COUPLED CHAOTIC SIMULATED ANNEALING PROCESSES

J.A.K. Suykens, M.E. Yalçın, J. Vandewalle

K.U. Leuven, ESAT-SCD-SISTA
Kasteelpark Arenberg 10
B-3001 Leuven (Heverlee), Belgium
Email: {johan.suykens, mey}@esat.kuleuven.ac.be

ABSTRACT

In this paper we formulate methods of chaotic simulated annealing within the context of coupled local minimizers. Interpreted within the cellular nonlinear networks context, coupled local minimizers consider local optimization algorithms as cells with local connections between the cells. As a result, information exchange is taking place between the minimizers. Instead of taking local optimization methods as individual cells we explore here the use of chaotic signals as additional driving force as in continuous simulated annealing where deterministic chaos now plays the role of noise. On a number of examples, improved results are obtained by coupled chaotic annealing. In general, the coupling of the minimizers also leads to a variance reduction on the optimal cost function values simulated for many different runs.

Keywords. Cellular nonlinear networks, coupled local minimizers, chaotic annealing, Lagrange programming neural networks, optimization.

1. INTRODUCTION

In [11] a method of coupled local minimizers has been recently proposed where local coupling between continuous time local optimization networks (Lagrange programming networks [5, 16]) has been defined. Interpreted within the cellular nonlinear networks framework [3, 4], cells correspond then to local optimization algorithms and information exchange between the cells takes place through state synchronization constraints that express that the minimizers should reach the same state asymptotically. This method has been successfully applied so far to e.g. the training of multilayer perceptron neural networks, global optimization problems in finite element modelling for civil engineering applications and optimization of Lennard-Jones clusters. The basic formulation of coupled local minimizers is formulated in terms of invariant coupling constants. However, optimal cooperative search is possible where the coupling constants are re-scheduled after solving linear programs in the coupling constants [11]. This implicitly incorporates a mechanism of master-slave dynamics as is known in chaos synchronization methods for secure communications [2, 15]. The performance can be further enhanced also by re-numbering the cells, motivated by principles of small-world networks [14].

The aim of this paper is to study the use of chaos as an additional driving force to the cells and create coupled chaotic annealing processes. This is similar to the use of noise in continuous simulated annealing [7] where the noise (within a stochastic differential equation) is replaced by deterministic chaos. For this purpose we use hypercube attractors [13] consisting of coupled Chua's circuits [9]. Other versions of chaotic annealing have been discussed e.g. in [1]. A number of simulation examples are reported where multistart local optimization (uncoupled), coupled local minimizers, chaotic annealing (uncoupled) and coupled chaotic annealing processes are compared. Coupled chaotic annealing processes perform well in these comparative studies and show a reduced variance on the obtained results for the cost function values.

This paper is organized as follows. In Section 2 we briefly review some basic formulations of coupled local minimizers. In Section 3 we discuss the proposed chaotic annealing method at the individual optimizer level. In Section 4 we formulate the coupled chaotic annealing processes and finally in Section 5 a number of illustrative examples are given.

2. COUPLED LOCAL MINIMIZERS (CLM)

Consider the unconstrained problem of minimization of a twice continuously differentiable cost function

$$\min_{\mathbf{x}\in\mathbb{R}^n} U(\mathbf{x}). \tag{1}$$

A simple continuous time local optimization algorithm is the gradient system (steepest descent method)

$$\dot{\mathbf{x}} = -\mu \nabla_{\mathbf{x}} U(\mathbf{x}) \tag{2}$$

0-7803-7761-3/03/$17.00 ©2003 IEEE

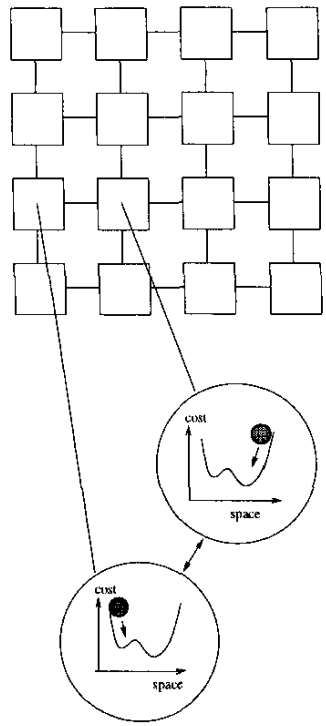

Figure 1: *CNN interpretation of CLMs.*

with $\mu \in \mathbb{R}^+$ the step size.

