A Neurodynamic Optimization Approach to Constrained Sparsity Maximization Based on Alternative Objective Functions

Zhishan Guo and Jun Wang

Abstract— **In recent years, constrained sparsity maximization problems received tremendous attention in the context of compressive sensing. Because the formulated constrained L_0 norm minimization problem is NP-hard, constrained L_1 norm minimization is usually used to compute approximate sparse solutions. In this paper, we introduce several alternative objective functions, such as weighted L_1 norm, Laplacian, hyperbolic secant, and Gaussian functions, as approximations of the L_0 norm. A one-layer recurrent neural network is applied to compute the optimal solutions to the reformulated constrained minimization problems subject to equality constraints. Simulation results in terms of time responses, phase diagrams, and tabular data are provided to demonstrate the superior performance of the proposed neurodynamic optimization approach to constrained sparsity maximization based on the problem reformulations.**

I. Introduction

Many signals are compressible which can be well-approximated by signals that have only a few non-zero variables on a suitable basis [1][2]. Effect sparsity [3] of a compressible signal is a measurement of the few non-zero coefficients that actually affect its responses. Compressive sensing [4][5] is a method which employs these few large coefficients that preserve the structure of the signal; the signal is then reconstructed from the constraints using an optimization method. This, in some sense, overcomes the limit of Shannon's sampling theorem omnipresent in electrical engineering: A signal has to be sampled at the Nyquist rate (i.e., proportional to its highest frequency) in order to be reconstructed perfectly. Various potential applications in engineering have been shown; e.g., [6][7][8].

Mathematically speaking, we would like to determine an object $x_0 \in \mathcal{R}^n$ from data $y = Ax_0$, where A is an $m \times n$ matrix with fewer rows than columns; i.e., $m < n$. A system with fewer equations than unknowns usually has an infinite number of solutions. An important criterion is to find the most sparse one subject to constraints. Ideally, we want to solve the following combinatorial optimization problem:

$$\begin{array}{ll} \text{minimize} & \|x\|_0, \\ \text{subject to} & Ax = b, \end{array} \tag{1}$$

where $x \in \mathcal{R}^n$, $A \in \mathcal{R}^{m\times n}$, and $\|x\|_0$ is defined as the number of non-zero elements of x.

Z. Guo and J. Wang are with the Department of Mechanical and Automation Engineering, The Chinese University of Hong Kong, Shatin, New Territories, Hong Kong (email: {zsguo, jwang}@mae.cuhk.edu.hk).

The work described in this paper was supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region, China (Project no. CUHK417608E).

The constrained sparsity maximization problem (1) is NP complete, since its solution usually requires an intractable combinatorial search [9]. As a result, a convex relaxation of the problem is usually solved instead [4]:

$$\begin{array}{ll} \text{minimize} & \|x\|_1, \\ \text{subject to} & Ax = b, \end{array} \tag{2}$$

where $\|x\|_1 = \sum_{i=1}^{n} |x_i|$. It was shown that if A has some restricted isometry property, the solutions to (1) and (2) are identical [11][12]. However, restricted isometry property may be too strong for practical basis design matrix A to hold [26], and thus (2) may not be a good formulation to the original problem (1). Many attempts on improving L_1 norm minimization have been done, such as adding weights [20][21][22], solving iteratively [23][24], using L_p norm $(p < 1)$ instead [25][26].

Since Tank and Hopfield's pioneering work on a neural network approaches to linear programming [13], many results have been reported in neurodynymic optimization. For example, Zhang and Constantinides [14] proposed the Lagrangian network for solving nonlinear programming problems with equality constraints. Forti et al. [15] proposed a generalized neural network for solving non-smooth nonlinear programming problems based on the gradient method. Recently, several recurrent neural networks for solving linear and quadratic programming problems with discontinuous activation functions have been proposed [16][17][18][19]. In particular, in [18], a two-layer recurrent neural network is presented for solving non-smooth convex optimization subject to linear equality and bound constraints. A one-layer recurrent neural network for solving the same problem is presented in [19].

In this paper, we present a neurodynamic optimization approach to constrained sparsity maximization based on alternative objective functions, such as weighted L_1 norm, Laplacian, hyperbolic secant, and Gaussian functions. The experimental results herein show that neurodynamic optimization based on the alternative objective functions outperforms L_1 norm minimization. In particular, successive Gaussian maximization performs the best, which obtains solution of the highest sparsity with the highest parametric robustness.

II. Problem Reformulation

In this section, four alternative objective functions are proposed as the approximations of the L_0 norm. Let

978-1-4244-8126-2/10/$26.00 ©2010 IEEE

Authorized licensed use limited to: N.C. State University Libraries - Acquisitions & Discovery S. Downloaded on April 28,2023 at 17:44:33 UTC from IEEE Xplore. Restrictions apply.

$$\delta(x_i) = \begin{cases} -1, & x_i = 0; \\ 0, & x_i \neq 0. \end{cases} \quad (3)$$

According to its definition, L_0 norm is equivalent to $\sum_{i=1}^{n} \delta(x_i)$.

Weighted L_1 norm minimization adds weights to each element of decision variable x. Instead of minimizing $\|x\|_1$ in (2), the following function is minimized:

$$f_w(x) = \sum_{i=1}^{n} a_i^* |x_i|, \quad (4)$$

where $a_i^* = \max_j |a_{ji}|$. Note that multiplying any positive constant will never change the convexity of a function. By multiplying the maximum element of each column of the matrix A, the sensitivity of solution may increase, while the convexity of the optimal function is maintained. Previous experiments [4][5][9][10][11][12] have shown that under many levels of sparsity and undeterminedness, the solution to (2) may be far away from the ones to (1). A problem for the solution $\hat{x}$ to (2) is that some elements in x with value of zero turn out to be non-zero elements to (1). Thus, multiplying bigger weights to these elements may increase the penalty to force them to be smaller.

