Reconstruction of Block-Sparse Signals by Using an $\ell_{2/p}$-Regularized Least-Squares Algorithm

Jeevan K. Pant, Wu-Sheng Lu, and Andreas Antoniou
Department of Electrical and Computer Engineering, University of Victoria
P. O. Box 3055 STN CSC, Victoria, B.C. V8W 3P6, CANADA
Email:{jkpant, wslu}@ece.uvic.ca, aantoniou@ieee.org

***Abstract*—A new algorithm for the reconstruction of so called block-sparse signals in a compressive sensing framework is presented. The algorithm is based on minimizing an $\ell_{2/p}$-norm regularized ℓ_2 error. The minimization is carried out by using a sequential conjugate-gradient algorithm where the line search involved is carried out using a technique based on Banach's fixed-point theorem. Simulation results are presented which show that for large-size data the proposed algorithm yields improved reconstruction performance and requires a reduced amount of computation relative to several known algorithms.**

I. Introduction

In conventional compressive sensing (CS), the algorithms used to recover sparse signals do not take into account the block structure of the signal components, where nonzero coefficients occur in cluster [1]-[7]. In [8] and [9], an algorithm based on an $\ell_{2/1}$-norm minimization is used to recover block-sparse signals where the $\ell_{2/1}$-norm minimization problem is solved by recasting it as a second-order cone-programming (SOCP) problem. In [10], a block orthogonal matching pursuit (BOMP) algorithm is proposed as an extension of the orthogonal matching pursuit (OMP) algorithm [11] for block-sparse signals.

In this paper, we propose a new algorithm for the reconstruction of so called *block-sparse* signals in the CS framework by minimizing an $\ell_{2/p}$-norm regularized ℓ_2 error with $p < 1$. The $\ell_{2/p}$ norm is used to promote inter-block sparsity in the signal. The optimization problem involved is solved using a sequential procedure where each optimization problem is solved by a conjugate-gradient algorithm known as the Fletcher-Reeves algorithm [12]. The approximation of the ℓ_p norm facilitates the application of the Fletcher-Reeves algorithm and helps to accelerate the convergence to the optimal solution. Simulation results are presented, which demonstrate that for large-size data the proposed algorithm yields improved reconstruction performance and requires a reduced amount of computation relative to several known algorithms.

II. Preliminaries

A. Compressive sensing

A real-valued discrete-time signal $\boldsymbol{x}$ of length N is said to be K-sparse if it has K non-zero components with $K \ll N$. The measurement process in CS is described by

$$\boldsymbol{y} = \boldsymbol{\Phi}\boldsymbol{x} \tag{1}$$

where $\boldsymbol{y}$ is a measurement vector of length M and $\boldsymbol{\Phi}$ is a measurement matrix of size $M \times N$. With $M < N$, the inverse-problem of recovering signal $\boldsymbol{x}$ from measurements $\boldsymbol{y}$ is ill-posed [13]. In principle, the sparsest solution of (1) can be estimated by minimizing the ℓ_0 norm of $\boldsymbol{x}$, i.e., $||\boldsymbol{x}||_0 = \sum_{i=1}^{N} |x_i|^0$, subject to the constraint in (1). Unfortunately, the problem of finding the solution with the minimum ℓ_0 norm requires a combinatorial search among all the solutions of (1) for which the computation required grows exponentially as N increases. A tractable algorithm for the reconstruction of signal $\boldsymbol{x}$ is the ℓ_1-minimization based *basis pursuit* (BP) algorithm which solves the convex optimization problem

$$\underset{\boldsymbol{x}}{\text{minimize}} \quad ||\boldsymbol{x}||_1 \quad \text{subject to:} \quad \boldsymbol{\Phi}\boldsymbol{x} = \boldsymbol{y} \tag{2}$$

where $||\boldsymbol{x}||_1$ is the ℓ_1 norm of $\boldsymbol{x}$ defined as $||\boldsymbol{x}||_1 = \sum_{i=1}^{N} |x_i|$ [4]. A key result in CS is that if signal $\boldsymbol{x}$ is K-sparse, the elements of $\boldsymbol{\Phi}$ are drawn from a Gaussian distribution $\mathcal{N}(0, 1/\mathrm{N})$, and the number of measurements M satisfies the condition

