New Improved Algorithms for Compressive Sensing Based on ℓ_p Norm

Jeevan K. Pant, *Member, IEEE*, Wu-Sheng Lu, *Life Fellow, IEEE*, and Andreas Antoniou, *Life Fellow, IEEE*

Abstract—A new algorithm for the reconstruction of sparse signals, which is referred to as the ℓ_p-*regularized least squares* (ℓ_p-RLS) algorithm, is proposed. The new algorithm is based on the minimization of a smoothed ℓ_p-norm regularized square error with $p < 1$. It uses a conjugate-gradient (CG) optimization method in a sequential minimization strategy that involves a two-parameter continuation technique. An improved version of the new algorithm is also proposed, which entails a bisection technique that optimizes an inherent regularization parameter. Extensive simulation results show that the new algorithm offers improved signal reconstruction performance and requires reduced computational effort relative to several state-of-the-art competing algorithms. The improved version of the ℓ_p-RLS algorithm offers better performance than the basic version, although this is achieved at the cost of increased computational effort.

Index Terms—Compressive sensing (CS), conjugate-gradient (CG) optimization, least squares optimization, sequential optimization, ℓ_p-norm.

I. INTRODUCTION

COMPRESSIVE sensing (CS) is a technique for the acquisition of sparse signals introduced by Candès, Romberg, and Tao [1], Donoho [2], and Candès and Tao [3] in 2006. The focus of CS research has been on the reconstruction of sparse signals from a small number of linear measurements that may be corrupted by noise. Several types of algorithms have been proposed for CS, but the ℓ_1-minimization-based algorithms in [4] are among the most successful. Recently, several researchers have investigated gradient-descent and iteratively reweighted least squares (IRLS) algorithms that are based on ℓ_p-norm minimization with $p < 1$ [5], [6]; ℓ_q-minimization (ℓ_q-Min) with $0 < q \leq 1$, which is based on the minimization of a nonconvex approximation of the ℓ_p-norm [7]; and smoothed ℓ_0-norm (Sℓ_0) minimization [8], which have been found to offer an improved performance relative to ℓ_1-minimization

based algorithms. An improved IRLS algorithm called the iteratively reweighted unconstrained ℓ_q (IRucLq) algorithm that takes noisy measurements into account has been proposed in [9]. A greedy algorithm such as the orthogonal matching pursuit (OMP) [10] is known for its computational efficiency. Recently, a computationally efficient approximate message-passing (AMP) algorithm based on the ℓ_1-minimization has been proposed in [11].

In this brief, a new algorithm based on the minimization of an ℓ_p-regularized square error, which is referred to as the ℓ_p-*regularized least squares* (ℓ_p-RLS) algorithm, is proposed. A smoothed approximate ℓ_p-norm is used, which facilitates the application of a conjugate-gradient (CG) method in a *sequential minimization strategy*, which involves a *new two-parameter continuation technique* that assures the robustness of the proposed algorithm with respect to measurement noise. We use the so-called *basic CG* (BCG) method [12], which is very efficient and can also handle large-size data. An improved version of the new algorithm is also proposed, which entails a bisection technique that optimizes an inherent regularization parameter. The basic version of the algorithm offers improved reconstruction performance and requires a reduced amount of computation relative to several state-of-the-art competing algorithms. The improved version of the algorithm provides even better signal reconstruction performance than the basic version but at the cost of increased computational effort.

II. ℓ_p-RLS ALGORITHM

A. Background and Previous Works

In CS, a typical data acquisition model that takes measurement noise into account is given by

$$y = \Phi x + w \tag{1}$$

where the components of w are independent and identically distributed Gaussian random variables with zero mean and variance σ^2. Signal x can be recovered by using an ℓ_1-RLS algorithm, which solves the problem

$$\underset{x}{\text{minimize}} \; \frac{1}{2}\|\Phi x - y\|_2^2 + \lambda\|x\|_1 \tag{2}$$

where $\lambda > 0$ is a regularization parameter. An excellent survey of several recently proposed state-of-the-art algorithms for solving the problem in (2) can be found in [13].

J. K. Pant was with the Department of Electrical and Computer Engineering, University of Victoria, Victoria, BC V8W 2Y2, Canada. He is currently with the Department of Electrical and Computer Engineering, Ryerson University, Toronto, ON M5B 2K3 Canada (e-mail: jpant@rnet.ryerson.ca).

W.-S. Lu and A. Antoniou are with the Department of Electrical and Computer Engineering, University of Victoria, Victoria, BC V8W 2Y2, Canada (e-mail: wslu@ece.uvic.ca; aantoniou@ieee.org).