To approximate $\delta(\cdot)$, we propose a group of inverted Laplacian functions, inverted hyperbolic secant functions, and inverted Gaussian functions:

$$f_l(x, \sigma_k) = -\sum_{i=1}^{n} \exp\left(-\frac{|x_i|}{\sigma_k}\right), \quad (5)$$

$$f_h(x, \sigma_k) = -2\sum_{i=1}^{n} \left[\exp\left(\frac{x_i}{\sigma_k}\right) + \exp\left(-\frac{x_i}{\sigma_k}\right)\right]^{-1}, \quad (6)$$

$$f_g(x, \sigma_k) = -\sum_{i=1}^{n} \exp\left(-\frac{x_i^2}{\sigma_k^2}\right); \quad (7)$$

where $k = 0, 1, 2, ...$; and $\sigma_0 > \sigma_1 > ... > 0$ for (5)-(7).

The three functions are shown in Figure 1. The approximation accuracy to L_0 norm by the given functions $f_l(\cdot)$, $f_h(\cdot)$, and $f_g(\cdot)$ varies with different values of σ_k; the smaller σ_k is, the closer these functions approach $\delta(\cdot)$, as shown in Figure 2 for the inverted Gaussian functions.

In this paper, the objective functions (4)-(7) are minimized successively. Hence the solutions will converge as inverted functions $f_l(\cdot)$, $f_h(\cdot)$, and $f_g(\cdot)$ narrow down and sequentially approach to the solutions to the L_0 norm.

Step 0: Set $k = 0$, σ_0 sufficiently large;

Step 1: Minimize (5), (6), or (7), subject to constraints;

Step 2: $\sigma_k = \Delta(\sigma_{k-1})$;

Step 3: If $\sigma_k < \sigma_{\min}$, end; else $k = k + 1$, go to step 1.

Here $\Delta(\cdot)$ is a decreasing function, such as $\Delta(\sigma) = \sigma - \delta_\sigma$, or $\Delta(\sigma) = \sigma/M$, where M is a positive number.

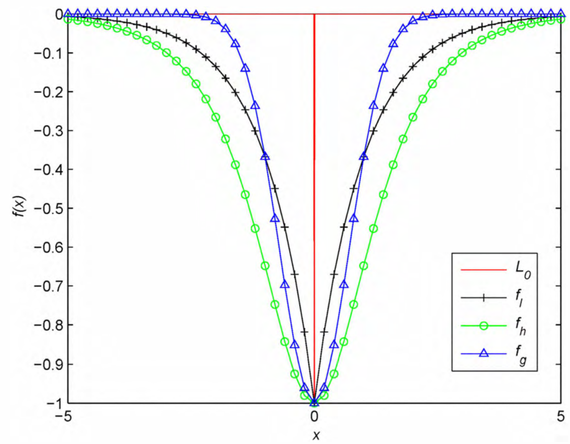

Fig. 1. Alternative objective functions $\delta(\cdot)$, $f_l(\cdot)$, $f_h(\cdot)$, and $f_g(\cdot)$ with $\sigma_k = 1$.

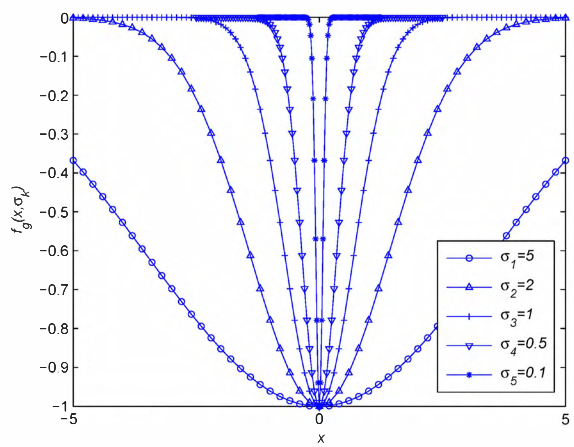

Fig. 2. A group of inverted Gaussian functions $f_g(x, \sigma_k)$ with several values of σ_k.

III. Model Description

In this section, a one-layer recurrent neural network [19] is introduced for constrained sparsity maximization based on the alternative objective functions (4)-(7). The neural network is described by the following differential equation:

$$\epsilon\frac{dx}{dt} = -Px - (I - P)\partial f(x) + q, \quad (8)$$

where x is the state vector, ϵ is a positive scaling constant, I is the identity matrix, $P = A^T(AA^T)^{-1}A$, $q = A^T(AA^T)^{-1}b$, and $\partial f(x)$ is the differential of $f(x)$ (i.e., (4), (5), (6), or (7)).

Specifically, to minimize objective functions (4)-(7), the one-layer recurrent neural network is describe by the following equations.

$$\epsilon\frac{dx}{dt} = -Px - (I - P)g_w(x) + q, \quad (9)$$

$$\epsilon\frac{dx}{dt} = -Px - (I - P)g_l(x, \sigma_k) + q, \quad (10)$$

Authorized licensed use limited to: N.C. State University Libraries - Acquisitions & Discovery S. Downloaded on April 28,2023 at 17:44:33 UTC from IEEE Xplore. Restrictions apply.

$$\epsilon \frac{dx}{dt} = -Px - (I-P)g_h(x, \sigma_k) + q, \quad (11)$$

$$\epsilon \frac{dx}{dt} = -Px - (I-P)g_g(x, \sigma_k) + q; \quad (12)$$

where $g_w(\cdot)$, $g_l(\cdot)$, $g_h(\cdot)$, and $g_g(\cdot)$ are vectors of activation functions with the ith element given by the following formulas, respectively, for $i = 1, ..., n$; as shown in Figure 3.

$$g_w(x)_i = a_i^* \mathrm{sgn}(x_i), \quad (13)$$

$$g_l(x, \sigma_k)_i = \frac{1}{\sigma_k} \mathrm{sgn}(x_i) \exp\left(-\frac{|x_i|}{\sigma_k}\right), \quad (14)$$

$$g_h(x, \sigma_k)_i = \frac{2}{\sigma_k} \frac{\exp\left(\frac{x_i}{\sigma_k}\right) - \exp\left(-\frac{x_i}{\sigma_k}\right)}{\left[\exp\left(\frac{x_i}{\sigma_k}\right) + \exp\left(-\frac{x_i}{\sigma_k}\right)\right]^2}, \quad (15)$$

$$g_g(x, \sigma_k)_i = 2\frac{x_i}{\sigma_k^2} \exp\left(-\frac{x_i^2}{\sigma_k^2}\right); \quad (16)$$

where $\mathrm{sgn}(x)$ is the sign function which return 1 if $x > 0$, -1 if $x < 0$, and 0 if $x = 0$.