$$M \geq cK \log(N/K) \tag{3}$$

with c a small constant, then $\boldsymbol{x}$ can be recovered by solving the problem in (2) [1] - [3]. In [5] and [6], several ℓ_p-minimization based algorithms that solve the optimization problem

$$\underset{\boldsymbol{x}}{\text{minimize}} \quad ||\boldsymbol{x}||_p^p \quad \text{subject to} \quad \boldsymbol{\Phi}\boldsymbol{x} = \boldsymbol{y} \tag{4}$$

with $p < 1$ are shown to offer improved signal reconstruction performance compared to the BP algorithm.

B. Block-sparse signals

Consider signal $\boldsymbol{x}$ of length N which is divisible by a positive integer d. We divide signal $\boldsymbol{x}$ into N/d blocks $\tilde{\boldsymbol{x}}_1$, $\tilde{\boldsymbol{x}}_2$, ..., $\tilde{\boldsymbol{x}}_{N/d}$ and denote $\boldsymbol{x}$ as

$$\boldsymbol{x} = \left[\tilde{\boldsymbol{x}}_1^T \ \tilde{\boldsymbol{x}}_2^T \ \cdots \ \tilde{\boldsymbol{x}}_{N/d}^T\right]^T \tag{5}$$

where

$$\tilde{\boldsymbol{x}}_i = \left[x_{(i-1)d+1} \ x_{(i-1)d+2} \ \cdots \ x_{(i-1)d+d}\right]^T$$

for $i = 1, 2, \ldots, N/d$ and x_i is the ith component of $\boldsymbol{x}$.

The signal $\boldsymbol{x}$ in (5) is said to be K-block sparse if $\boldsymbol{x}$ has K nonzero blocks with $K \ll N/d$. Note that the definition of K-sparse in the conventional CS is the special case of K-block sparse with $d = 1$. Recently, it has been shown that improved performance for the reconstruction of K-block sparse signals can be achieved by solving the $\ell_{2/1}$-norm minimization problem

$$\begin{array}{ll}\underset{\boldsymbol{x}}{\text{minimize}} & ||\boldsymbol{x}||_{2/1} \\ \text{subject to} & \boldsymbol{\Phi x} = \boldsymbol{y}\end{array} \tag{6}$$

where $||\boldsymbol{x}||_{2/1}$ is the $\ell_{2/1}$ norm of $\boldsymbol{x}$ defined as

$$||\boldsymbol{x}||_{2/1} = \sum_{i=1}^{N/d} ||\tilde{\boldsymbol{x}}_i||_2 \tag{7}$$

where $||\tilde{\boldsymbol{x}}_i||_2$ is the ℓ_2 norm of the ith block $\tilde{\boldsymbol{x}}_i$ [8], [9], [10]. The problem in (6) can be recast as

$$\begin{array}{lll}\underset{\boldsymbol{x},\boldsymbol{t}}{\text{minimize}} & \sum_{i=1}^{N/d} t_i & \\ \text{subject to} & \boldsymbol{\Phi x} = \boldsymbol{y} & \\ & ||\tilde{\boldsymbol{x}}_i||_2 \le t_i & 1 \le i \le N/d \\ & 0 \le t_i & 1 \le i \le N/d\end{array} \tag{8}$$

which can be solved using an SOCP solver. In the rest of the paper, the algorithm used to solve the above problem will be referred to as the $\ell_{2/1}$-SOCP algorithm. In [10], the BOMP algorithm was shown to offer improved signal reconstruction performance relative to that of the $\ell_{2/1}$-SOCP algorithm.