Color versions of one or more of the figures in this brief are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TCSII.2013.2296133

The IRLS and Sℓ_0 algorithms in [5] and [8], which solve the problems

$$\underset{\boldsymbol{x}}{\text{minimize}} \quad \|\boldsymbol{x}\|_p^p = \sum_{i=1}^{N} |x_i|^p \quad \text{subject to} \quad \boldsymbol{y} = \boldsymbol{\Phi}\boldsymbol{x} \qquad (3)$$

$$\underset{\boldsymbol{x}}{\text{minimize}} \quad \sum_{i=1}^{N} \left(1 - e^{-x_i^2/2\sigma^2}\right) \quad \text{subject to} \quad \boldsymbol{y} = \boldsymbol{\Phi}\boldsymbol{x} \qquad (4)$$

respectively, with $p < 1$ and a small positive σ using iterative procedures have been found to offer improved reconstruction performance relative to the ℓ_1-minimization-based algorithm. However, these algorithms are not very effective for the case of noisy measurements, i.e., for $\boldsymbol{w} \neq \boldsymbol{0}$ in (1).

In some preliminary work we reported in [14] and [15], we have investigated the minimization of the approximate ℓ_p-norm, i.e.,

$$\|\boldsymbol{x}\|_{p,\,\epsilon}^p = \sum_{i=1}^{N} \left(x_i^2 + \epsilon^2\right)^{p/2} \qquad (5)$$

for the reconstruction of sparse signals and found out that an improved signal reconstruction performance can be achieved relative to that achieved with several state-of-the-art algorithms. The algorithms in [14] and [15] involve the evaluation of the QR and singular-value decompositions; hence, they are not suitable for large signals.

B. Approximate ℓ_p-Norm and Problem Formulation

Consider the ℓ_p-norm $\|\boldsymbol{x}\|_p$ whose pth power $\|\boldsymbol{x}\|_p^p$ is given in (3). Each term $|x_i|^p$ in the summation in (3) is discontinuous for $x_i \approx 0$ and, consequently, it is not differentiable. It follows that the ℓ_p-norm is not differentiable; therefore, the gradient and Hessian of the objective function in (3) cannot be evaluated.

The proposed algorithm is based on the approximate ℓ_p-norm of $\boldsymbol{x}$ given in (5), where $p < 1$, and $\epsilon > 0$ is an approximation parameter. The inclusion of parameter ϵ in (5) introduces three highly desired properties to the approximate ℓ_p-norm. First, we have

$$\lim_{\epsilon \to 0} \|\boldsymbol{x}\|_{p,\,\epsilon} = \|\boldsymbol{x}\|_p.$$

Hence, with a small ϵ, $\|\boldsymbol{x}\|_{p,\,\epsilon}$ accurately approximates the ℓ_p-norm of $\boldsymbol{x}$. Second, with a nonzero ϵ, $\|\boldsymbol{x}\|_{p,\,\epsilon}$ becomes smooth; thus, it is differentiable. Third, the Hessian of $\|\boldsymbol{x}\|_{p,\,\epsilon}^p$ can be easily evaluated using the closed-form equation

$$\nabla^2 \|\boldsymbol{x}\|_{p,\,\epsilon}^p = \text{diag}\{u_1, u_2, \ldots, u_N\} \qquad (6a)$$

where

$$u_i = p\left(x_i^2 + \epsilon^2\right)^{p/2-2} \cdot \left[(p-1)x_i^2 + \epsilon^2\right]. \qquad (6b)$$

In effect, $\|\boldsymbol{x}\|_{p,\,\epsilon}^p$ is convex over the region

$$\left\{\boldsymbol{x} : |x_i| \leq \epsilon/\sqrt{1-p}\right\}.$$

By using this smooth and locally convex approximate ℓ_p-norm as the regularization term, signal $\boldsymbol{x}$ can be reconstructed from measurement $\boldsymbol{y}$ by solving the ℓ_p-RLS problem

$$\underset{\boldsymbol{x}}{\text{minimize}} \, f_{p,\,\epsilon}(\boldsymbol{x}) = \frac{1}{2}\|\boldsymbol{\Phi}\boldsymbol{x} - \boldsymbol{y}\|_2^2 + \lambda\|\boldsymbol{x}\|_{p,\,\epsilon}^p. \qquad (7)$$

Two remarks pertaining to the given problem are in order. First, for a very small ϵ, the problem in (7) becomes the ℓ_p formulation of the problem in (2), which assures a more accurate recovery of a sparse signal, as is well known [5], [6], [14]. Consequently, better performance can be achieved by solving the problem in (7) relative to that achieved by using an ℓ_1-RLS algorithm. Second, because of the differentiability and local convexity of $f_{p,\,\epsilon}(\boldsymbol{x})$, the solution of the problem in (7) can be approached using a variety of well-established optimization methods [12].