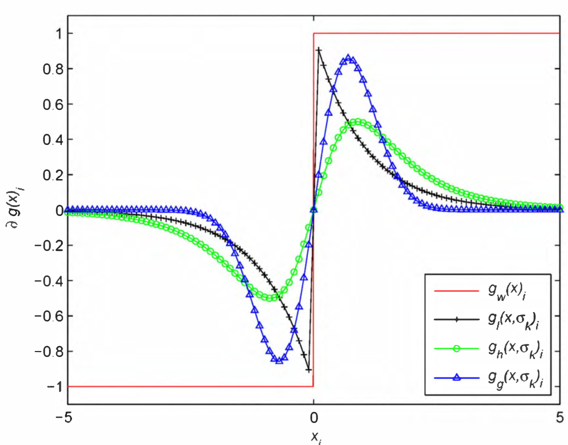

Fig. 3. Activation functions in (9)-(12) where $a_j^* = 1$ and $\sigma_k = 1$.

IV. Convergent Behaviors

In this section, the convergence behaviors of the one-layer recurrent neural networks (9)-(12) are shown in terms of their state variables and two performance indices by using the simulation results of randomly generated examples.

A performance index is given in [9] as:

$$\gamma = \frac{\|x - x^*\|_2}{\|x^*\|_2}, \quad (17)$$

where x is a solution to (2) and x^* is a known solution to (1).

Since sometimes (1) may have different sparsest solutions, γ may not represent sparsity well, especially for tasks other than signal reconstruction. In this paper, we also use another performance index, sparsity level ξ, to evaluate the performance of solutions:

$$\xi = \log \frac{\|x\|_0}{\|x^*\|_0}. \quad (18)$$

In order to compare fairly, we specified the problem to be $b = Ax$, where x has k non-zero elements (e.g., $k = 5$) drawn randomly from the uniform distribution over $(0, 100)$; $n = 100, m = 20$ fixed and each a_{ij} drawn from the normal distribution $N(0, 1)$. The parameter ϵ is fixed as 10^{-6}, and 10^{-7} in the one-layer recurrent neural network (9)-(11), and (12) respectively. Figure 4 shows a convergence process of the weighted L_1 norm minimization based on the neural network, and Figure 5 illustrates the performance indices.

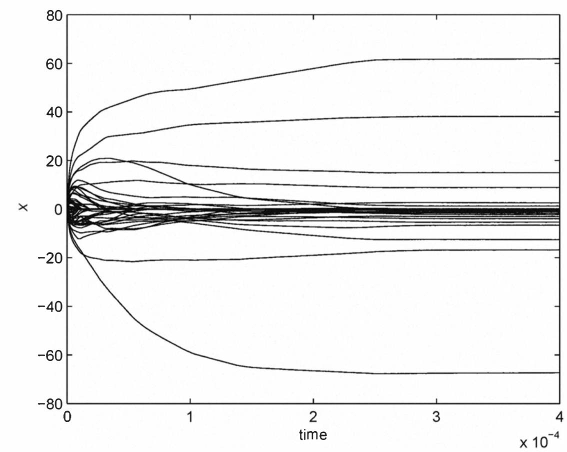

Fig. 4. Transient states of the one-layer recurrent neural network (9) for solving a randomly generated weighted L_1 norm minimization problem.

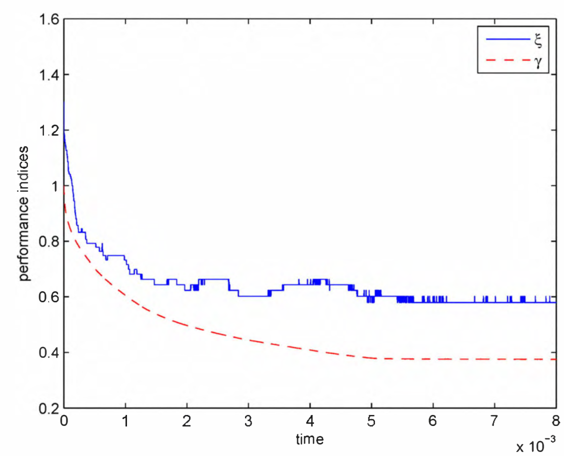

Fig. 5. Transients of the performance indices in weighted L_1 norm minimization using the one-layer recurrent neural network (9).

Simulation results with two fixed values σ_k for $f_g(x, \sigma_k)$ of system (12) are shown in Figures 6 and 7. It is obvious that both processes converge to an unsatisfactory state with low sparsity level. When σ_k is small; i.e., $\sigma_k = 0.5$ in Figure 6, though inverted Gaussian function approaches $\delta(\cdot)$ well, the system converges too slowly and many elements stay far

Authorized licensed use limited to: N.C. State University Libraries - Acquisitions & Discovery S. Downloaded on April 28,2023 at 17:44:33 UTC from IEEE Xplore. Restrictions apply.

away from zero (i.e., $> 3\sigma_k$) where the value of activation function is almost zero in a short period of time. If a larger value of σ_k is selected and fixed; i.e., $\sigma_k = 20$ in Figure 7, $g_g(x, \sigma_k)$ does not approximate $\delta(x)$ well, and many non-zero elements stay inside the region $(-10, 10)$. As a result, x is far from x^*. On the other hand, each of them has benefits: When σ_k is large, the system will obtain a most sparsity solution by gathering all most likely elements near zero, while when σ_k is small, the neurodynamics pull the pseudo sparse elements away from zero. As shown in Section 2, smaller σ_k leads to better approximation to L_0 norm, and at the same time, it takes much more time for the neurodynamic system to converge to a stable state. A simple idea is to iteratively reduce the value of σ_k as system converges, which is the procedure stated in the end of Section 2. Thus, the system converges fast all the time and approximates L_0 norm well in the end.

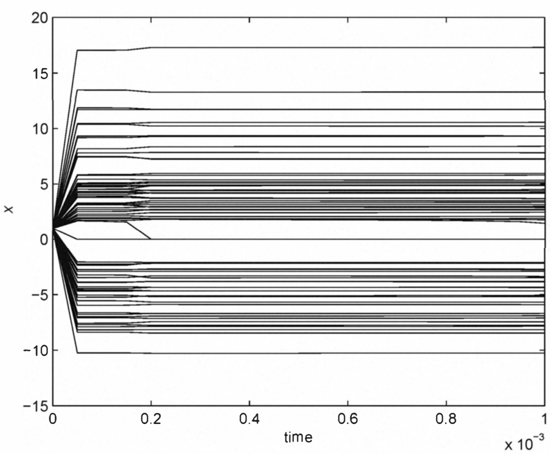

Fig. 6. Transient states of the one-layer recurrent neural network (12) with $\sigma_k = 0.5$ for solving a randomly generated problem.