III. Reconstruction of Block-Sparse Signals by Using an $\ell_{2/p}$-Regularized Least-Squares Optimization

A. Approximate $\ell_{2/p}$ norm and problem conversion

ℓ_p minimization with $p < 1$ is known to offer improved signal reconstruction performance relative to ℓ_1 minimization. In what follows, we propose a method for the reconstruction of block-sparse signals by solving the optimization problem

$$\underset{\boldsymbol{x}}{\text{minimize}} \quad F_\epsilon(\boldsymbol{x}) = \tfrac{1}{2}\,||\boldsymbol{\Phi x} - \boldsymbol{y}||_2^2 + \lambda\,||\boldsymbol{x}||_{2/p,\epsilon} \tag{9}$$

with $p < 1$, where $\lambda > 0$ is a regularization parameter and $||\boldsymbol{x}||_{2/p,\epsilon}$ is an approximate $\ell_{2/p}$ norm of $\boldsymbol{x}$ defined as

$$||\boldsymbol{x}||_{2/p,\epsilon} = \sum_{i=1}^{N/d} \left(||\tilde{\boldsymbol{x}}_i||_2^2 + \epsilon^2\right)^{p/2} \tag{10}$$

Good signal reconstruction performance is expected when the proposed method is applied with very small p and ϵ because in that case function $||\boldsymbol{x}||_{2/p,\epsilon}$ accurately approximates the ℓ_0 norm of signal $\boldsymbol{x}$. However, function $||\boldsymbol{x}||_{2/p,\epsilon}$ becomes highly nonconvex and nearly nondifferentiable if p and ϵ are too small, which makes the problem in (9) a difficult one to solve. In our experiments, we found $p = 0.1$ and $\epsilon = 10^{-5}$ appropriate for a wide range of signals and for noiseless measurements.

Note that function $||\boldsymbol{x}||_{2/p,\epsilon}$ remains differentiable and so is function $F_\epsilon(\boldsymbol{x})$ in (9) as long as ϵ is kept positive. In effect, for $\epsilon > 0$ the gradient of $F_\epsilon(\boldsymbol{x})$ is given by

$$\boldsymbol{g} = \boldsymbol{\Phi}^T(\boldsymbol{\Phi x} - \boldsymbol{y}) + \lambda \boldsymbol{u} \tag{11}$$

where $\boldsymbol{u}$ denotes the gradient of $||\boldsymbol{x}||_{2/p,\epsilon}$ and assumes the form

$$\boldsymbol{u} = \left[\tilde{\boldsymbol{u}}_1^T \ \tilde{\boldsymbol{u}}_2^T \ \cdots \ \tilde{\boldsymbol{u}}_{N/d}^T\right]^T \tag{12}$$

where $\tilde{\boldsymbol{u}}_i$ is the ith block of $\boldsymbol{u}$. The $\{(i-1)d+j\}$th component of $\boldsymbol{u}$ in the ith block $\tilde{\boldsymbol{u}}_i$ is determined as

$$\tilde{u}_{(i-1)d+j} = p\left(||\tilde{\boldsymbol{x}}_i||_2^2 + \epsilon^2\right)^{p/2-1} x_{(i-1)d+j}$$

for $j = 1, 2, \ldots, d$ and $i = 1, 2, \ldots, N/d$ where $\tilde{\boldsymbol{x}}_i$ is the ith block of $\boldsymbol{x}$.

By examining the Hessian of $F_\epsilon(\boldsymbol{x})$, it can be shown that the region where $F_\epsilon(\boldsymbol{x})$ is convex is proportional to the value of ϵ, namely, the larger the ϵ, the larger the convex region. Thus if a sufficiently large value of ϵ is used, a gradient descent based algorithm will find the global solution of the problem in (9). On the other hand, the desired solution is the global minimizer of the objective function in (9) with a small value of ϵ. A good optimal solution can, therefore, be obtained by using a sequential optimization whereby a series of objective functions are minimized starting with a large value of ϵ and gradually decreasing ϵ to a very small value ϵ_T.