The StSALq algorithm [6] solves an optimization problem that is equivalent to solving the problem in (7). However, the optimization procedure used in this algorithm involves repetitive matrix inversion whose computational effort increases rapidly with the length of the signal. The IRucLq algorithm [9] also solves the problem in (7) but repetitive matrix inversion is avoided by iteratively constructing a set of linear equations and then solving it using matrix factorization. Consequently, the IRucLq algorithm offers reduced computational effort relative to the StSALq algorithm although the computational effort continues to be relatively large for long signals because of the matrix factorization involved.

C. Use of BCG Method

CG methods perform optimization by using search directions known as *conjugate* directions that are generated using the gradient of the objective function [12]. These methods have been found to be effective in many types of problems and have been used extensively in many applications other than CS in the past. In our proposed ℓ_p-RLS algorithm, we use the BCG algorithm described on p. 149 of [12] for the solution of the problem in (7) because of its simplicity and because it was found to offer improved reconstruction performance relative to several state-of-the-art reconstruction algorithms.

The gradient and Hessian of $f_{p,\,\epsilon}(\boldsymbol{x})$ can be readily expressed in closed form as

$$\boldsymbol{g} = \boldsymbol{\Phi}^T(\boldsymbol{\Phi}\boldsymbol{x} - \boldsymbol{y}) + \lambda\boldsymbol{g}_p \qquad (8)$$

where

$$\boldsymbol{g}_p = [g_{p_1} \, g_{p_2} \, \cdots \, g_{p_N}]^T$$

with

$$g_{p_i} = p\left(x_i^2 + \epsilon^2\right)^{p/2-1} x_i \qquad (9)$$

$$\boldsymbol{H} = \boldsymbol{\Phi}^T\boldsymbol{\Phi} + \lambda\boldsymbol{U}. \qquad (10)$$

In (10), $\boldsymbol{U} = \text{diag}\{u_1 u_2 \cdots u_N\}$ where u_i is given by (6b).

The problem at hand can be solved by using a sequential p, ϵ-*continuation* strategy as detailed in the next paragraph.

Given a pair of small target values p_T and ϵ_T, and a pair of sufficiently large initial values of parameters p and ϵ, i.e., p_1 and

ϵ_1, monotonically decreasing sequences $\{p_t : t = 1, 2, \ldots, T\}$ and $\{\epsilon_t, t = 1, 2, \ldots, T\}$ are generated as

$$p_t = p_1 e^{-\gamma(t-1)} \quad \text{for} \quad t = 1, 2, \ldots, T \tag{11}$$

$$\epsilon_t = \epsilon_1 e^{-\beta(t-1)} \quad \text{for} \quad t = 1, 2, \ldots, T \tag{12}$$

where $\gamma = \log(p_1/p_T)/(T-1)$ and $\beta = \log(\epsilon_1/\epsilon_T)/(T-1)$.

The problem in (7) is then solved sequentially for (p_t, ϵ_t) with $t = 1, 2, \ldots, T$. For each value of t, the problem is solved using a finite number of iterations, say, L_t, using the BCG algorithm.

From (10) and (6b), it is clear that $f_{p,\epsilon}(\boldsymbol{x})$ is convex over the region $\{\boldsymbol{x} : |x_i| \leq \epsilon/\sqrt{1-p}\}$. Therefore, for $p = 1$ function $f_{p,\epsilon}(\boldsymbol{x})$ is convex and for any $\epsilon > 0$ it is smooth; the larger the ϵ is, the smoother the function. If $|x_m| = \max_i |x_i|$, $\epsilon \geq |x_m|\sqrt{1-p_T}$ would make the function $f_{p,\epsilon}(\boldsymbol{x})$ sufficiently smooth, where p_T is the target value of p used in (11). Consequently, the global minimizer $\boldsymbol{x}_{1,\epsilon}^*$ of this function can be readily obtained by running the BCG algorithm over a sufficiently large number of iterations, say, L_1. By using this point as the initial point for the next round of L_2 BCG iterations to minimize $f_{p_2,\epsilon_2}(\boldsymbol{x})$, an improved solution is generated. This procedure is continued until function $f_{p_T,\epsilon_T}(\boldsymbol{x})$ is minimized using L_T BCG iterations.