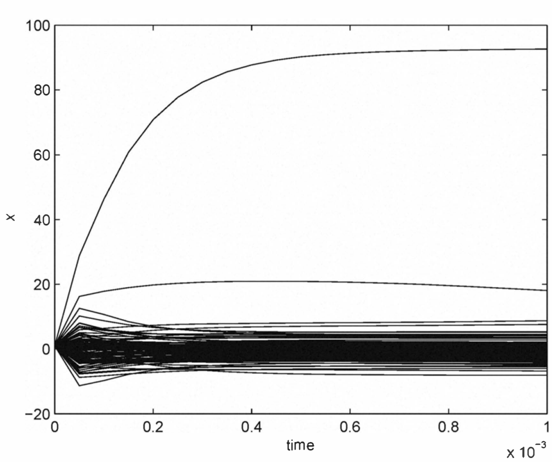

Fig. 7. Transient states of the one-layer recurrent neural network (12) with $\sigma_k = 20$ for solving a randomly generated problem.

Figure 8 shows the transient states with $\sigma_0 = 20$, $\Delta(\sigma) = \sigma/2$, and $\sigma_{\min} = 0.5$, and Figure 9 shows the performance indices. In Figure 9, the curves are not smooth because σ_k changes at $t = 4, 8, \ldots \times 10^{-4}$. These two figures show that some elements are pulled away from zero as σ_k deceases, and finally all the states are stable, which means the sensitivity of the neurodynamics to the pseudo sparse elements increases as σ_k deceases.

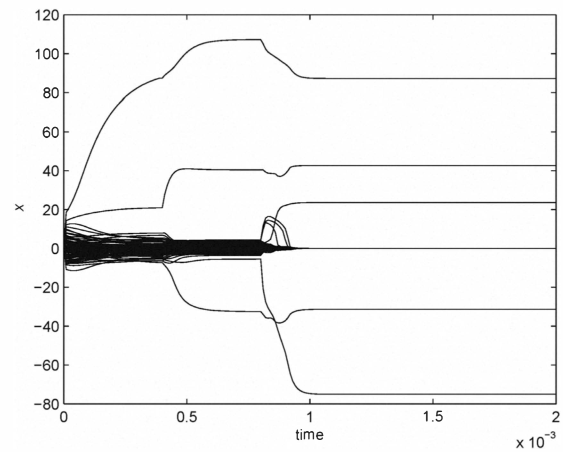

Fig. 8. Transient states of the one-layer recurrent neural network (12) with $\sigma_0 = 20$, $\Delta(\sigma) = \sigma/2$, and $\sigma_{\min} = 0.5$ for solving a randomly generated problem.

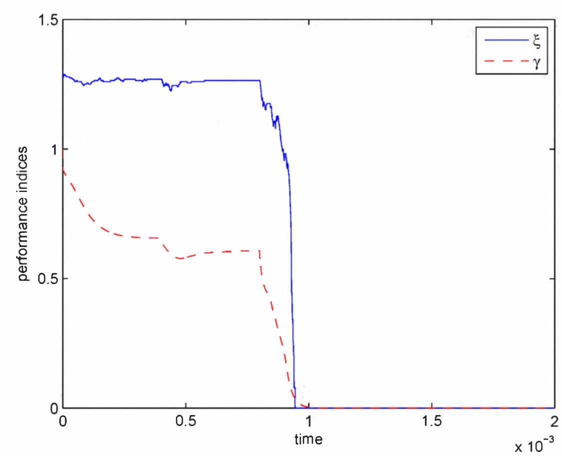

Fig. 9. Transients of the performance indices using the one-layer recurrent neural network (12) with $\sigma_0 = 20$, $\Delta(\sigma) = \sigma/2$, and $\sigma_{\min} = 0.5$.

Figures 10-13 show the transient states x and performance indices γ and ξ of the neural networks (10) and (11) for solving randomly generated problems based on (5) and (6), respectively.

According to Theorem 3 in [18], the one-layer neural network is globally convergent if $f(x)$ is convex. Figures 4-13 show that the one-layer recurrent neural network (9) converges to global minimum for the weighted L_1 norm minimization, since the weighted L_1 norm is convex. For the other three quasi-convex objective functions, the given

Authorized licensed use limited to: N.C. State University Libraries - Acquisitions & Discovery S. Downloaded on April 28,2023 at 17:44:33 UTC from IEEE Xplore. Restrictions apply.

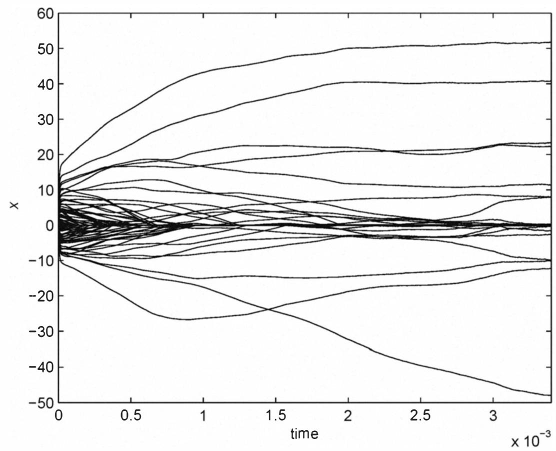

Fig. 10. Transient states of the one-layer recurrent neural network (10) with $\sigma_0 = 20$, $\Delta(\sigma) = \sigma/2$, and $\sigma_{min} = 3$ for solving a randomly generated problem.

