

## ESTUDO DO ESTADO FUNDAMENTAL DE AGLOMERADOS DE SILÍCIO VIA REDES NEURAIS

Maurício Ruv Lemes\* e Arnaldo Dal Pino Júnior

Departamento de Física, Instituto Tecnológico de Aeronáutica, Pça Marechal do Ar Eduardo Gomes, 50, 12228-900 São José dos Campos - SP

Recebido em 10/4/01; aceito em 22/11/01

STUDY OF THE GROUND-STATE GEOMETRY OF SILICON CLUSTERS THROUGH ARTIFICIAL NEURAL NETWORKS. We introduce a global optimization method based on the cooperation between an Artificial Neural Net (ANN) and Genetic Algorithm (GA). We have used ANN to select the initial population for the GA. We have tested the new method to predict the ground-state geometry of silicon clusters. We have described the clusters as a piling of plane structures. We have trained three ANN architectures and compared their results with those of pure GA. ANN strongly reduces the total computational time. For  $Si_{10}$ , it gained a factor of 5 in search speed. This method can be easily extended to other optimization problems.

Keywords: silicon clusters; genetic algorithm; neural network.

### INTRODUÇÃO

Redes Neurais Artificiais (RNA) vêm sendo muito utilizadas em Química. Recentemente Gasteiger e Zupan<sup>1</sup> publicaram uma monografia sobre aplicações de RNA em Química. As RNA mostram uma grande capacidade de reconhecimento de padrões; alguns exemplos incluem a identificação automática de grupos de espectro molecular e a determinação da estrutura de uma proteína e da sequência de um amino-ácido<sup>1</sup>. Uma outra aplicação importante foi feita por Sigman e Rives<sup>2</sup>, que utilizaram uma RNA para prever potenciais de ionização atômica usando camadas de valência como dados de entrada para a rede. Estas aplicações encorajam-nos a explorar o potencial das RNA numa área diferente da Química: a geometria de sistemas poliatômicos, em particular a otimização da geometria de aglomerados de silício. As condições experimentais de geração e seleção<sup>3,4</sup> destes aglomerados inviabilizam a determinação experimental de sua geometria. Como consequência, devemos inferir a estrutura dos aglomerados, através de evidências indiretas ou através de cálculos teóricos. Mesmo teoricamente, encontrar a geometria referente ao estado fundamental de uma grande coleção de átomos é uma tarefa extremamente complicada devido a duas razões: i) a energia do aglomerado não é função apenas das coordenadas de cada núcleo, mas seus elétrons também devem ser considerados. Assim, para tratar estes problemas, a energia deve ser calculada através dos princípios da mecânica quântica sendo cada cálculo de energia computacionalmente bastante delicado<sup>5</sup>, e ii) a superfície de energia depende de um número grande de variáveis e apresenta incontáveis mínimos locais. Por exemplo, um aglomerado de gases nobres<sup>6</sup> de cerca de 100 átomos tem um número estimado de  $10^{60}$  mínimos! Obviamente determinar o mínimo global entre tantos mínimos locais é uma tarefa muito difícil.

Redes neurais artificiais (RNA) demonstram um grande potencial de aplicação em modelagem, simulação, controle e predição<sup>7-9</sup> em diversas áreas de ciências e química. Na maioria destas aplicações, as RNA's são treinadas com dados coletados durante operações ou experimentos. Concluído o treinamento, a capacidade natu-

ral de generalização das redes permite que as RNA's realizem as previsões desejadas.

Tradicionalmente, vários problemas práticos de Química e Física são transformados em problemas de otimização. Existem muitos algoritmos para a solução de problemas de otimização e podemos dividi-los em dois grupos: i) Aqueles baseados no cálculo de gradientes<sup>10,11</sup> e ii) os que independem do gradiente<sup>12</sup>. As diversas variantes do método dos gradientes conjugados<sup>11</sup> são exemplos de procedimentos que usam a derivada para sistematicamente aproximar-se das condições ótimas desejadas. Estes métodos não possuem a capacidade de evitar as armadilhas de mínimos locais e por isso costumam ser repetidos diversas vezes a partir de diferentes pontos de partida. O melhor resultado de uma série de procedimentos iterativos deste tipo é assumido como sendo a solução procurada. Por outro lado, algoritmos genéticos<sup>13-15</sup> e recozimento simulado<sup>12,16</sup> são procedimentos que não se baseiam no cálculo de gradientes. Eles imitam a natureza e têm a característica de poderem evitar sua captura em mínimos locais. Qualquer que seja o método, a escolha do ponto (ou pontos) de partida para o cálculo é de vital importância para que se encontre a solução desejada e para evitar os mínimos locais.