Instead of running such local optimization methods in a multi-start fashion (i.e. independently from each other for different initial conditions), in coupled local minimizers (CLM) an ensemble of local minimizers has been proposed that interacts through state synchronization constraints. A group consisting of a number of q minimizers is considered for which one aims at minimizing the average energy cost $\langle U \rangle = \frac{1}{q}\sum_{i=1}^{q} U[\mathbf{x}^{(i)}]$ subject to pairwise state synchronization [11]

$$\min_{\mathbf{x}^{(i)} \in \mathbb{R}^n} \langle U \rangle \text{ such that } \mathbf{x}^{(i)} - \mathbf{x}^{(i+1)} = 0 \qquad (3)$$

with individual states $\mathbf{x}^{(i)} \in \mathbb{R}^n$ for $i = 1, 2, ..., q$ and boundary conditions $\mathbf{x}^{(0)} = \mathbf{x}^{(q)}$, $\mathbf{x}^{(q+1)} = \mathbf{x}^{(1)}$. The synchronization constraints have to be achieved in an asymptotic sense, i.e. the individuals have to reach the same final state.

One defines the augmented Lagrangian [6]

$$\mathcal{L}(\mathbf{x}^{(i)}; \lambda^{(i)}) = \frac{\eta}{q}\sum_{i=1}^{q} U[\mathbf{x}^{(i)}] + \frac{1}{2}\sum_{i=1}^{q} \gamma_i \|\mathbf{x}^{(i)} - \mathbf{x}^{(i+1)}\|_2^2 + \sum_{i=1}^{q} \langle \lambda^{(i)}, [\mathbf{x}^{(i)} - \mathbf{x}^{(i+1)}] \rangle \qquad (4)$$

with Lagrange multipliers $\lambda^{(i)} \in \mathbb{R}^n$ $(i = 1, 2, ..., q)$. The penalty factors γ_i emphasize the importance of each of the soft synchronization constraints.

From (4), one obtains the Lagrange programming network [16]

$$\begin{cases} \dot{\mathbf{x}}^{(i)} &= -\nabla_{\mathbf{x}^{(i)}} \mathcal{L}(\mathbf{x}^{(i)}; \lambda^{(i)}) \\ \dot{\lambda}^{(i)} &= \nabla_{\lambda^{(i)}} \mathcal{L}(\mathbf{x}^{(i)}; \lambda^{(i)}) \end{cases}, \; i = 1, 2, ..., q. \qquad (5)$$

This gives the following array of coupled local minimizers

$$\begin{cases} \dot{\mathbf{x}}^{(i)} &= -\frac{\eta}{q}\nabla_{\mathbf{x}^{(i)}} U[\mathbf{x}^{(i)}] + \gamma_{i-1}[\mathbf{x}^{(i-1)} - \mathbf{x}^{(i)}] - \\ & \quad \gamma_i[\mathbf{x}^{(i)} - \mathbf{x}^{(i+1)}] + \lambda^{(i-1)} - \lambda^{(i)} \\ \dot{\lambda}^{(i)} &= \mathbf{x}^{(i)} - \mathbf{x}^{(i+1)} \end{cases}, \; i = 1, 2, ..., q \qquad (6)$$

with learning rate η. This array consists of a number of q coupled nonlinear cells and can be considered as a cellular neural network (CNN) [3, 4] (Fig. 1). Other CLM formulations have been given in [12].

3. CHAOTIC ANNEALING (CA) METHOD

In the chaotic annealing method studied within this paper, a continuous time optimization algorithm of the following form is considered:

$$\dot{\mathbf{x}} = -\mu \nabla_{\mathbf{x}} U(\mathbf{x}) + T\mathbf{d}(t)\, e^{-\tau t} \qquad (7)$$

with $T \in \mathbb{R}$, $\mathbf{d}(t) \in \mathbb{R}^n$ is an external driven force vector (taken as output from a hyperchaotic system [13], in this case a hypercube attractor consisting of coupled Chua's circuits [9]) (Fig.2). The value τ is a cooling speed time constant. Additional componentwise scaling can be taken into account if needed. The hypercube attractor is obtained from

$$\begin{cases} \dot{x}^{(j)} &= \alpha[y^{(j)} - h(x^{(j)})] \\ \dot{y}^{(j)} &= x^{(j)} - y^{(j)} + z^{(j)} + K(y^{(j)} - y^{(j-1)}) \\ \dot{z}^{(j)} &= -\beta y^{(j)}, \; j = 1, 2, ..., L \end{cases} \qquad (8)$$

with L the number of cells in the one-dimensional CNN array. The nonlinearity can be either related to Chua's circuit or generalized n-scroll circuits. In this paper we take the Chua circuit with $h(x) = m_1 x + \frac{1}{2}(m_0 - m_1)((x+1) - (x-1))$ where $m_0 = -1/7, m_1 = 2/7$ and $\alpha = 9, \beta = 14.286$, $K = 0.06$. Here $\mathbf{d}(t) = [x^{(1)}; x^{(2)}; ...; x^{(L)}]$ is taken as driving force for the optimization algorithm. The equations (7)(8) are simulated simultaneously.