B. Use of Fletcher-Reeves algorithm

In the optimization procedure described above, the problem in (9) is solved for a set of values ϵ. For each value of ϵ, we propose to use a finite number of iterations of the Fletcher-Reeves algorithm [12] to minimize the objective function. The Fletcher-Reeves algorithm belongs to the class of conjugate gradient methods where search directions are conjugate directions computed based on the gradient of $F_\epsilon(\boldsymbol{x})$. In the kth iteration, iterate $\boldsymbol{x}_k$ is updated as

$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \alpha_k \boldsymbol{d}_k \quad \text{for} \quad k = 0, 1, \ldots, L-1 \tag{13}$$

where $\boldsymbol{d}_0$, $\boldsymbol{d}_1$, ..., $\boldsymbol{d}_{L-1}$ are L conjugate directions and α_0, α_1, ..., α_{L-1} are L step sizes. The conjugate directions are computed as

$$\boldsymbol{d}_k = \begin{cases} -\boldsymbol{g}_0 & \text{for} \quad k = 0 \\ -\boldsymbol{g}_k + \beta_{k-1}\boldsymbol{d}_{k-1} & \text{for} \quad k = 1, 2, \ldots, L-1 \end{cases} \tag{14}$$

where $\boldsymbol{g}_k$ is the gradient computed by using $\boldsymbol{x} = \boldsymbol{x}_k$ in (11) and $\beta_k = \beta_n/\beta_d$ where

$$\beta_n = \boldsymbol{g}_{k+1}^T \boldsymbol{g}_{k+1} \quad \text{and} \quad \beta_d = \boldsymbol{g}_k^T \boldsymbol{g}_k$$

for $k = 0, 1, \ldots, L-1$.

C. Line search

Given parameters p, ϵ, and λ, the step size α_k in the Fletcher-Reeves algorithm is obtained by solving the one-dimensional optimization problem

$$\underset{\alpha}{\text{minimize}} \quad f(\alpha)$$

where

$$f(\alpha) = \frac{1}{2}\,||\boldsymbol{\Phi}(\boldsymbol{x}_k + \alpha \boldsymbol{d}_k) - \boldsymbol{y}||_2^2 + \lambda\,||\boldsymbol{x}_k + \alpha \boldsymbol{d}_k||_{2/p,\epsilon}$$

By setting the first derivative of $f(\alpha)$ to zero, we obtain an equation of the form $\alpha = G(\alpha)$ where

$$G(\alpha) = -\frac{\boldsymbol{d}_k^T \boldsymbol{\Phi}^T (\boldsymbol{\Phi}\boldsymbol{x}_k - \boldsymbol{y}) + \lambda \cdot p \cdot \sum_{i=1}^{N/d} \gamma_i \cdot (\tilde{\boldsymbol{x}}_{ki}^T \tilde{\boldsymbol{d}}_{ki})}{||\boldsymbol{\Phi}\boldsymbol{d}_k||_2^2 + \lambda \cdot p \cdot \sum_{i=1}^{N/d} \gamma_i \cdot (\tilde{\boldsymbol{d}}_{ki}^T \tilde{\boldsymbol{d}}_{ki})}$$

In $G(\alpha)$, $\tilde{\boldsymbol{x}}_{ki}$ and $\tilde{\boldsymbol{d}}_{ki}$ are the ith blocks of vectors $\boldsymbol{x}_k$ and $\boldsymbol{d}_k$, respectively, and

$$\gamma_i = \left(||\tilde{\boldsymbol{x}}_i + \alpha \tilde{\boldsymbol{d}}_i||_2^2 + \epsilon^2\right)^{p/2-1} \quad \text{for} \quad i = 1, 2, \ldots, N/d$$

Therefore, finding a minimizer of $f(\alpha)$ amounts to finding a fixed point of function $G(\alpha)$ which, according to the Banach fixed-point theorem [14], can be done by using a sufficient number of recursions of the recursive relation

$$\alpha_{l+1} = G(\alpha_l) \quad \text{for} \quad l = 1, 2, \ldots \tag{15}$$

D. Algorithm

The proposed $\ell_{2/p}$-regularized least-squares ($\ell_{2/p}$-RLS) algorithm is summarized in Table I. Parameter p, the number of iterations T, the length of block d, the initial value ϵ_1 and target value ϵ_T of ϵ, and parameter λ are supplied as input.