Recentemente, os pesquisadores Cundari e Moody<sup>17</sup> testaram a utilização de redes neurais para predição de propriedades moleculares de uma série de moléculas diatômicas. Eles usaram informações de número atômico, massa atômica, configuração eletrônica de cada átomo e verificaram que as RNA's tinham uma capacidade de previsão de grandezas como frequência de vibração, energia de ligação e distância de equilíbrio comparáveis ou superiores aos cálculos de primeiros princípios.

Este trabalho, propõe-se a avaliar a associação de uma rede neural a cálculos de química quântica para determinação da geometria do estado fundamental de aglomerados de silício. Aqui, aproveitamos as redes neurais para selecionar pontos de partida para métodos iterativos de otimização. Especificamente, uma RNA classificará estruturas candidatas para o algoritmo genético. Enquanto Cundari e Moody utilizaram RNA's para compará-las com cálculos de primeiros princípios, nós utilizamos as redes para acelerar o cálculo de química quântica e não para substituí-lo. Queremos ainda aplicar o método a um importante problema de Química de materiais. A miniaturização de dispositivos estimula o interesse nas propriedades

\*e-mail: ruv@uol.com.br

de aglomerados<sup>18</sup>, e estruturas envolvendo o silício são particularmente interessantes, uma vez que o silício ainda é o mais importante elemento para o desenvolvimento de dispositivos eletrônicos. Logo, a busca por modelos estruturais de aglomerados de silício é motivada pelo desenvolvimento tecnológico. Afinal a estrutura determina, em boa parte, as características elétricas e mecânicas do material<sup>19</sup>.

Vários trabalhos tentam descrever, teoricamente, as conformações tridimensionais e as propriedades dos aglomerados de Silício. Métodos de primeiros princípios<sup>20-22</sup> estão limitados a poucos átomos. Somente os menores (até 10 átomos) podem ser estudados pela associação destes métodos com algoritmos de busca<sup>20</sup> e explorar todo o espaço conformacional. No caso de aglomerados maiores não podemos realizar tal estudo. Assim sendo, para aglomerados com mais de 10 átomos as buscas são restritas a modelos estruturais fisicamente motivados<sup>23 e 24</sup>. Tentativas anteriores restringem-se a determinadas simetrias ou foram baseadas na geometria do cristal<sup>25</sup>, ou ainda, nas superfícies de reconstrução de Silício<sup>26</sup>.

Neste trabalho utilizamos as RNA's para distinguir a afinidade existente entre diferentes camadas atômicas. A partir de estruturas de aglomerados pequenos, cujas energias foram previamente calculadas, desejamos que a RNA identifique quais camadas tendem a se ligar mais fortemente entre si. Composto seqüências de camadas que a RNA prevê como energeticamente favoráveis, podemos evitar que o algoritmo de busca desperdice o tempo de cálculo com estruturas quimicamente desfavoráveis. Neste caso, a RNA substitui a experiência do investigador por sua capacidade de aprendizado e de generalização.

Considerando a alta dimensão do espaço de configurações seria impossível esperar que com um número pequeno de dados de treinamento<sup>27</sup> a RNA fosse capaz de determinar a estrutura correspondente ao mínimo global. Contudo, constatamos que ela é capaz de selecionar eficientemente estruturas para a subsequente otimização global pelo algoritmo de busca, em nosso caso: o algoritmo genético. Nossos resultados mostram que a RNA aumenta de forma significativa a eficiência do algoritmo de otimização.

A seguir apresentaremos o procedimento pelo qual transformamos o problema químico em um problema de classificação. Discutiremos, também a arquitetura da rede e os resultados obtidos pela combinação do classificador – RNA com o algoritmo genético.

