# Predicting the equilibrium structure of organic semiconductors with genetic algorithms

Luiz Antonio Ribeiro Jr., Pedro Henrique de Oliveira Neto, Wiliam Ferreira da Cunha, Ricardo Gargano *, Geraldo Magela e Silva

*Institute of Physics, University of Brasilia, Brasilia 70.919-970, Brazil*

### ABSTRACT

We propose a genetic algorithm optimization technique (GAOT) designed to obtain the lattice geometry parameters and the electronic structure of $\pi$-conjugated polymer systems. Particularly, we obtain the ground state structure of a completely dimerized trans-polyacetylene chain $(CH)_x$ using this approach. We also use this approach to determine the polaron and soliton solutions. These solutions are then compared to the traditional self-consistent-field (SCF) solutions. We conclude that the optimization technique proposed here generates solutions as those obtained using traditional (SCF) techniques, with the important advantage of reducing significantly the number of iterations needed to reach convergence. Besides, the convergence of the GAOT technique is more robust, reaching solutions independently of initial guesses.

© 2012 Elsevier B.V. All rights reserved.

## 1. Introduction

The potential of organic semiconductors in the development of new optoelectronic devices has attracted the interest of the scientific community in recent years [1,2]. Applications of these materials include light emitting diodes (OLEDs) [3,4], thin film transistors [5,6], photodiodes and photovoltaics (OPVs) [7,8]. A significant challenge to make efficient use of these new materials is the accurate prediction of their geometrical and electronic properties. It is known that, in $\pi$-conjugated polymers, the low-energy structure displays a pattern of alternate single and double bonds forming a dimerized polymeric chain with two phases [9]. The determination of this ground state structure is crucial to the analysis of different optical and electronic properties in these materials.

Genetic algorithm optimization techniques (GAOTs) are a class of computational methods based on evolution to quickly determine a potential solution for a specific problem. These algorithms allow the development of generalized methods of search and optimization that attempts to solve by trial and error, complex problems. The underline idea is to simulate the natural processes of evolution without human intervention. Genetic algorithms have been applied successfully in the description of a variety of global minimization and optimization problems [10–17], geometrical structures predictions [18], and fitting of the potential energy surfaces [19–23]. Theoretical studies about ground state structures and low-energy configuration predictions have been extensively performed recently [24–28].

Another interesting feature of genetic algorithms is that they allow mapping of the phase space. This feature allows for the location not only of global minimum but also of all local minima that describes the system. In this sense, Xiang et al. analyzed the structure and the electronic properties of metallic and organic systems using a genetic algorithm to search the global lowest-energy structures [29,30]. In the case of organic systems, a genetic algorithm was used in conjunction with DFT to search for the global minimum structures of oxidized graphene in different oxidation states [29]. Initially, they consider only an arrangement of epoxide groups on single-layer graphene. Using this starting point, they found two new low-energy semiconducting phases of the fully oxidized graphene.

In the present work we developed a program endowed with the ability to self-adapt as it searches for the ground state of both neutral and charged trans-polyacetylene molecules. Specifically, a systematic numerical investigation is performed to find the positions of each monomer describing the minimal energy (equilibrium) configuration within the Su–Schrieffer–Heeger (SSH) model. In these materials, solitons and polarons are the typical structures responsible for charge transport (free charge carriers) [31–33]. All the properties of the charge carrier including dynamics [34,35], photogeneration mechanisms [36], processes of excitons dissociation, and charge carriers recombination [36–38] use ground state structures as a starting point. The advantage of our method is that GA convergence is reached in a considerably smaller number of iterations when compared with the usual SCF.

We show that results obtained using GA are equivalent to those obtained using conventional SCF models. To do this comparison we used polaron and soliton type solutions. In some cases, SCF-based

* Corresponding author. Fax: +55 61 33072363.
  *E-mail addresses:* gargano@unb.br, gargano@fis.unb.br (R. Gargano).

methods has convergence difficulties because these methods walk through the phase space based on the initial configuration guess. On the other hand, GAOT is designed to search the global minimum of the system regardless of the initial conditions. Due to its robustness, GAOT always converges to the global minimum avoiding the premature convergence to a local minima. These results show that GAOT combined with the SSH model can be successfully applied to find the ground state structure of organic semiconductors reducing the number of iterations needed to reach convergence.