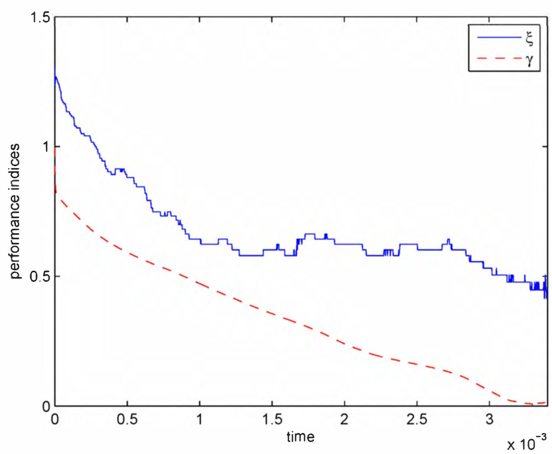

Fig. 11. Transients of the performance indices using the one-layer recurrent neural network (10) with $\sigma_0 = 20$, $\Delta(\sigma) = \sigma/2$, and $\sigma_{\min} = 3$.

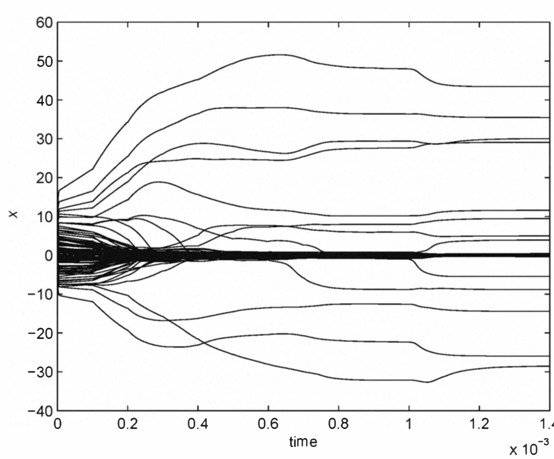

Fig. 12. Transient states of the one-layer recurrent neural network (11) with $\sigma_0 = 10$, $\Delta(\sigma) = \sigma/2$, and $\sigma_{\min} = 0.5$ for solving a randomly generated problem.

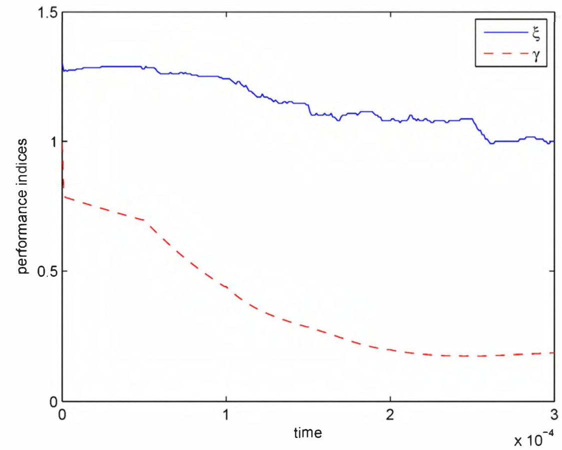

Fig. 13. Transients of the performance indices using the one-layer recurrent neural network (11) with $\sigma_0 = 10$, $\Delta(\sigma) = \sigma/2$, and $\sigma_{\min} = 0.5$.

network is not guaranteed to converge to global minimum. However, convexity of the objective function is only a sufficient condition, with the neural network (10)-(12) still converges for suitable parameters.

V. Experimental Results

In this section, experimental results are reported and compared in terms of the predefined performance indices by using phase diagrams and tabular data. We show how undeterminedness m/n and sparsity level k/m affect the performance indices of our algorithms. In the experiments, the performance indices are shown with various combinations of undeterminedness and sparsity level with the following procedure:

$Step\,0$: Set n (e.g., $n = 100$), upper bound of loop round N (e.g., $N = 20$), iteration counter $s = 0$;

$Step\,1$: Generate a problem randomly: $b = Ax$, where A is drawn from the normal distribution $N(0, 1)$, and x is sparse; i.e., has k non-zero elements drawn from the uniform distribution over $(0, 100)$;

$Step\,2$: Use the the one-layer recurrent neural network (9)-(12) with parameter ϵ (e.g., 10^{-6}, or 10^{-7}) to solve (4)-(6), or (7), obtaining $\hat{x}$;

$Step\,3$: Evaluate performance indices γ in (17) and ξ in (18);

$Step\,4$: If $s \le N$, $s = s + 1$, go to Step 1. Else, calculate the mean value of γ and ξ which is the final performance indices of a sample (k/m,m/n).

Samples with undeterminedness and sparsity level of 1% and $5l\%$ (where l is a positive integer, and $l \le 20$) are computed, while the performance indices of the rest combinations of m/n and k/m are fitting results by cubic spline interpolation. By averaging the results over numerous Monte-Carlo tests, phase diagram shows the performance indices of all different levels of sparsity and indeterminacy, thus illustrates how the success of L_1 optimization (2) is affected by sparsity and indeterminacy [10].

Authorized licensed use limited to: N.C. State University Libraries - Acquisitions & Discovery S. Downloaded on April 28,2023 at 17:44:33 UTC from IEEE Xplore. Restrictions apply.

Phase diagrams of performance index resulted from the neural network with different activation functions are shown in Figures 14-18. From the figures, we can see that basically all five cases have the same performance shape, which are all consistent with the boundary of the L_1 approximation (2) to L_0 optimization (1).

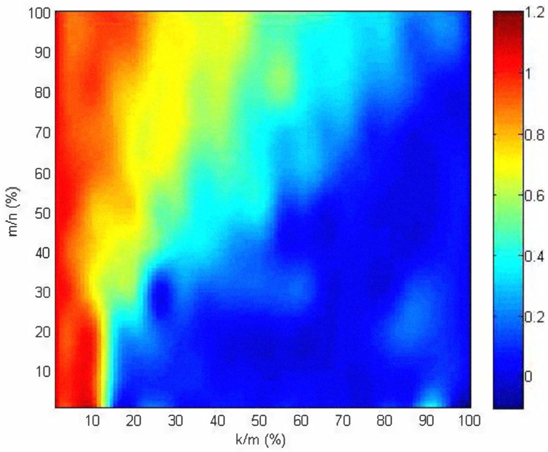

Fig. 14. Phase diagram of the solutions to L_1 norm minimization (2) in terms of performance index γ.