A total of $J - 2$ values of ϵ, for which the optimization in (9) is carried out, are set between the initial value ϵ_1 and target value ϵ_T as

$$\epsilon_t = \epsilon_1 e^{-\beta(t-1)} \quad \text{for } t = 2, 3, \ldots, T-1 \tag{16}$$

where $\beta = \log(\epsilon_1/\epsilon_T)/(T-1)$. The initial conjugate direction is set to $-\boldsymbol{g}$ at the beginning of the optimization for each value of ϵ.

For noise-free measurement $\boldsymbol{y}$, a large initial value λ_1 and a small target value λ_T are supplied as input instead of λ. A total of $T - 2$ values of λ lying between λ_1 and λ_T are computed as

$$\lambda_t = \lambda_1 e^{-\sigma(t-1)} \quad \text{for } t = 2, 3, \ldots, T-1 \tag{17}$$

where $\sigma = \log(\lambda_1/\lambda_T)/(T-1)$.

TABLE I
$\ell_{2/p}$-RLS ALGORITHM

Step 1
Input: p, T, L, ϵ_1, ϵ_T, $\boldsymbol{\Phi}$, $\boldsymbol{y}$, E_t.
  λ if measurement $\boldsymbol{y}$ is noisy
  λ_1 and λ_T if measurement $\boldsymbol{y}$ is noiseless.
Set $\boldsymbol{x}_s = \boldsymbol{0}$.
Step 2
Compute ϵ_t for $t = 2, 3, \ldots, T-1$ using (16).
  λ_t for $t = 2, 3, \ldots, T-1$ using (17)
  if measurement $\boldsymbol{y}$ is noiseless.
Step 3
For $t = 1, \ldots, T$
  i) Set $\epsilon = \epsilon_t$, $L_t = 3 + round(t/4)$.
  ii) If measurement is noiseless, set $\lambda = \lambda_t$.
  iii) Set $k = 0$, $\boldsymbol{x}_0 = \boldsymbol{x}_s$, $E_r = 10^{10}$.
  iv) While $E_r > E_t$,
    a) Compute $\boldsymbol{g}_k$ using (11).
    b) Compute $\boldsymbol{d}_k$ using (14).
    c) Compute α_k using (15).
    d) Compute $\boldsymbol{x}_{k+1}$ using (13).
    e) Set $k = k + 1$.
    f) Exit loop if $k > L_t$.
    g) Compute $E_r = ||\alpha_k \boldsymbol{d}_k||_2$.
  viii) Set $\boldsymbol{x}_s = \boldsymbol{x}_{L_t}$.
Step 4
Output $\boldsymbol{x}^* = \boldsymbol{x}_s$ and stop.

IV. EXPERIMENTAL RESULTS

To demonstrate the effectiveness of the proposed method, we have carried two experiments, as detailed below.

In the first experiment, the signal length N, the number of measurements M, and block length d were set to 512, 100, and 8, respectively. A total of sixteen block-sparsity levels $K = 1, 2, \ldots, 16$ were chosen. A K-block sparse signal $\boldsymbol{x}$ was constructed by assigning random values drawn from a normal distribution $\mathcal{N}(0, 1)$ to all the components of K randomly selected blocks of a zero vector of length N. Measurement matrix $\boldsymbol{\Phi}$ of size $M \times N$ was constructed by drawing its elements from $\mathcal{N}(0, 1)$ followed by an orthonormalization step where the rows of $\boldsymbol{\Phi}$ were made orthonormal to each other. The measurement was obtained as $\boldsymbol{y} = \boldsymbol{\Phi}\boldsymbol{x}$. With $p = 0.1$, $T = 80$, $\epsilon_1 = 1$, $\epsilon_T = 1e-5$, $\lambda_1 = 1$, $\lambda_T = 1e-10$, and $E_t = 1e-25$, the $\ell_{2/p}$-RLS algorithm was applied and compared with $\ell_{2/1}$-norm minimization using $\ell_{2/1}$ SOCP [9], iterative re-weighted (IR) with $p = 0.1$ [6], smoothed ℓ_0 norm (SL0) [7], BOMP [10], and BP [4] algorithms. Reconstruction was deemed successful if the maximum absolute error between the original signal, $\boldsymbol{x}$, and the recovered signal, $\hat{\boldsymbol{x}}$, measured as $\max_i |x_i - \hat{x}_i|$ was smaller than 0.09, where x_i and $\hat{x}_i$ are the ith components of $\boldsymbol{x}$ and $\hat{\boldsymbol{x}}$, respectively. The percentage of the number of successful reconstructions over 100 runs is plotted in Fig. 1. It is observed that the performance of the $\ell_{2/p}$-RLS algorithm is significantly better than that of the other algorithms.