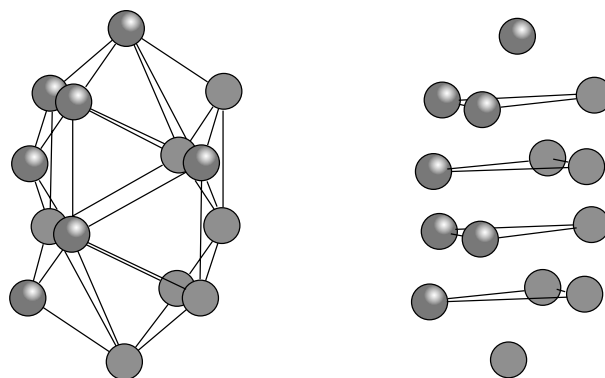
## ADEQUAÇÃO DAS GEOMETRIAS ÀS REDES NEURAIS

Para que as RNA's possam atuar com eficiência, interpretamos a estrutura geométrica de um aglomerado como sendo um empilhamento de camadas planas de átomos. Este tratamento assemelha-se ao apresentado por Grossman e Mitas<sup>23</sup>. Eles sugeriram uma descrição geométrica dos aglomerados de silício como uma sobreposição de elementos triangulares, com alguns átomos nas extremidades, conforme Figura 1.

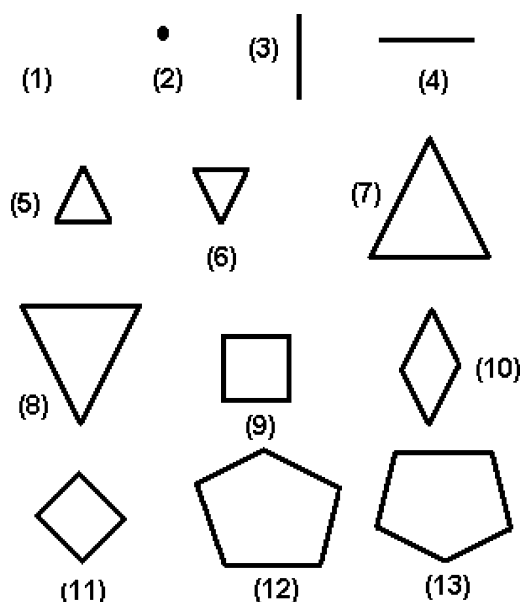
Na Figura 1, estruturas tridimensionais são descritas como uma série de camadas sobrepostas, cada qual contendo três átomos. Neste trabalho seguimos a idéia de que aglomerados podem ser descritos por uma sobreposição de camadas poliatômicas, eliminando, porém, com a limitação de três átomos por camada. Descrevemos cada aglomerado como um empilhamento de camadas poliatômicas com até cinco átomos. Esta escolha tornou o algoritmo bastante eficiente, pois a descrição através do empilhamento de camadas planas restringe o número de maneiras pelas quais um aglomerado com  $n$  átomos pode ser construído.

Criamos um conjunto de estruturas possíveis para a formação das camadas (Figura 2).

Na Figura 3 mostramos um exemplo de aglomerado de  $Si_6$  representado por 5 distintas descrições, a partir dos elementos da figu-



**Figura 1.** Aglomerado ( $Si_{14}$ ) como uma sobreposição de camadas poliatômicas. À esquerda, temos a visão global do aglomerado; à direita, destacamos a superposição de camadas com três átomos



**Figura 2.** O elemento (1) descreve uma camada com zero átomos, representando o espaço livre (vácuo) que envolve os aglomerados estudados. Os demais elementos são as camadas que empilhadas formam os aglomerados. Note que cada vértice corresponde à posição de um átomo

ra anterior. É importante ressaltar que os elementos geométricos utilizados em cada descrição serão utilizados como dados de entrada da RNA.

O classificador neural foi construído com o objetivo de filtrar as configurações que seriam fornecidas como possíveis candidatas ao algoritmo genético. A RNA deve distinguir quais empilhamentos de camadas atômicas têm alta energia de ligação, isto é, correspondem a estruturas mais estáveis. Deve ainda ser possível treiná-la com um número bastante restrito de elementos no conjunto de treinamento. Dividimos sua preparação em três partes: (a) geração de dados de entrada; (b) treinamento e; (c) predição.