4. COUPLED LOCAL MINIMIZERS WITH CHAOTIC ANNEALING (CLM-CA)

Instead of taking steepest descent gradient systems (2) within the CLM, we couple now different chaotic an-

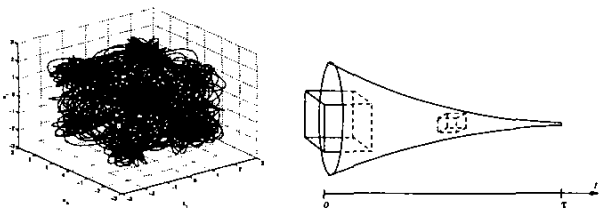

Figure 2: *Hypercube attractor for chaotic annealing.*

nealing systems (7) within a CLM:

$$\left\{\begin{array}{rcl} \dot{\mathbf{x}}^{(i)} & = & -\frac{\eta}{q}\nabla_{\mathbf{x}^{(i)}}U[\mathbf{x}^{(i)}] + T\mathbf{d}^{(i)}(t)\,e^{-\tau t} + \\ & & \gamma_{i-1}[\mathbf{x}^{(i-1)} - \mathbf{x}^{(i)}] - \\ & & \gamma_i[\mathbf{x}^{(i)} - \mathbf{x}^{(i+1)}] + \lambda^{(i-1)} - \lambda^{(i)} \\ \dot{\lambda}^{(i)} & = & \mathbf{x}^{(i)} - \mathbf{x}^{(i+1)}\,,\; i = 1,2,...,q \end{array}\right. \tag{9}$$

where $L = q$ and the hypercube attractors are simulated simultaneously with the CLM dynamics.

5. ILLUSTRATIVE EXAMPLES

We illustrate and compare the methods on a number of illustrative examples. A first test function is

$$U(\mathbf{x}) = \frac{a}{2n}\sum_{i=1}^{n} x_i^2 + 4n - 4n\prod_{i=1}^{n}\cos(\omega x_i) \tag{10}$$

with $a = 0.01$, $\omega = 0.2$ which has been optimized here for $n = 3$ and $n = 10$ with random initial conditions in $[-10000, 10000]^n$. This test function has been reported and investigated in [10] for simulated annealing and stochastic approximation methods. For the annealing scheme was taken $T = 50$, $\tau = 0.01$, $\mu = 65$ (case $n = 3$), $T = 150$, $\tau = 0.01$, $\mu = 20$ (case $n = 10$), and $\mu = 65$, $\gamma_i = 0.01\forall i$. A number of $q = 3$ minimizers has been taken and 100 different runs with random initial conditions were taken. The equations were simulated in Matlab using ode23 (in order to avoid storing state variables for all time instants the Matlab option [t,x] = ode23('filename', tspan, x0) is used where tspan specifies the time instants to be returned). Boxplots of the simulation results with 100 runs are shown in Fig.3-4. Each time, uncoupled gradient networks (GN), CLM, CA and CLM-CA are compared for the same set of initial conditions. One observes a significant improvement on these test functions of CLM, CA, CLM-CA wrt GN. The use of chaotic annealing leads to improvements in performance and CLM-CA leads to best variance reduction on the results.

In Fig.5 results of the method are shown on a 10-city Travelling Salesman Problem (TSP) problem. Let N be the number of cities and $d_{x,y}$ be distance between cities x and y. A tour of the TSP can then be represented by a $N \times N$ permutation matrix [8], where each row and column is associated to a particular city and order within the tour, respectively. As cost function is then taken here

$$\begin{array}{cccc} U & \mathbf{v} & \frac{A}{2}\sum_x\sum_i\sum_{j\neq j} v_{x,i}v_{x,j} & \frac{B}{2}\sum_i\sum_x\sum_{y\neq x} v_{x,i}v_{x,j} \\ \frac{C}{2}\left(\sum_x\sum_i v_{x,i} - N\right)^2 & \frac{D}{2}\sum_x\sum_{y\neq x}\sum_j d_{x,y}v_{x,i}\; v_{y,i+1} & & v_{y,i-1} \end{array} \tag{11}$$

where $v_{x,i}$ represents visiting city x as i-th city and the indices $i+1$ and $i-1$ subscripts are modulo N. This has been analysed and proposed in the area of Hopfield network [8]. We consider here implementations which are different from the Hopfield networks (dynamics in state variables $u_{x,i}$ with $v_{x,i} = \frac{1}{2}(1 + \tanh(u_{x,i}))$) for comparison reasons (comparisons made between GN, CLM, CA, CLM-CA). The simulation results for 100 different runs and $D = 1, C = 30, A = 20, B = 20$, $T = 0.4$, $\tau = 0.5$, $\gamma = 40$, $q = 5$, $\mu = 5$ with random initial conditions indicate a good performance with low variance of the CLM-CA technique (Fig.5).