In the BCG algorithm, iterate $\boldsymbol{x}_k$ is updated as [12]

$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \alpha_k d_k \tag{13a}$$

where

$$d_k = -\boldsymbol{g}_k + \beta_{k-1} d_{k-1} \tag{13b}$$

$$\beta_{k-1} = \frac{\|\boldsymbol{g}_k\|_2^2}{\|\boldsymbol{g}_{k-1}\|_2^2} \tag{13c}$$

$$\alpha_k = \frac{\|\boldsymbol{g}_k\|_2^2}{d_k^T \boldsymbol{H}_k d_k}. \tag{13d}$$

In (13), $\boldsymbol{g}_k$ and $\boldsymbol{H}_k$ are the gradient and Hessian of $f_{p,\epsilon}(\boldsymbol{x})$ evaluated at $\boldsymbol{x} = \boldsymbol{x}_k$ using (8) and (10), respectively. In order for α_k in (13d) to be positive, $\boldsymbol{H}_k$ must be positive definite (PD). It follows from (10) that $\boldsymbol{H}_k$ is PD if $\boldsymbol{U}_k$ is PD and the positive definiteness of $\boldsymbol{U}_k$ can be assured by letting

$$u_i = \begin{cases} u_i & \text{if } u_i > \delta \\ \delta & \text{if } u_i \leq \delta \end{cases} \tag{14}$$

where δ is a small positive constant typically of the order of 10^{-5}. The denominator in (13d) can be evaluated efficiently as

$$d_k^T \boldsymbol{H}_k d_k = \|\boldsymbol{\Phi} d_k\|_2^2 + \lambda \|\boldsymbol{b}_k\|_2^2 \tag{15}$$

where $\boldsymbol{b}_k = [b_{k1} b_{k2} \cdots b_{kN}]^T$ with $b_{ki} = \sqrt{u_i} d_{ki}$, and from (13d), we have

$$\alpha_k = \frac{\|\boldsymbol{g}_k\|_2^2}{\|\boldsymbol{\Phi} d_k\|_2^2 + \lambda \|\boldsymbol{b}_k\|_2^2}. \tag{16}$$

Based on the given principles, the basic ℓ_p-RLS algorithm summarized in Table I can be constructed, where parameter E_t is a sufficiently small threshold.

TABLE I
BASIC ℓ_p-RLS ALGORITHM

Step 1: Input T, p_1, p_T, ϵ_1, ϵ_T, λ, E_t, δ, $\boldsymbol{\Phi}$, $\boldsymbol{y}$ and set $\boldsymbol{x}_s = \boldsymbol{0}$.
Step 2: Compute p_t and ϵ_t for $t = 2, 3, \ldots, T - 1$ using (11) and (12), respectively.
Step 3: For $t = 1, 2, \ldots, T$
    i) Set $\epsilon = \epsilon_t$, $p = p_t$, and $L_t = 6 + \text{round}(t/5)$.
    ii) Set $k = 0$, $\boldsymbol{x}_0 = \boldsymbol{x}_s$, and $E_r = 10^{10}$.
    iii) While $E_r > E_t$.
      a) Compute $\boldsymbol{x}_{k+1}$ using (13), (14), (16).
      b) Set $k = k + 1$.
      c) Exit loop if $k > L_t$.
      d) Compute $E_r = \|\alpha_k d_k\|_2$.
    iv) Set $\boldsymbol{x}_s = \boldsymbol{x}_k$.
Step 4: Output $\boldsymbol{x}^* = \boldsymbol{x}_s$.

III. OPTIMIZATION OF PARAMETER λ

The performance of the proposed ℓ_p-RLS algorithm is closely related to the selection of regularization parameter λ as it controls a tradeoff between the approximate ℓ_p-norm and the closeness of $\boldsymbol{\Phi}\boldsymbol{x}$ to measurement $\boldsymbol{y}$. For a solution $\boldsymbol{x}$ of the problem in (7), (1) implies that $\|\boldsymbol{\Phi}\boldsymbol{x} - \boldsymbol{y}\|_2^2 \approx M\sigma^2$, where M is the size of vector $\boldsymbol{y}$. If λ is large, then the weight used in the approximate ℓ_p-norm will be large, which will cause the fidelity term $\|\boldsymbol{\Phi}\boldsymbol{x} - \boldsymbol{y}\|_2^2$ to exceed $\|\boldsymbol{w}\|_2^2 = M\sigma^2$. On the other hand, if λ is small, then the fidelity term becomes overweighted, which will result in a value of $\|\boldsymbol{\Phi}\boldsymbol{x} - \boldsymbol{y}\|_2^2$ that is considerably smaller than $M\sigma^2$, thereby violating (1) again. In effect, $\|\boldsymbol{\Phi}\boldsymbol{x} - \boldsymbol{y}\|_2^2$ is a monotonic function of λ. In such a situation, a bisection technique such as that in [12] can be used, as detailed in the next paragraph, to obtain a good value of λ.