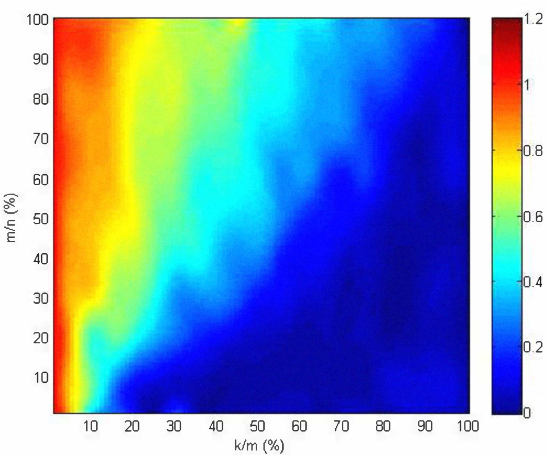

Fig. 15. Phase diagram of the solutions to weighted L_1 norm minimization (4) in terms of performance index γ.

The area in phase diagrams with performance index $\gamma \leq 0.3$ (blue color blocks) can be compared. The weighted L_1 norm minimization method results in an increase of approximation area by about 5%, while successive Gaussian maximization method increases the area by more than 15%. By using the proposed models, we can obtain a nearer solution to the constrained maximization problem (1) than L_1 minimization.

Figures 19-23 show phase diagrams of performance index ξ. For easy comparisons, we use the same color-bar for these three phase diagrams. For this purpose, performance indices

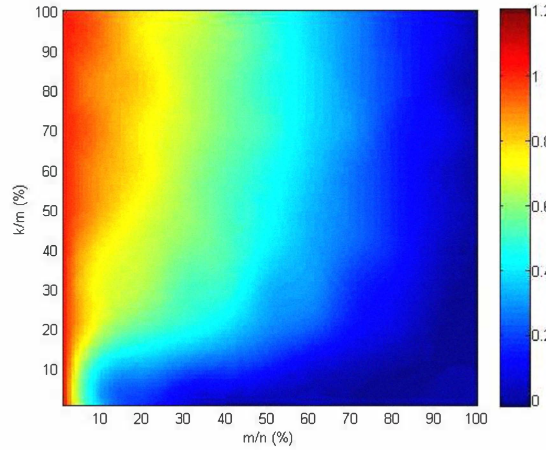

Fig. 16. Phase diagram of the solutions to successive inverted Laplacian minimization (5) in terms of performance index γ.

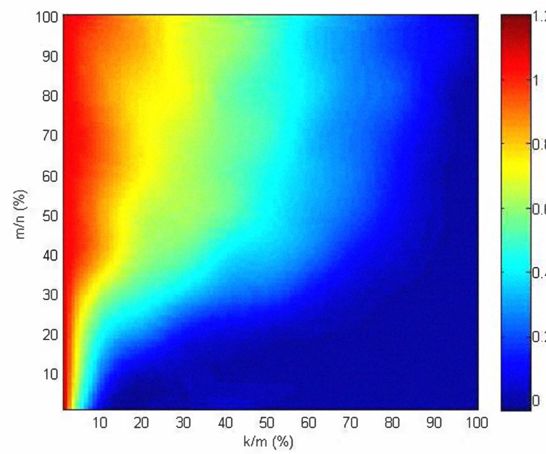

Fig. 17. Phase diagram of the solutions to successive inverted hyperbolic secant minimization (6) in terms of performance index γ.

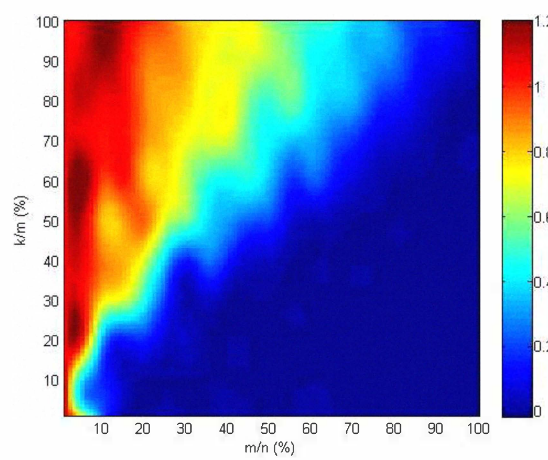

Fig. 18. Phase diagram of the solutions to successive inverted Gaussian minimization (7) in terms of performance index γ.

Authorized licensed use limited to: N.C. State University Libraries - Acquisitions & Discovery S. Downloaded on April 28,2023 at 17:44:33 UTC from IEEE Xplore. Restrictions apply.

of the blocks with $\xi \geq 1$ are set to be 1. Fortunately, except for weighted L_1 norm minimization, ξ rarely reach 1.

Figures 19 and 20 show similar patterns like a branch of hyperbola. The shape indicates that the solutions to both L_1 minimization and weighted L_1 norm minimization may not be very sparse when undeterminedness or sparsity level is low. Even for undeterminedness and sparsity level are both 0.5, ξ in these two phase diagrams is only around 0.5. From the definition of ξ, $\xi = 0.5$ means that the sparsity level of $\hat{x}$ is about 3 times larger than that of x.

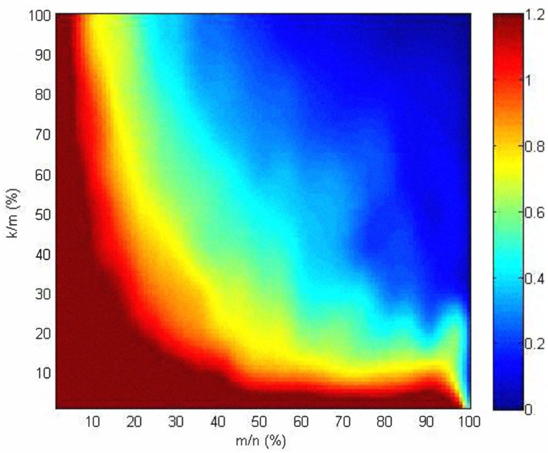

Fig. 19. Phase diagram of the solutions to L_1 norm minimization (2) in terms of performance index ξ.