A primeira etapa consiste em obter as informações e elementos necessários à descrição e caracterização do tema. Fornecemos como dados de entrada para a rede uma tabela previamente calculada com as energias de ligação de 110 aglomerados. O conjunto de treinamento é composto de estruturas de até 9 átomos correspondentes a mínimos locais, globais e outras estruturas cujas energias nada possuem em especial (veja três exemplos na Tabela 1). É importante

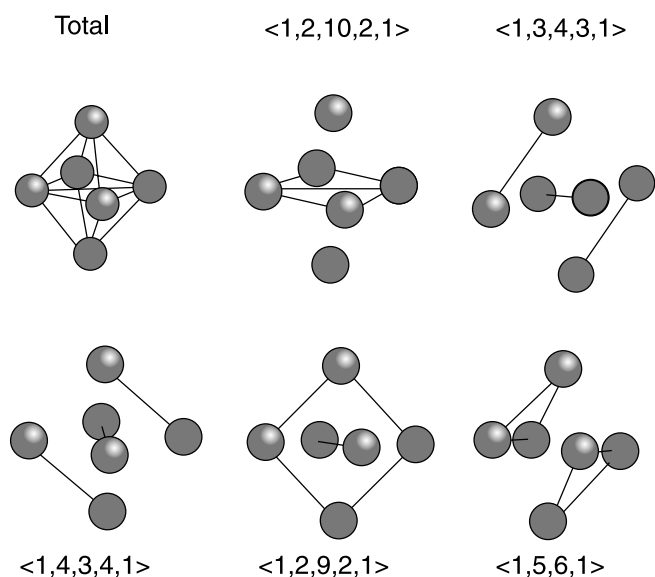


Figura 3. Configurações do  $Si_6$ , com suas respectivas estruturas formadoras

ressaltar que a aplicação do classificador neural exige sua combinação com um método de cálculo de energia total, mas qualquer método que escolhêssemos seria igualmente conveniente. Neste trabalho utilizamos o método semi-empírico Tight Binding (TB) cuja descrição detalhada pode ser encontrada nas referências<sup>25-26,28-30</sup>.

Tabela 1. Estruturas de formação e energia

Estrutura	Aglomerado Correspondente	Energia de ligação por átomo (eV)
1 4 3 4 1	$Si_6$	3.0710
1 5 10 1	$Si_7$	3.2670
1 1 1 1	—	0.0000

Na segunda etapa realizamos o treinamento, adotando como dados de entrada as 110 estruturas e como dados de saída suas respectivas energias. Nessa etapa, que é de extrema importância porque determinará a qualidade das predições a serem realizadas, empregamos o método de treinamento conhecido como back propagation<sup>31</sup>. Treinamos a RNA para classificar as estruturas como adequadas ou inadequadas para a minimização global. Esse critério é baseado no conhecimento prévio de estruturas menores usadas como dados de treinamento da rede. Assim, as estruturas cujas energias previstas estão abaixo de um valor limite são consideradas inadequadas, enquanto aquelas acima deste valor serão utilizadas pelo algoritmo de busca do mínimo global. Podemos observar na Tabela 1 que a energia de ligação por átomo para o  $Si_6$  é maior que 3 eV. Como é esperado que a energia de ligação por átomo aumente com o número de átomos, escolhemos 2,8 eV como fator limitante. O valor foi escolhido levando em consideração que a RNA foi treinada com poucos dados de entrada e portanto é esperado que seu resultado seja qualitativamente bom, mas quantitativamente impreciso.

O treinamento pode ser readaptado de maneira rápida e simples a outros aglomerados. É fundamental dizer aqui também que o treinamento não precisa ser extremamente longo para ser confiável, mesmo um treinamento breve melhora o desempenho do Algoritmo Genético, conforme mostraremos na seção de resultados.

Na terceira e última etapa fazemos a predição. Geramos todas as combinações de camadas possíveis do aglomerado  $Si_n$  ( $n > 9$ ). A seguir o preditor seleciona as estruturas classificadas como adequadas e elimina as demais. As estruturas adequadas são utilizadas de duas formas: (i) um certo número delas é utilizado como população inicial para o algoritmo genético; (ii) os aglomerados restantes são introduzidos na população do algoritmo através de mutações, ou seja, a cada  $n$  gerações duas novas estruturas são introduzidas na população. É importante dizer que o cálculo da energia do aglomerado dentro do algoritmo genético é feito através da aproximação TB. Com o objetivo de testar o método aqui apresentado, escolhemos determinar a geometria do estado fundamental do aglomerado composto por 10 átomos de silício. A escolha do tamanho do aglomerado deve-se ao fato que: (i) este sistema possui muitos mínimos locais; (ii) dada uma configuração geométrica, a energia pode ser calculada rapidamente na aproximação TB.