Let $[\lambda_l, \lambda_u]$ be an interval that contains the optimal value of λ. Since λ is nonnegative, initially we can set $\lambda_l = 0$ and λ_u sufficiently large. The length of this interval is the initial range of uncertainty (ROU). By using the midpoint of the interval $[\lambda_l, \lambda_u]$, i.e., $\lambda = (\lambda_l + \lambda_u)/2$, the problem in (7) is solved, and its solution $\boldsymbol{x}_\lambda$ is used to update λ_l and λ_u as

$$\begin{aligned} \text{set} \quad &\lambda_l = \lambda \quad \text{if} \quad \|\boldsymbol{\Phi}\boldsymbol{x}_\lambda - \boldsymbol{y}\|_2^2 < M\sigma^2 \\ \text{set} \quad &\lambda_u = \lambda \quad \text{if} \quad \|\boldsymbol{\Phi}\boldsymbol{x}_\lambda - \boldsymbol{y}\|_2^2 > M\sigma^2. \end{aligned} \tag{17}$$

In this way, the ROU is reduced by half. This procedure is repeated until the ROU is reduced to a level below a prescribed tolerance, and the midpoint of the last interval is taken to be an estimate of the optimized value of λ. Because of the exponential convergence rate 0.5^k of the procedure, a good value of λ can be found with a small number of iterations.

The given bisection technique can be readily incorporated in the ℓ_p-RLS algorithm, as shown in Table II, where parameter E_{to} is a small threshold. The two versions of the proposed algorithm will be referred to as the *basic* and *improved* ℓ_p-RLS algorithm, respectively, hereafter.

IV. SIMULATION RESULTS

In our experiments, a K-sparse signal of length N with an energy value of 100 units was constructed as follows: 1) A zero vector $\boldsymbol{x}$ of length N was constructed; 2) a random vector of length K was constructed by drawing its components from a normal distribution $\mathcal{N}(0, 1)$ and then applying a normalization step to ensure that the ℓ_2-norm of the resulting vector is $\sqrt{100}$;

TABLE II
IMPROVED ℓ_p-RLS ALGORITHM

Step 1: Input T, p_1, p_T, ϵ_1, ϵ_T, λ_u, λ_l, $\mathbf{\Phi}$, $\mathbf{y}$, E_t, E_{to}, δ, σ
and set $\mathbf{x}_s = \mathbf{0}$, $E_{ro} = 10^{10}$.
Step 2: Repeat the following while $E_{ro} > E_{to}$.
   i) Compute $\lambda = (\lambda_l + \lambda_u)/2$.
   ii) Solve (7) using the basic ℓ_p-RLS algorithm with T,
      p_1, p_T, ϵ_1, ϵ_T, λ, $\mathbf{\Phi}$, $\mathbf{y}$, E_t, δ as inputs and $\mathbf{x}_s$ as the
      initializer. Denote the resulting solution as $\mathbf{x}_\lambda$.
   iii) Update λ_u or λ_l using (17).
   iv) Compute $E_{ro} = \lambda_u - \lambda_l$.
   v) Set $\mathbf{x}_s = \mathbf{x}_\lambda$.
Step 3: Compute $\lambda = (\lambda_l + \lambda_u)/2$,
solve the problem in (7) using the basic ℓ_p-RLS
algorithm, and denote the resulting solution as $\mathbf{x}_\lambda$.
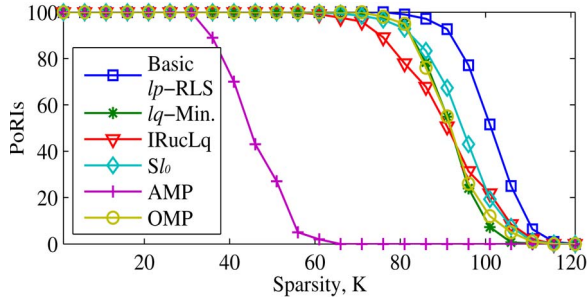**Step 4:** Output $\mathbf{x}^* = \mathbf{x}_\lambda$ and stop.

Fig. 1. PoRIs versus sparsity for basic ℓ_p-RLS, ℓ_q-Min, IRucLq, Sℓ_0, AMP, and OMP algorithms (Experiment 1).