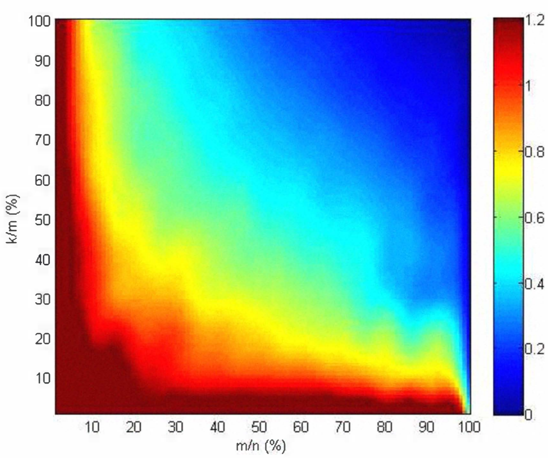

Fig. 20. Phase diagram of the solutions to weighted L_1 norm minimization (4) in terms of performance index ξ.

Figures 21-23 show that the models (10)-(12) can obtain a quite sparse solution mostly, with the successive Gaussian maximization performing the best. From Figure 23, whatever values of undeterminedness and sparsity level are at, the solutions for (12) have high sparsity levels. Sometimes these three methods may even obtain sparser solutions than x, where $\xi \leq 0$ in the phase diagrams.

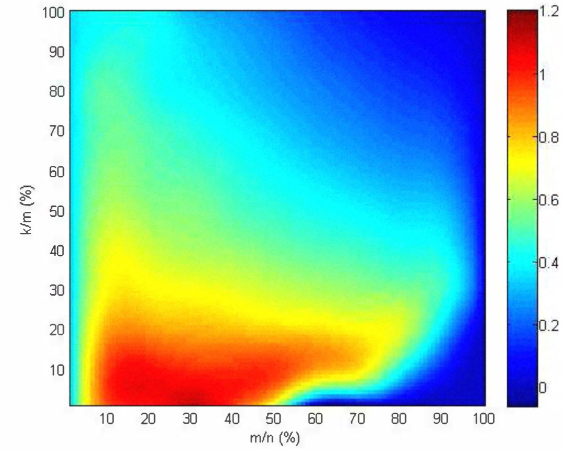

Fig. 21. Phase diagram of the solutions to successive inverted Laplacian minimization (5) in terms of performance index ξ.

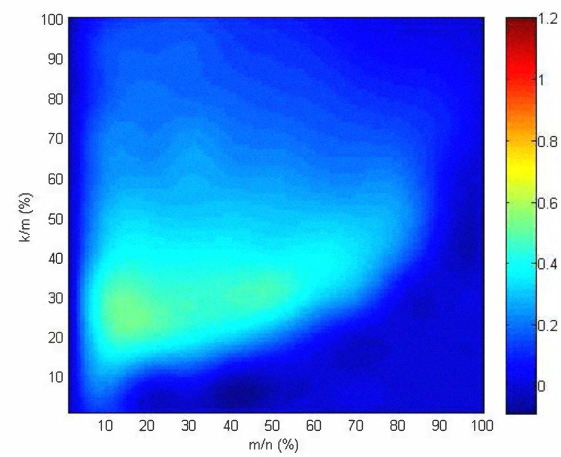

Fig. 22. Phase diagram of the solutions to successive inverted hyperbolic secant minimization (6) in terms of performance index ξ.

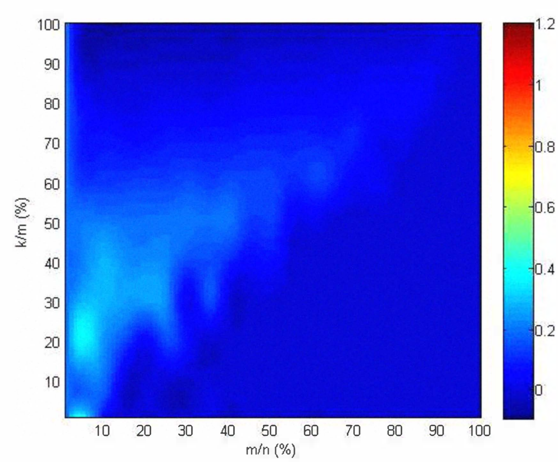

Fig. 23. Phase diagram of the solutions to successive inverted Gaussian minimization (7) in terms of performance index ξ.

Authorized licensed use limited to: N.C. State University Libraries - Acquisitions & Discovery S. Downloaded on April 28,2023 at 17:44:33 UTC from IEEE Xplore. Restrictions apply.

Moreover, from Figures 14-23 we can see that γ and ξ are not well correlated. Sometimes an $\hat{x}$ very near to x under performance index of γ can be obtained, but their levels of sparsity vary significantly from each other. Tables I and II show the exact performance indices γ and ξ of typical points (with different undeterminedness and sparsity levels).

TABLE I

COMPARISON OF PERFORMANCE INDEX γ AT TYPICAL SAMPLES

k/m	m/n	L_1	f_w	f_l	f_h	f_g
10%	10%	0.9432	0.5506	0.4022	0.2431	0.1678
10%	50%	0.7758	0.8271	0.8367	0.8430	0.8172
30%	10%	0.0819	0.0713	0.2358	0.0138	0.0000
30%	20%	0.1073	0.2563	0.4633	0.1193	0.0415
50%	20%	0.0095	0.0568	0.2865	0.0283	0.0000
50%	40%	0.1622	0.2892	0.4095	0.3490	0.0424
50%	60%	0.3745	0.3920	0.4560	0.4378	0.3373
80%	20%	0.1302	0.0148	0.0788	0.0000	0.0000
80%	40%	0.0183	0.0208	0.1554	0.0504	0.0000
80%	60%	0.0338	0.0941	0.1599	0.1263	0.0000

TABLE II

COMPARISON OF PERFORMANCE INDEX ξ AT TYPICAL SAMPLES

k/m	m/n	L_1	f_w	f_l	f_h	f_g
10%	10%	1.653	1.375	0.965	0.263	0.190
10%	50%	1.125	0.920	0.614	0.289	0.204
30%	10%	1.274	1.033	0.979	0.106	−0.013
30%	20%	0.943	0.967	0.817	0.394	0.053
50%	20%	0.708	0.792	0.768	0.221	0.000
50%	40%	0.512	0.574	0.534	0.371	0.068
50%	60%	0.342	0.415	0.393	0.247	0.137
80%	20%	0.464	0.649	0.659	0.000	−0.004
80%	40%	0.212	0.348	0.430	0.192	−0.003
80%	60%	0.210	0.286	0.274	0.203	−0.001

VI. CONCLUSIONS

In this paper, a neurodynamic optimization approach is presented for solving the NP-hard constrained sparsity maximization problem. Four approximations to L_0 norm, weighted L_1 norm, and other three quasi-convex functions, are proposed. A one-layer recurrent neural network is applied to compute the sparse solutions to the approximate norm minimization problems. Experimental results show that the recurrent neural network with appropriate selections of its parameters generates better sparse solutions than L_1 norm minimization. Phase diagrams based on two performance indices show that the successive Gaussian maximization performs the best among the four functions, which yields solutions of the highest sparsity with the highest robustness to parameters. Our further investigations include theoretical analysis to the convergence of the one-layer recurrent neural network with non-convex objective functions (e.g., inverted Gaussian functions).

