Materials Characterization 100 (2015) 143–151

Contents lists available at ScienceDirect

Materials Characterization

journal homepage: www.elsevier.com/locate/matchar

Differential evolution: Global search problem in LEED-IV surface structural analysis

V.B. Nascimento [a,b,1], E.W. Plummer [a,*]

[a] Department of Physics and Astronomy, Louisiana State University, Baton Rouge, LA 70803, United States
[b] Center for Computation and Technology, Louisiana State University, Baton Rouge, LA 70803, United States

ARTICLE INFO

Article history:
Received 4 August 2014
Received in revised form 14 December 2014
Accepted 16 December 2014
Available online 18 December 2014

Keywords:
Low energy electron diffraction
Surface structure
Search problem
Global optimization

ABSTRACT

The search process associated with the quantitative theory–experiment comparison in Low Energy Electron Diffraction surface structural analysis can be very time consuming, especially in the case of complex materials with many atoms in the unit cell. Global search algorithms need to be employed to locate the global minimum of the reliability factor in the multi-dimensional structural parameter space. In this study we investigate the use of the Differential Evolution algorithm in Low Energy Electron Diffraction structural analysis. Despite the simplicity of its mechanism the Differential Evolution algorithm presents an impressive performance when applied to ultra-thin films of $BaTiO_3(001)$ in a theory–theory comparison. A scaling relation of $N^{(1.47 \pm 0.08)}$ was obtained, where N is the total number of parameters to be optimized.

1. Introduction

The Low Energy Electron Diffraction (LEED) technique is the most used experimental approach for surface structure determination. Impinging low energy electrons (20 to 500 eV) interact strongly with the atoms in the top surface layers, drastically reducing their effective penetration depth and consequently making LEED very surface sensitive. However this strong interaction (multiple scattering) excludes the possibility of a simple direct inversion of experimental data like in a Patterson function formalism [1]. Consequently, LEED structure determination needs to be performed in an indirect way.

LEED surface structural determination is performed by a quantitative comparison of experimentally collected curves of the intensity from diffracted spots as a function of incident electron energy [I(V)] with theoretically calculated ones as function of structure. This comparison is made quantitative by the use of a so-called reliability factor (R-factor). The lower the final R-factor achieved, the more reliable are the results of the structural determination. In summary, the surface structure determination by LEED turns into a search problem in which one needs to explore a multi-dimensional parameter space of surface structural parameters and find the global minimum of the R-factor among all existing local minima.

This search process is usually performed by applying local search algorithms (Powell-Directed Search, Simplex, Levenberg-Marquardt) [2], which are able to explore only a small fraction of the parameter space and usually locate only a minimum (most likely local) close to the search starting point. As a consequence the employed local search method needs to be launched from different initial points in the parameter space in order to be able to locate the global minimum among all other local minima, turning the search process into a very time consuming process. An alternative is the use of global search algorithms, like genetic algorithms [3] and simulated annealing method [4], that are able to efficiently explore the multi-dimensional parameter space and locate the global minimum of the R-factor without getting trapped in local minima.

In order to be suitable to the LEED search problem a global search algorithm needs to possess two main features. 1) It needs to present a high probability of locating the global minimum among all existing local minima in the multi-dimensional parameters. 2) The method needs to present a favorable scaling behavior with the number of total structural parameters being optimized. This scaling will evaluate the computational effort to locate the global minimum as a function of the number of optimized structural parameters.

The typical large size and complexity of the structure of new materials like transition metal oxides makes the LEED structural search process even more difficult. The surface unit cell of these materials will be large, with many distinct atoms. For example, a reconstructed surface for a layered transition metal compound may have over 20 atoms in the unit cell, with around 60 structural parameters (depending on symmetry) to be determined by the experiment–theory comparison. After one finds a physically feasible structural model for the surface (hopefully the actual structure), based on all available information based on experimental and/or theoretical results, a large number of structural parameters still need to be optimized in order to obtain the final structure. The use of global search algorithms for the surface analysis of such materials

* Corresponding author.
E-mail address: wplummer@phys.lsu.edu (E.W. Plummer).
[1] Present address: Departamento de Física, ICEx-UFMG, CP 702, Belo Horizonte, MG CEP 30123-970, Brazil.

http://dx.doi.org/10.1016/j.matchar.2014.12.020
1044-5803/

turns out even more important than for less complex metallic or some semiconductor systems for instance. A review of all the global search methods applied to the LEED structural determination can be found in the recently published article by Soares and collaborators [5].