Fig. 2. Average CPU time versus signal length for basic ℓ_p-RLS, ℓ_q-Min, IRucLq, Sℓ_0, AMP, and OMP algorithms (Experiment 2).

and 3) the components of the resulting vector were set to K randomly chosen locations of vector $\mathbf{x}$. Measurement matrix $\mathbf{\Phi}$ of size $M \times N$ was constructed by selecting its components from $\mathcal{N}(0, 1)$ and then by applying an orthonormalization step to ensure that the rows of $\mathbf{\Phi}$ are orthonormally related to each other. A noisy measurement was taken using $\mathbf{y} = \mathbf{\Phi}\mathbf{x} + \mathbf{w}$, where $\mathbf{w}$ is a measurement noise vector of length M, whose components were drawn from $\mathcal{N}(0, \sigma^2)$.

In Experiment 1, the signal length N, number of measurements M, noise variance σ^2, and λ were set to $N = 512$, $M = 200$, $\sigma^2 = 10^{-4}$, and $\lambda = 7 \times 10^{-3}$, respectively. The basic ℓ_p-RLS algorithm was run with parameters $T = 80$, $p_1 = 1$, $p_T = 0.1$, $\epsilon_1 = 1$, $\epsilon_T = 10^{-2}$, $E_t = 10^{-25}$, and $\delta = 10^{-5}$. The IRucLq algorithm [9] was run with $\lambda = 7 \times 10^{-3}$. The ℓ_q-Min algorithm [7] was run with $n = 10$, $\epsilon = 1/(k+2)$, and $q \in \{0, 0.05, 0.1, 0.2\}$. The S$\ell_0$ algorithm [8] was run with $\sigma_1 = 2$, $r = 1.001$, $L = 2$, and $\sigma_T = 10^{-2}$. The SNR was measured as $20 \log(\|\mathbf{x}\|_2/\|\mathbf{x} - \hat{\mathbf{x}}\|_2)$, where $\hat{\mathbf{x}}$ is a vector representing the reconstructed signal. The signal was deemed to be reconstructed when the SNR was greater than 29 dB. The percentage of recovered instances (PoRIs) of the required signal for the basic ℓ_p-RLS, ℓ_q-Min, IRucLq, Sℓ_0, AMP, and OMP algorithms over 1000 runs is plotted versus sparsity in Fig. 1. As can be seen, the basic ℓ_p-RLS algorithm yields increased PoRIs for sparsity values in excess of 80 relative to the competing algorithms. The performance of the proposed algorithm is better than that of the AMP algorithm for values of K in excess of 35.

In Experiment 2, the number of measurements and sparsity value were set to $M = N/2$ and $K = \text{round}(M/2.5)$, and the noise variance was set to $\sigma^2 = 10^{-4}$. The CPU time was measured on a desktop PC with Intel Core i7-3632QM CPU
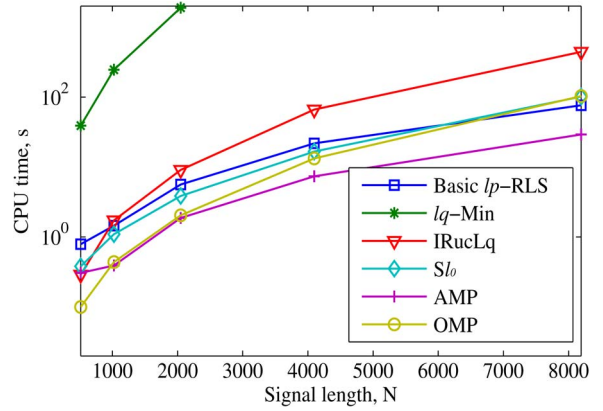
2.2-GHz processor and 8-GB RAM using MATLAB command `cputime`. The average CPU time over 1000 runs versus the signal length is plotted in Fig. 2. As can be seen, the CPU time required by the basic ℓ_p-RLS algorithm is less than that required by the ℓ_q-Min and IRucLq algorithms but more than that required by the AMP algorithm. The ℓ_q-Min algorithm involves the use of a primal–dual interior-point method, which requires a significantly larger computational effort than the technique used in the basic ℓ_p-RLS algorithm. Unlike the IRucLq algorithm, the basic ℓ_p-RLS algorithm does not involve repetitive matrix factorization and, consequently, the rate of increase in the computational effort with respect to the signal length N for the basic ℓ_p-RLS algorithm is smaller than that for the IRucLq algorithm. The basic ℓ_p-RLS algorithm required less CPU time than the OMP and Sℓ_0 algorithms for values of N above 6400. This is due to the fact that the OMP method entails the use of QR factorization, and the Sℓ_0 algorithm entails evaluation of the pseudoinverse of matrix $\mathbf{\Phi}$, which requires a large amount of computation for large signals.