REFERENCES

[1] A. Gersho, R.M. Gray, *Vector Quantization and Signal Compression*, Boston, MA: Kluwer, 1992.

[2] R.G. Baraniuk, "Compressive sensing," *IEEE Signal Processing Magazine*, vol. 24, no. 4, pp. 118-124, 2007.

[3] G. Box, D. Meyer, "An analysis for unreplicated fractional factorials," *Technometrics*, vol. 28, no. 1, pp 11-18, 1986.

[4] E. Candès, J. Romberg and T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inform. Theory*, vol. 52, no. 2, pp. 489-509, 2006.

[5] D. Donoho, "Compressed sensing," *IEEE Trans. Inform. Theory*, vol. 52, no. 4, pp. 1289-1306, 2006.

[6] M. Lustig, J. M. Santos, J.-H. Lee, D. L. Donoho and J. M. Pauly, "Application of compressed sensing for rapid MR imaging," *SPARS*, (Rennes, France), 2005.

[7] D. Takhar, J. N. Laska, M. B. Wakin, M. F. Duarte, D. Baron, S. Sarvotham, K. F. Kelly, and R. G. Baraniuk, "A new compressive imaging camera architecture using optical-domain compression," *Computational Imaging IV - Proc. of SPIE-IS and T Electronic Imaging*, vol. 6065, pp. 43-52, 2006.

[8] M. F. Duarte, S. Sarvotham, D. Baron, M. B. Wakin, and R. G. Baraniuk, "Distributed compressed sensing of jointly sparse signals," *39th Asilomar Conference on Signals, Systems and Computers*, pp. 1537-1541, 2005.

[9] D. Donoho, "High-dimensional centrosymmetric polytopes with neighborliness proportional to dimension," *Discrete and Computational Geometry*, online first edition, 2005.

[10] D. Donoho and V. Stodden, "Breakdown point of model selection when the number of variables exceeds the number of observations," In *Proc. of the IJCNN*, pp. 1916-1921, 2006.

[11] E. Candès and T. Tao, "Decoding by linear programming," *IEEE Trans. Inform. Theory*, vol. 51, pp. 4203-4215, 2005.

[12] D. Donoho, "For most large underdetermined systems of linear equations the minimal L_1 norm solution is also the sparsest Solution," *Technical Report*, Stanford University, 2004.

[13] D.W. Tank and J.J. Hopfield, "Simple neural optimization networks: an A/D converter, signal decision circuit, and a linear programming circuit," *IEEE Trans. Circuits and Systems*, vol. 33, pp. 533-541, 1986.

[14] S. Zhang and A.G. Constantinides, "Lagrange programming neural networks," *IEEE Trans. Circuits and Systems II*, vol. 39, pp. 441-452, 1992.

[15] M. Forti, P. Nistri and M. Quincampoix, "Generalized neural network for nonsmooth nonlinear progranmming problems," *IEEE Trans. Circuis Syst. I*, vol. 51, pp. 1741-1754, 2004.

[16] Q. Liu and J. Wang, "A one-layer recurrent neural network with a discontinuous hard-limiting activation function for quadratic programming," *IEEE Trans. on Neural Networks*, vol. 19, pp. 558-570, 2008.

[17] Q. Liu and J. Wang, "A one-layer recurrent neural network with a discontinuous activation function for linear programming," *Neural Computation*, vol. 20, pp. 1366-1383, 2008.

[18] Q. Liu and J. Wang, "A recurrent neural network for non-smooth convex programming subject to linear equality and bound constraints," In *I. King et al. (eds.): ICONIP 2006*, Part II, LNCS 4233, pp. 1004-1013, 2006.

[19] Q. Liu and J. Wang, "A one-layer recurrent neural network for non-smooth convex optimization subject to linear equality constraints," In *M. Köppen et al. (eds.): ICONIP 2008*, Part II, LNCS 5507, pp. 1003-1010, 2009.

[20] I. F. Gorodnitsky and B. D. Rao, "Sparse signal reconstruction from limited data using FOCUSS: A re-weighted minimum norm algorithm," *IEEE Trans. Signal Processing*, vol. 45, no. 3, pp. 600-616, 1997.

[21] E. Candès, M. Wakin and S. Boyd, "Enhancing sparsity by reweighted L_1 minimization," *Journal of Fourier Analysis and Applications*, vol. 14, pp. 877-905, 2008.

[22] M. Khajehnejad, W. Xu, S. Avestimehr, B. Hassibi, "Weighted L_1 minimization for sparse recovery with prior information," *IEEE Intl. Symp. Info. Theory*, Seoul, Korea, pp. 483-487, 2009.

[23] I. Daubechies, R. DeVore, M. Fornasier, and S. Gunturk, "Iteratively reweighted least squares minimization for sparse recovery, *Comm. Pure Appl. Math.*, vol. 63, no. 1, pp. 1-38, 2010.

[24] R. Chartrand and W. Yin, "Iteratively reweighted algorithms for compressive sensing, *Proc. of Int. Conf. on Acoustics, Speech, Signal Processing (ICASSP)*, pp. 3869-3872, 2008.

[25] G. Vossen and H. Maurer, "On L_1-minimization in optimal control and applications to robotics," *Optim. Control Appl. Meth.*, vol. 27, pp. 301-321, 2006.

[26] X. Jiang and Y. Ye, "A note on complexity of L_p minimization," *Technical Report*, Stanford University, 2009.

Authorized licensed use limited to: N.C. State University Libraries - Acquisitions & Discovery S. Downloaded on April 28,2023 at 17:44:33 UTC from IEEE Xplore. Restrictions apply.