## APLICAÇÃO E RESULTADOS

Definimos a arquitetura de nossas 3 RNA's da seguinte maneira: uma camada de entrada com 11 elementos, uma camada intermediária e uma camada de saída com 2 elementos. Na camada intermediária, empregamos 12 (RNA12), 6 (RNA6) e 3 (RNA3) neurônios, obtendo resultados que apresentaremos a seguir.

Independente da rede, um treinamento rápido é capaz de identificar uma porcentagem elevada de estruturas inadequadas. Decidimos encerrar os treinamentos das redes quando 60 % das estruturas do conjunto inicial eram reconhecidas como inadequadas.

Para cada uma das redes RNA12, RNA6 e RNA3 realizamos o seguinte procedimento: a rede fornece um conjunto de 10 estruturas escolhidas aleatoriamente que constituiu a população inicial de um cálculo com o algoritmo genético. Além disso, o algoritmo genético necessita de um critério para a inclusão de mutações nos indivíduos da população. Fizemos com que a cada 10 gerações, duas novas estruturas, escolhidas ao acaso, dentre as consideradas adequadas, substituíssem as estruturas de menor energia de ligação por átomo da população naquele instante.

O algoritmo genético (GA), pré-condicionado por cada uma das 3 RNA's, foi rodado um total de 3000 gerações e seus resultados comparados aos obtidos pelo GA Puro. Como a aplicação do algoritmo genético depende do uso de números aleatórios realizamos 10 cálculos distintos para cada uma das nossas RNA's. Desse modo acreditamos que o resultado da média dos 10 cálculos reflete de maneira mais confiável as características deste novo procedimento. Nas Figuras 4, 5 e 6, comparamos o melhor desses cálculos e a média dos mesmos, com a média de dez cálculos de algoritmo genético puro ou seja, algoritmo genético sem RNA's. Note que estes gráficos apresentam a evolução do oposto da energia de ligação por átomo como função do número de gerações. Desta forma as estruturas mais estáveis, que correspondem a maior energia de ligação por átomo, são representadas pelos menores valores de energia da figura.

Na Figura 4 analisamos a RNA12. Enquanto que o cálculo genético puro leva em torno de 4000 gerações para encontrar estruturas com energia de ligação por átomo maior que 3 eV, nosso melhor cálculo com esse tipo de rede obteve essa mesma marca com apenas 500 gerações! Com o cálculo médio atingimos este valor com cerca de 1650 gerações.

Na Figura 5 analisamos a RNA6. Podemos notar que o nosso melhor cálculo leva cerca de 200 gerações para atingir a energia de ligação por átomo maior que 3 eV. Com o cálculo médio obtivemos a mesma marca com 300 gerações.

Na Figura 6 analisamos a RNA3. Podemos notar que o nosso melhor cálculo leva cerca de 800 gerações para encontrar estruturas