In Experiment 3, the signal length, the number of measurements, and sparsity values were selected as in Experiment 1. The basic ℓ_p-RLS algorithm with p *fixed* was implemented as in Table I with the following modifications: 1) $p = 0.1$ was supplied as input instead of p_1 and p_T in Step 1; 2) the computation of p_t in Step 2 was removed; and 3) the statement $p = p_t$ in Step 3.i was removed. The basic ℓ_p-RLS algorithm with p, ϵ continuation, exactly as described in Table I, and with p fixed were run with $\lambda = 9 \times 10^{-3}$, 2×10^{-2}, and 4×10^{-2} for $\sigma = 3 \times 10^{-2}$, 5×10^{-2}, and 7×10^{-2}, respectively. The average SNR obtained is plotted in Fig. 3 versus sparsity. As can be seen, for sparsity values up to 80, the SNR for the basic ℓ_p-RLS algorithm with p, ϵ continuation is larger than that with p fixed for all three values of σ. Moreover, the difference between the SNRs for the two implementations of the algorithm is greatest when $\sigma = 7 \times 10^{-2}$, which implies that the basic ℓ_p-RLS algorithm with p, ϵ continuation is more robust with respect to the measurement noise than the algorithm with p fixed.

In Experiment 4, the signal length, the number of measurements, and the sparsity values were set as in Experiment 1. The improved ℓ_p-RLS algorithm was run with $\lambda_l = 0$, $\lambda_u =$
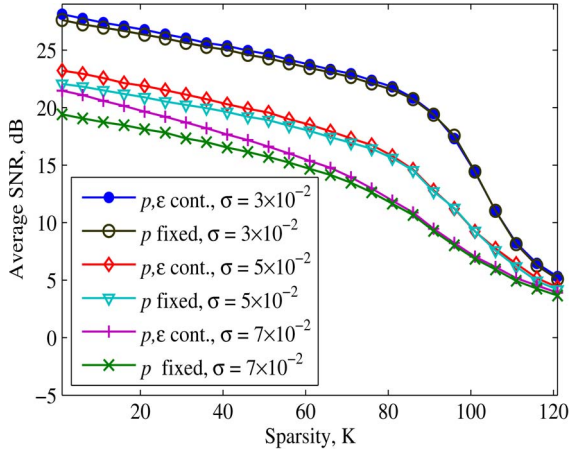
Fig. 3.   Average SNR versus sparsity for basic ℓ_p-RLS algorithm with p fixed and p, ϵ continuation (Experiment 3).
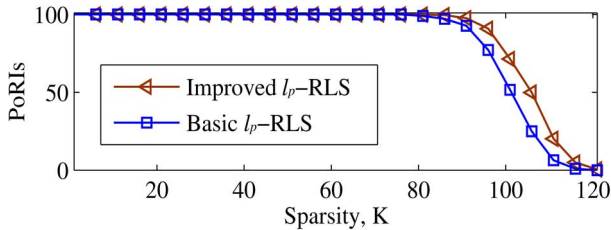

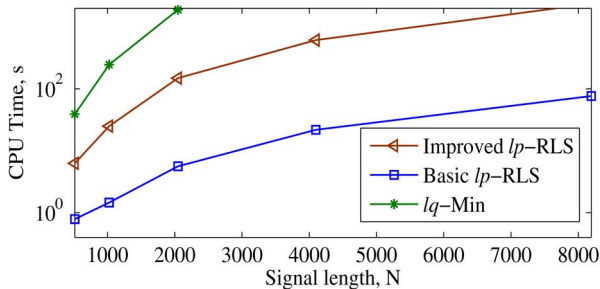Fig. 4.   PoRIs versus sparsity for basic and improved ℓ_p-RLS algorithms (Experiment 4).

Fig. 5.   Average CPU time versus signal length for basic and improved ℓ_p-RLS algorithms and ℓ_q-Min algorithm (Experiment 5).

5×10^{-3}, $E_{\text{to}} = 1 \times 10^{-4}$, and $\sigma^2 = 10^{-4}$. For comparison, the basic ℓ_p-RLS algorithm in Table I was also run with $\lambda = 7 \times 10^{-4}$. The results obtained are plotted in Fig. 4. As can be seen, optimizing parameter λ by using the proposed bisection technique increases the PoRIs quite significantly for values of K over 80 relative to that achieved with the basic ℓ_p-RLS algorithm. Step 2 of the improved ℓ_p-RLS algorithm took a total of six iterations.