In the second experiment, the average CPU time required by the algorithms to converge was measured over 100 runs for typical instances with M = round$(N/2)$, K = round$(M/2.5d)$, and $d = 8$ for N = 128, 256, 512, 1024, 2048, 4096, and 8192. The CPU time was measured on a PC desktop with an Intel Core 2 CPU E6850 3.00 GHz processor. The CPU times for the six algorithms are plotted in Fig. 2 for values of N in the range 123 to 1024. The CPU times for the $\ell_{2/p}$-RLS algorithm with $p = 0.1$, and the SL0, BP, and BOMP algorithms for values of N in the range 1024 to 8192 are plotted in Fig. 3. As can be seen, for $N > 5000$ the proposed $\ell_{2/p}$-RLS algorithm requires the least amount of computation among the algorithms tested.

V. CONCLUSION

We have proposed an algorithm for the reconstruction of block-sparse signals in the CS framework. The algorithm minimizes an $\ell_{2/p}$-norm regularized ℓ_2 error with $p < 1$ by using

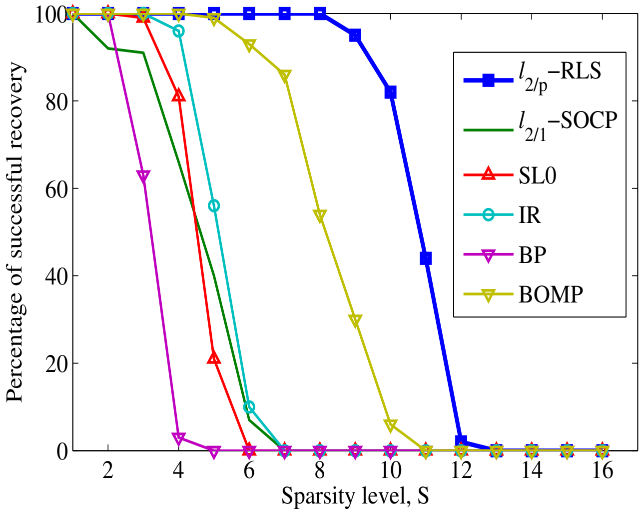

Fig. 1. Percentage of perfect reconstructions for $\ell_2, \ell_{2/p}$-RLS ($p = 0.1$), $\ell_{2/1}$-SOCP, SL0, IR ($p = 0.1$), BP, and BOMP algorithms over 100 runs with $N = 512$, $M = 100$, $d = 8$.

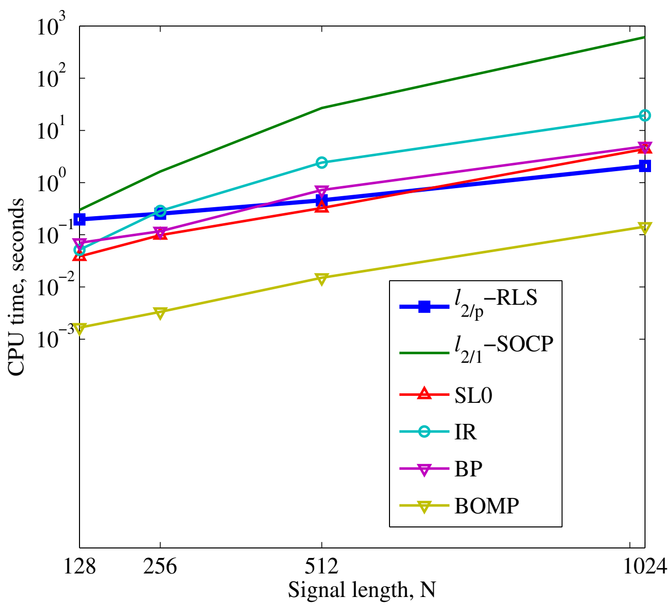

Fig. 2. Average CPU time required for $\ell_{2,p}$-RLS ($p = 0.1$), $\ell_{2/1}$-SOCP, SL0, IR ($p = 0.1$), BP, and BOMP algorithms over 100 runs with $M = N/2$, $K = M/2.5d$.

a sequential optimization in conjunction with the Fletcher-Reeves algorithm. Simulation results show that the proposed algorithm yields significantly improved signal reconstruction performance and requires a reduced amount of computation for large sized data relative to several contemporary competing algorithms.