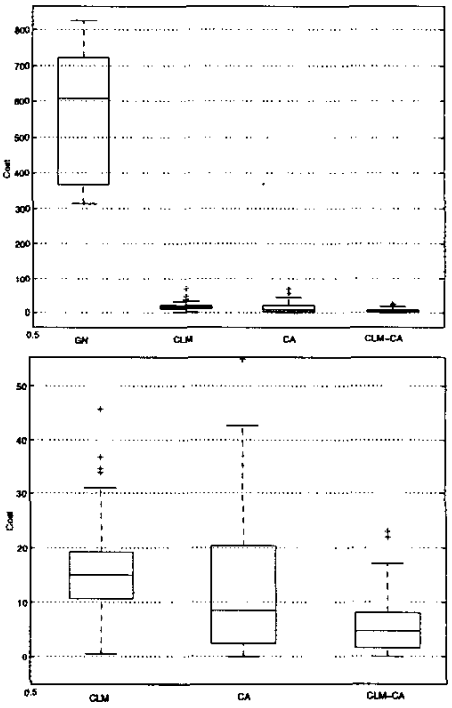

Figure 3: *Test function for $n = 3$: (Top) comparison between GN, CLM, CA, CLM-CA (from left to right); (Bottom) enlarged top-figure for CLM, CA, CLM-CA.*

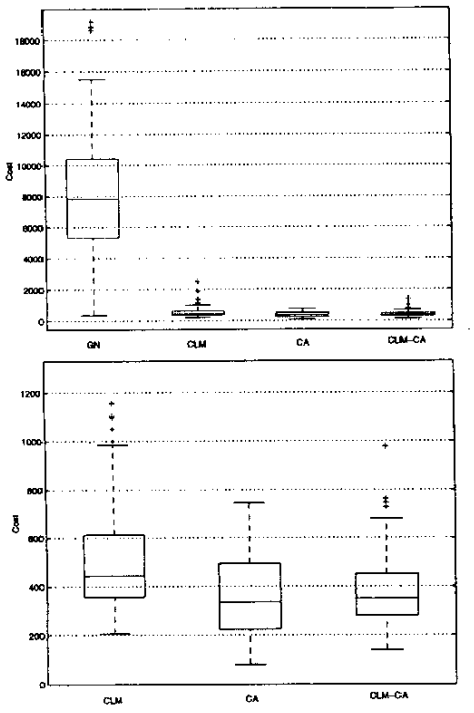

Figure 4: *Test function for $n = 10$: (Top) comparison between GN, CLM, CA, CLM-CA (from left to right); (Bottom) enlarged top-figure for CLM, CA, CLM-CA.*

6. CONCLUSIONS

Methods of chaotic annealing have been explored in this paper for optimization within the context of cellular nonlinear networks and coupled local minimizers. Simulation results on a number of test functions indicate a good performance of CLM-CA with low variance on many different runs.

Acknowledgments.

2

2

Collective Behavior and:Optimization: an Interdisciplinary Approach

2

7. REFERENCES

11

Proceedings of the IEEE **90** 1 2 2

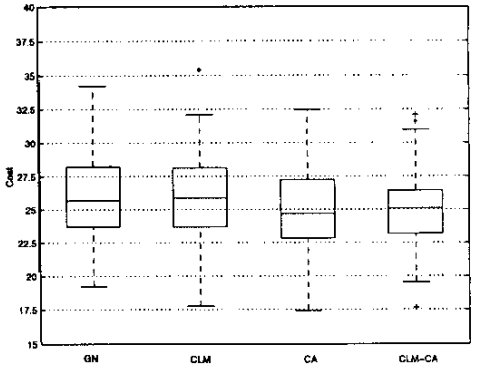

Figure 5: *10-city TSP problem: comparison between GN, CLM, CA, CLM-CA.*

2 *From Chaos to Order - Perspectives, Methodologies, and Applications* 1

IEEE Trans. Circuits and Systems-I **40** 1 1 1

CNN: a Paradigm for Complexity 1

Neural Networks for Optimization and Signal Processing **1**

Practical Methods of Optimization 1

$\mathbb{R}^d$ *SIAM Journal on Control and Optimization* **31** 1 111 1 1 **1**

Biological Cybernetics **52** 1 2 1

Int. J. Bifurcation and Chaos **4** 2 2 1 1

Neural Networks **3** 1 11

Int. J. Bifurcation and Chaos **11** 21 21 2 1 12

Proc. of the World Congress on Computational Intelligence - International Joint Conference on Neural Networks 2 2 2 2 2 2 1

1 *Int. J. Bifurcation and Chaos* **7** 1 1 1 1

Nature **393** 2 1 1

Int. J. Bifurcation and Chaos **4** 1 1

IEEE Trans. Circuits and Systems-II **39** 1 2 1 2