In Experiment 5, the number of measurements and sparsity value were set as in Experiment 2. The average CPU times for basic and improved ℓ_p-RLS and ℓ_q-Min algorithms are plotted in Fig. 5. As can be seen, the use of the bisection technique has increased the computational effort required. Nevertheless, the improved ℓ_p-RLS algorithm continues to be more efficient than the ℓ_q-Min algorithm.

Some additional results that cannot be included in this brief due to page limitations can be found in [16].

It should be mentioned that the algorithms proposed in this brief require the variance of the noise. State-of-the-art techniques for the estimation of the variance of the noise are available in the literature (see, for example, the references included in [16]).

V. CONCLUSION

A new algorithm for CS, i.e., the basic ℓ_p-RLS algorithm, has been proposed. It uses a CG optimization method in a sequential minimization strategy that involves a two-parameter continuation technique. An improved ℓ_p-RLS algorithm has also been proposed that uses in addition a bisection technique for the optimization of an inherent regularization parameter. Extensive simulation results have shown that the basic ℓ_p-RLS algorithm offers improved PoRIs and reduced computational effort for large data relative to the competing algorithms. The continuation of parameter p improves the robustness of the basic ℓ_p-RLS algorithm. The improved ℓ_p-RLS algorithm yields an increase in the PoRIs with respect to the basic ℓ_p-RLS algorithm, although this is achieved at the cost of increased computational effort.

REFERENCES

[1] E. J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 489–509, Feb. 2006.

[2] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.

[3] E. J. Candès and T. Tao, "Near-optimal signal recovery from random projections: Universal encoding strategies," *IEEE Trans. Inf. Theory*, vol. 52, no. 12, pp. 5406–5425, Dec. 2006.

[4] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM J. Sci. Comput.*, vol. 20, no. 1, pp. 33–61, 1998.

[5] R. Chartrand and W. Yin, "Iteratively reweighted algorithms for compressive sensing," in *Proc. IEEE Int. Conf. Acoust, Speech, Signal Process.*, 2008, pp. 3869–3872.

[6] R. Saab, R. Chartrand, and O. Yilmaz, "Stable sparse approximations via nonconvex optimization," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Mar./Apr. 2008, pp. 3885–3888.

[7] S. Foucart and M.-J. Lai, "Sparsest solutions of underdetermined linear systems via ℓ_q-minimization for $0 < q \leq 1$," *Appl. Comput. Harmon. Anal.*, vol. 26, no. 3, pp. 395–407, May 2009.

[8] H. Mohimani, M. Babie-Zadeh, and C. Jutten, "A fast approach for overcomplete sparse decomposition based on smoothed ℓ_0-norm," *IEEE Trans. Signal Process.*, vol. 57, no. 1, pp. 289–301, Jan. 2009.

[9] M.-J. Lai, Y. Xu, and W. Yin, "Improved iteratively reweighted least squares for unconstrained smoothed ℓ_q minimization," *SIAM J. Numer. Anal.*, vol. 51, no. 2, pp. 927–957, 2013.

[10] Y. C. Pati, R. Rezaiifar, and P. S. Krishnaprasad, "Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition," in *Proc. 27th Annu. Asilomar Conf. Signals Syst. Comput.*, Nov. 1993, vol. 1, pp. 40–44.

[11] D. L. Donoho, A. Maleki, and A. Montanari, "Message-passing algorithms for compressed sensing," *Proc. Natl. Acad. Sci.*, vol. 106, no. 45, pp. 18 914–18 919, Nov. 2009.

[12] A. Antoniou and W.-S. Lu, *Practical Optimization: Algorithms and Engineering Applications*.   New York, NY, USA: Springer, 2007.

[13] M. Zibulevsky and M. Elad, "L1-L2 optimization in signal and image processing," *IEEE Signal Process. Mag.*, vol. 27, no. 3, pp. 76–88, May 2010.

[14] J. K. Pant, W.-S. Lu, and A. Antoniou, "Unconstrained regularized ℓ_p-norm based algorithm for the reconstruction of sparse signals," in *Proc. IEEE Int. Symp. Circuits Syst.*, May 2011, pp. 1740–1743.

[15] J. K. Pant, W.-S. Lu, and A. Antoniou, "Recovery of sparse signals from noisy measurements using an ℓ_p regularized least-squares algorithm," in *Proc. IEEE Pacific Rim Conf. Commun., Comput., Signal Process.*, Aug. 2011, pp. 48–53.

[16] J. K. Pant, W.-S. Lu, and A. Antoniou, Additional results: New improved algorithms for compressive sensing based on ℓ_p norm. [Online]. Available: http://www.ece.uvic.ca/~andreas/JournalPapers/New_Improved_Algorithms_for_CS-Additional_Results.pdf