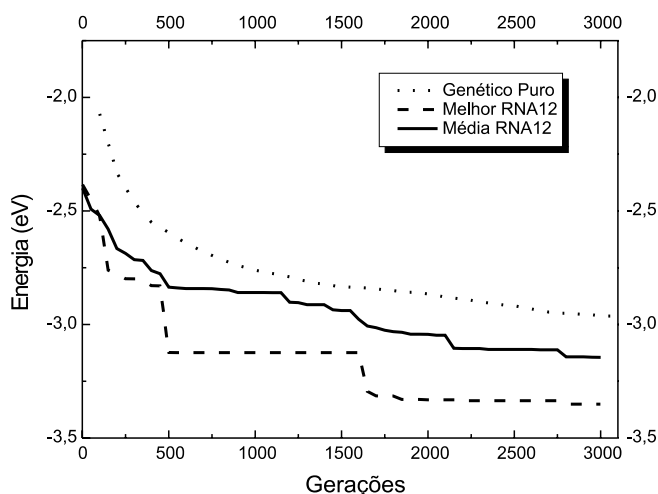


Figura 4. A evolução comparativa para camada escondida com 12 neurônios

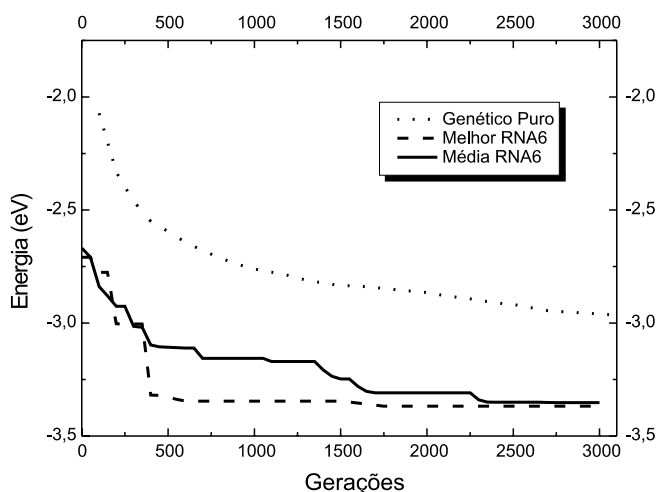


Figura 5. A evolução comparativa para camada escondida com 6 neurônios

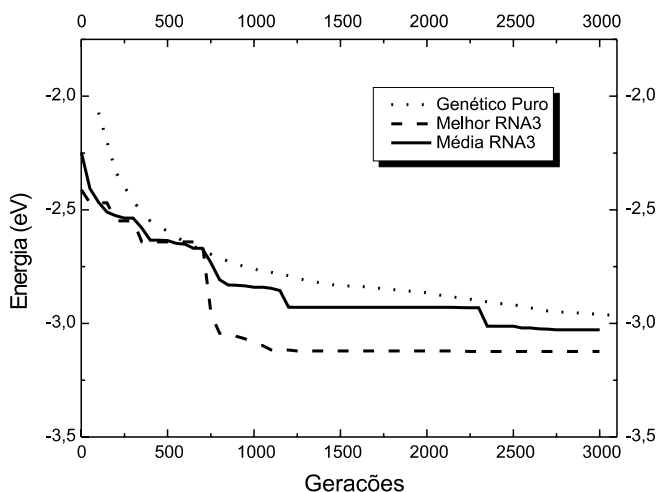


Figura 6. A evolução comparativa para camada escondida com 3 neurônios

com energia de ligação por átomo maior que 3 eV. Com o cálculo médio atingimos este valor com cerca de 2000 gerações.

Os diferentes gráficos mostram que a introdução de RNA's nos cálculos efetuados com algoritmos genéticos faz com que o número de gerações a serem realizadas seja diminuído grandemente. Outra constatação interessante é que quando reduzimos drasticamente o número de neurônios na rede, passamos a ter dificuldades em obter resultados satisfatórios (RNA3). É nítido também que não é necessário possuímos um grande número de neurônios na camada intermediária para obtermos bons resultados (basta comparar os resultados da RNA12 com aqueles da RNA6).

## CONCLUSÕES

Os nossos resultados mostraram que uma rede neural pode ser treinada para incorporar informações de mecânica quântica e agilizar os cálculos de energia total de sistemas poliatômicos. A RNA mostrou-se capaz de distinguir os empilhamentos de camadas atômicas que geram estruturas mais estáveis. Após a aplicação de um treinamento rápido as RNA's conseguem selecionar pontos de partida eficientes para métodos de otimização global.

No caso do  $\text{Si}_{10}$ , o tempo utilizado para o treinamento de uma RNA foi menor do que o despendido em 100 gerações do algoritmo Genético Puro. Para todas as RNA's empregadas, o número de gerações necessários para atingir uma energia de ligação por átomo de 3 eV é pelo menos 2000 gerações<sup>32</sup> mais rápido que o Genético Puro. Deste modo, o tempo gasto para treinar a rede é quase que desprezível.

Os resultados apresentados na seção anterior mostraram que a utilização de redes neurais para seleção de candidatos para problemas de otimização de geometria, agilizam sensivelmente os cálculos. Quando o número de átomos do aglomerado,  $N$ , aumenta, o número de mínimos locais cresce exponencialmente com  $N$  e o tempo de cada iteração, por envolver uma diagonalização matricial, cresce com  $N^3$ . Por outro lado o tempo de treinamento da rede mantém-se praticamente inalterado. Por isso consideramos o método muito promissor para ser adaptado para aglomerados maiores ( $\text{Si}_n$ ,  $n > 10$ ). O nosso algoritmo pode ainda ser adaptado para outros materiais, para outros métodos de cálculo de energia total e ainda para outros problemas de otimização.