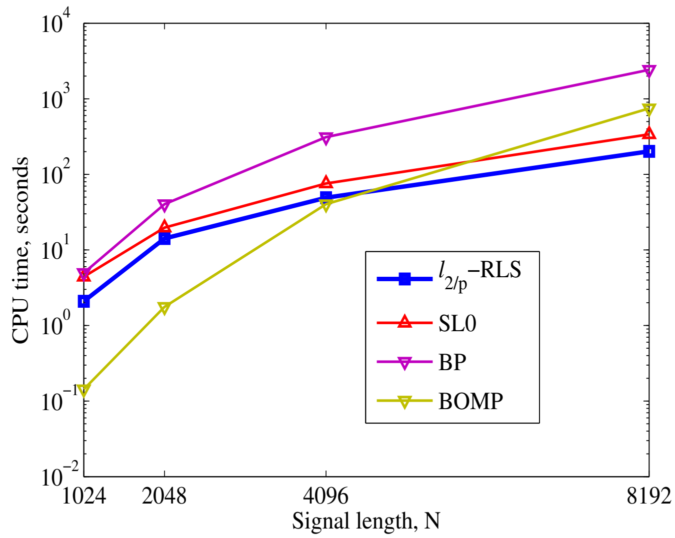

Fig. 3. Average CPU time required for $\ell_{2/p}$-RLS ($p = 0.1$), SL0, BP, and BOMP algorithms over 100 runs with $M = N/2$, $K = M/2.5d$.

ACKNOWLEDGEMENT

The authors would like to thank the Natural Sciences and Engineering Research Council of Canada for supporting this research.

REFERENCES

[1] E. 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. 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. Scientif. Comput.*, vol. 20, no. 1, pp. 33-61, 1999.
[5] R. Chartrand, “Exact reconstruction of sparse signals via nonconvex minimization,” *IEEE Signal Process. Lett.*, vol. 14, no. 10, pp. 770-710, Oct. 2007.
[6] R. Chartrand and W. Yin, “Iteratively reweighted algorithms for compressive sensing,” in Proc. *IEEE Inter. Conf. Acoustics, Speech, Signal Process.*, pp. 3869-3872, 2008.
[7] 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.
[8] M. Stojnic, F. Parvaresh, and B. Hassibi, “On the reconstruction of block-sparse signals with an optimal number of measurements,” *IEEE Trans. Signal Process.*, vol. 57, no. 8, pp. 3075-3085, Aug. 2009.
[9] Y. C. Eldar and M. Mishali, “Robust recovery of signals from a structured union of subspaces,” *IEEE Trans. Inf. Theory*, vol. 55, no. 11, Nov. 2009.
[10] Y. C. Eldar, P. Kuppinger, and H. Bolcskei, “Block-sparse signals: uncertainty relations and efficient recovery,” vol. 58, no. 6, pp. 3042-3054, Jun. 2010.
[11] J. A. Tropp and A. C. Gilbert, “Signal recovery from random measurements via orthogonal matching pursuit,” *IEEE Trans. Inf. Theory*, vol. 53, no. 12, pp. 4655-4666, Dec. 2007.
[12] A. Antoniou and W.-S. Lu, *Practical Optimization: Algorithms and Engineering Applications,* Springer, 2006.
[13] G. H. Golub and C. F. V. Loan, *Matrix Computations,* The Johns Hopkins University Press, 3rd ed., 1996.
[14] R. P. Agrawal, M. Meehan, and D. O’Regan, *Fixed Point Theory and Applications,* Cambridge University Press, 2001.