, F_1 is the most numerically costly operator and we choose to activate it only every 5 iterations, i.e.,

$$I_n = [\,\{2\},\ \text{if } n \not\equiv 0 \mod 5;\ \{1, 2\},\ \text{if } n \equiv 0 \mod 5\,]. \tag{17}$$

Figure 3(b) displays the time savings resulting from the use of (17) compared to full activation (both activation strategies yield visually indistinguishable recoveries). Notice that, while the observation in Figure 2(2b) is virtually illegible, many of the words in the recovery of Figure 2(2c) can be discerned.

5. REFERENCES

[1] J. S. Abel and J. O. Smith, Restoring a clipped signal, *Proc. Int. Conf. Acoust. Speech Signal Process.*, vol. 3, pp. 1745–1748, 1991.

[2] F. R. Ávila, M. P. Tcheou, and L. W. P. Biscainho, Audio soft declipping based on constrained weighted least squares, *IEEE Signal Process. Lett.*, vol. 24, pp. 1348–1352, 2017.

[3] R. E. Barlow and H. D. Brunk, The isotonic regression problem and its dual, *J. Amer. Stat. Assoc.*, vol. 67, pp. 140–147, 1972,

[4] H. Boche, M. Guillemard, G. Kutyniok, and F. Philipp, Signal recovery from thresholded frame measurements, *Proc. 15th SPIE Wavelets Sparsity Conf.*, vol. 8858, pp. 80–86, 2013.

[5] J. A. Cadzow, Signal enhancement – A composite property mapping algorithm, *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 36, pp. 49–62, 1988.

[6] E. Candès and T. Tao, Near-optimal signal recovery from random projections: Universal encoding strategies? *IEEE Trans. Inform. Theory*, vol. 52, pp. 5406–5425, 2006.

[7] Y. Censor, T. Elfving, N. Kopf, and T. Bortfeld, The multiple-sets split feasibility problem and its applications for inverse problems, *Inverse Problems*, vol. 21, pp. 2071–2084, 2005.

[8] P. L. Combettes and J.-C. Pesquet, Deep neural network structures solving variational inequalities, *Set-Valued Var. Anal.*, vol. 28, pp. 491–518, 2020.

[9] P. L. Combettes and Z. C. Woodstock, A fixed point framework for recovering signals from nonlinear transformations, *Proc. Europ. Signal Process. Conf.*, pp. 2120–2124, 2020.

[10] P. L. Combettes and Z. C. Woodstock, Reconstruction of functions from prescribed proximal points, *J. Approx. Theory*, vol. 268, art. 105606, 26 pp., 2021.

[11] P. L. Combettes and Z. C. Woodstock, A variational inequality model for the construction of signals from inconsistent nonlinear equations, *SIAM J. Imaging Sci.*, vol. 15, pp. 84–109, 2022.

[12] S. J. Dilworth, N. J. Kalton, D. Kutzarova, and V. N. Temlyakov, The thresholding greedy algorithm, greedy bases, and duality, *Constr. Approx.*, vol. 19, pp. 575–597, 2003.

[13] S. Foucart and J. Li, Sparse recovery from inaccurate saturated measurements, *Acta Appl. Math.*, vol. 158, pp. 49–66, 2018.

[14] A. Marmin, A. Jezierska, M. Castella, and J.-C. Pesquet, Global optimization for recovery of clipped signals corrupted with Poisson-Gaussian noise, *IEEE Signal Process. Lett.*, vol. 27, pp. 970–974, 2020.

[15] K. Nasrollahi and T. B. Moeslund, Super-resolution: A comprehensive survey, *Mach. Vis. Appl.*, vol. 25, pp. 1423–1468, 2014.

[16] T. Tao and B. Vidakovic, Almost everywhere behavior of general wavelet shrinkage operators, *Appl. Comput. Harmon. Anal.*, vol. 9, pp. 72–82, 2000.

[17] E. Tarr, *Hack Audio*. Routledge, New York, 2019.

[18] V. N. Temlyakov, The best m-term approximation and greedy algorithms, *Adv. Comput. Math.*, vol. 8, pp. 249–265, 1998.

[19] T. Teshima, M. Xu, I. Sato, and M. Sugiyama, Clipped matrix completion: A remedy for ceiling effects, *Proc. AAAI Conf. Artif. Intell.*, pp. 5151–5158, 2019.

[20] R. Tibshirani, Regression shrinkage and selection via the lasso: A retrospective, *J. R. Statist. Soc. B*, vol. 73, pp. 273–282, 2011.

[21] D. C. Youla, Generalized image restoration by the method of alternating orthogonal projections, *IEEE Trans. Circuits Syst.*, vol. 25, pp. 694–702, 1978.

[22] M. Yuan and Y. Lin, Model selection and estimation in regression with grouped variables, *J. R. Stat. Soc. Ser. B Stat. Methodol.*, vol. 68, pp. 49–67, 2006.