## REFERÊNCIAS

- Gasteiger, J.; Zupan, J.; *Neural Networks for Chemists*, VCH: Weinheim, 1993.
- Sigman, M. E.; Rives, S. S.; *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 617.
- Bloomfield, L. A.; Freeman, R. R.; Brown, W. L.; *Phys. Rev. Lett.* **1985**, *54*, 2246.
- Jarrod, M. F.; Constant, V. A.; *Phys. Rev. Lett.* **1991**, *67*, 2994.
- Remler, D. K.; Madden, P. A.; *Mol. Phys.* **1990**, *70*, 921.
- Wales, D. J.; Doye, J. P. K.; *J. Phys. Chem. A* **1997**, *101*, 5111.
- He, F.; Sung, A. H.; *IASTED International Conference on Control*, New Mexico, Mexico, 1997.
- Narendra, K. S.; Parthasarathy, K.; *IEEE Trans. on Neural Networks* **1990**, *1*, 4.
- Drossu, R.; Obradovic, Z.; *IEEE Computational Sciences and Engineering* **1996**, *3*, 78.
- Polak, E.; *Computational Methods in Optimization: A Unified Approach*, New York and London: Academic Press, 1971.
- Press, W. H.; Teukolsky, S. A.; Vetterling, W. T.; Flannery B. P.; *Numerical Recipes: The Art of Scientific Computing*, 2<sup>nd</sup> ed., Cambridge University Press: Cambridge, 1992.
- Lemes, M. R.; Zacharias, C. R.; Dal Pino, Jr. A.; *Phys. Rev. B* **1997**, *56*, 9279.
- Holland, J. H.; *Adaptation in Natural and Artificial Systems*, Ann Arbor: University of Michigan Press, 1989.
- Holland, J. H.; *Sci. Am.* **1992**, *July*, 44.
- Zacharias, C. R.; Lemes, M. R.; Dal Pino, Jr. A.; *J. Mol. Struct. (Theochem)* **1998**, *430*, 29.

16. Kirkpatrick, S.; Gellar, D. C.; Vecchi, M. P.; *Science* **1983**, 220, 671.
17. Cundari, T. R.; Moody, E. W.; *J. Chem. Inf. Comput. Sci.* **1997**, 37, 871.
18. Ho, K. M.; *Nature* **1998**, 392, 582.
19. Somorjai, G. A.; *Chemistry in Two Dimensions: Surfaces*, Ithaca: Cornell Univ. Press, 1981.
20. Raghavachari, K.; Rohlfing, C. M.; *J. Chem. Phys.* **1988**, 89, 2219.
21. Car, R.; Parrinello, M.; *Phys. Rev. Lett.* **1985**, 55, 2471.
22. Grossman, J. C.; Mitas, L.; *Phys. Rev. Lett.* **1995**, 74, 1323.
23. Grossman, J. C.; Mitas, L.; *Phys. Rev. B* **1995**, 52, 16735.
24. Kaxiras, E.; Jakson, K. A.; *Z. Phys. D.* **1993**, 26, 346.
25. Chadi, D. J.; Cohen, M. L.; *Phys. Stat. Sol. B* **1975**, 68, 405.
26. Chadi, D. J.; *Phys. Rev. B* **1984**, 29, 785.
27. Lanoutte, R.; Thibault, J.; Valade, J. L.; *Comput. Chem. Eng.* **1999**, 23, 1167.
28. Laasonen, K.; Nieminen, R. M.; *J. Phys. B:Condens. Matter* **1990**, 2, 1509.
29. Menon, M.; Subbaswamy, K. R.; *Phys. Rev. B* **1993**, 47, 12754.
30. Wang, C. Z.; Chan, C.T.; Ho K. M.; *Phys. Rev. Lett.* **1991**, 66, 189.
31. Hertz, J.; Krog, A.; Palmer, R. G.; *Introduction to the Theory of Neural Computation*, Addison-Wesley Publishing Company: Redwood City CA, 1991.
32. Para efeito de comparação, utilizamos o tempo gasto pelos cálculos de uma geração do algoritmo genético como unidade de tempo. Isso permite que as comparações sejam percebidas independentemente do computador utilizado.