## 2. Methodology

We choose a polyacetylene chain in trans configuration to demonstrate the prediction of the ground state structure of a conjugated polymer chain of finite length, using the methodology developed here. An SSH-type Hamiltonian [39,40] is used to describe the polymer,

$$H = -\sum_{n,s} t_{n,n+1}\left(C_{n+1,s}^{\dagger}C_{n,s} + h.c\right) + \sum_n \frac{K}{2}y_n^2 + \sum_n \frac{p_n^2}{2M}. \tag{1}$$

Here $n$ indexes the sites. The operator $C_{n+1,s}^{\dagger}$ creates and the operator $C_{n,s}$ annihilates a $\pi$-electron at the $n$-th site with spin $s$. $K$ is the harmonic constant due to the $\sigma$ bonds. $M$ is the mass of a CH group. $y_n \equiv u_{n+1} - u_n$, where $u_n$ is the lattice displacement of the atom at the $n$-th site along the chain. $p_n$ is the momentum conjugated to $u_n$ and $t_{n,n+1}$ is the hopping integral, given by

$$t_{n,n+1} = (t_0 - \alpha y_n), \tag{2}$$

with $t_0$ being the hopping integral of a $\pi$-electron between nearest neighbor sites in the undimerized chain, and $\alpha$ is the electron-lattice coupling constant. The parameters chosen in this work are $t_0 = 2.5$ eV, $M = 1349.14$ eV $\times$ fs$^2$/Å$^2$, $K = 21$ eVÅ$^{-2}$ and $\alpha = 4.1$ eVÅ$^{-1}$. These values have been used elsewhere [34–38] with a good track record and the results are expected to be valid for other conjugated polymers. Nevertheless, it should be pointed out that the use of the GAOT in three-dimensional models of organic semiconductors is straightforward. One should just substitutes the SSH by the chosen model and takes $y_n$ as the distance between atoms.

The GAOT finds the set of $u_n$'s that provides the lowest energy starting from a randomly generated set of $u_n$'s. The population is composed of $N_{pop}$ individuals. Each individual has a genetic code with $N$ genes ($u_n$). These genes are represented by real numbers. The initial population genes are randomly generated. Figure 1 presents a diagrammatic representation explaining how the GAOT works.

Our GAOT uses common genetic algorithm operations: selection, matching with recombination and mutation. The probability of each individual to be selected for recombination takes place by a roulette spinning depending on its fitness [34]. We selected $N_{pop}/2$ individuals (parents) that will generate, through the recombination operation, $N_{pop}/2$ new individuals (offspring). The new generation, with $N_{pop}/2$ parents and $N_{pop}/2$ new strings (offsprings) maintains the population with a fixed number $N_{pop}$.

The original aspect of our genetic algorithm concerns the crossing of genes. Since each gene ($u_n$) is represented by a real number, whenever two corresponding genes of two parents are chosen to recombine, the offspring genes are determined as follows: each pair of real numbers generates two new numbers following a GAUSSIAN probability distribution centered at their mean value and with a dispersion proportional to their difference. This procedure warrants the similarity between the genes of the parents and offsprings, but it also enables the possibility of the offspring genes being higher or lower than their parent values.

The mutation operation is also performed following a GAUSSIAN probability distribution. This time it is centered at a given gene, $u_n$. The program selects randomly one gene belonging to a given individual to mutate. The probability of mutation of a given gene equals 0.01% per generation.

It is quite convenient that the global minimum solution should be found only by a group of individuals in the population. When the entire population converges prematurely to a single solution, this solution may be a local minimum. This is called premature convergence and it can be avoided by the linear scaling. This procedure enhances the probability that several minima will coexists in the population [41]. It introduces a parameter to control the selection pressure, defined as the degree to which the better
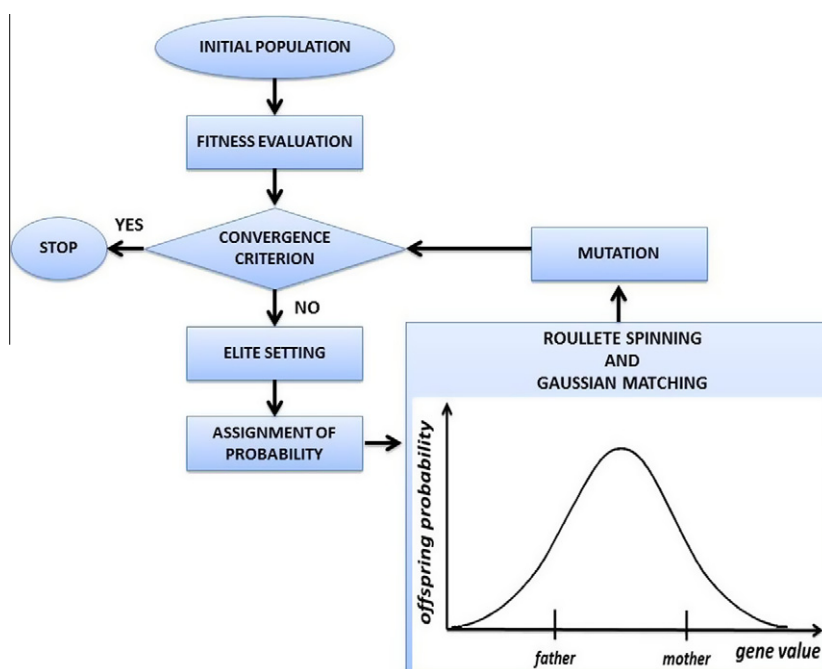


**Figure 1.** A diagrammatic representation explaining how the GAOT works.