Several distinct approaches have been proposed for the use of global search algorithms in LEED structural analysis.

- The simulated annealing (SA) algorithm was investigated by Rous [6] in a pioneering attempt to apply a global search method to the LEED search problem. A theory–theory comparison for the Ir(110)(2 × 1) system was used to evaluate the scaling relation, yielding an unfavorable scaling given by $N^{6.0}$, where N stands for the total number of optimized structural parameters.
- Motivated by this first SA approach for the LEED global search problem, Nascimento and co-workers [7,8] investigated the fast simulated annealing (FSA) approach [9]. A Cauchy–Lorentz random step distribution function is employed in the FSA algorithm contrasting with the Gaussian or uniform distributions typically used in the SA approach. By employing the Cauchy–Lorentz distribution random large steps will be taken during the search process. A very favorable scaling given by $N^{1.0}$ was obtained in a theory–theory comparison for the CdTe(110) system, with a drawback of a decreasing convergence probability as N increases. Correa et al. worked on implementing and testing the generalized simulated annealing (GSA) approach for the LEED search problem [10]. The LEED GSA approach was based on the non-extensive statistical mechanics by Tsallis [11–13]. Within this approach several distinct distribution functions defined by the q_V parameter have been tested. This parameter basically defines the distribution function for the size of random steps taken during the search process. The obtained results indicated that the FSA approach (special case for q_V=2.0, defining a Cauchy–Lorentz distribution) had the optimal performance, again indicating a linear scaling $N^{1.0}$ in agreement with the previous results obtained by Nascimento et al. [8].
- Simulated annealing was recently incorporated in the computational code implementing the new frozen LEED (FL) perturbative approach for tensor calculation proposed by Yu and Tong [14].
- A simulated annealing algorithm is also one of the global search methods implemented in a LEED code completely written in the C language (CLEED) by Held et al. [15].
- A random sampling algorithm was proposed by Kottcke and Heinz [16]. This method allows only for downhill moves during the search process, in contrast with the simulated annealing approach and represents a compromise between global and local search. An effective $N^{2.5}$ scaling was obtained for the Ir(110)(2x1) system.
- The combinatorial simultaneous optimization (SO) method was proposed by Blanco-Rey and de Andres [17]. Results obtained with a theory–theory comparison for the Ir(110)(2x1) missing-row surface indicates a very high convergence probability but a not effective scaling of $N^{4.1}$.
- Döll and Van Hove [18] investigated the application of the genetic algorithm [3] to the LEED search problem. The genetic algorithm mimics the mechanism of natural evolution. No scaling relation was explored in this work.
- A more recent work by Alvarenga and co-workers [19] has indicated that the genetic algorithm can be a very effective tool for the search problem if combined with local search methods since under this combination a very favorable scaling of $N^{1.5}$ was obtained for the Ni(111)($\sqrt{3}x\sqrt{3}$)–R30^0–Sn system.
- Another interesting approach to the LEED search problem, based on pattern search methodology (generalized pattern search), was recently proposed by Zhao et al. [20]. Obtained results indicated a better performance than the genetic algorithm, but no scaling was reported.

In spite of promising results obtained with the previously mentioned global search methods, it is important to explore new possible algorithms to tackle the difficult LEED search problem. In this work we investigate the use of the Differential Evolution algorithm (DE) [21,22] for the LEED search problem. As an initial test the Differential Evolution method has been applied to the optimization of the surface structure of the $BaTiO_3$(001) ultra-thin films [23] in a theory–theory comparison. A scaling relation has been obtained and showed a promising performance as characterized by a $N^{(1.47 \pm 0.08)}$ scaling, which rivals performance results from previously investigated global search algorithms (see Table 1).

2. Differential evolution algorithm

Differential Evolution (DE) algorithm is a new heuristic approach for minimization/optimization problems that belongs in the evolutionary group of algorithms [21,22,24,25]. Its main feature consists on the fact that the vectors in the N-dimensional parameter space are mutated by adding weighted random vector differentials. This feature contrasts with regular evolutionary algorithms that use a probability function (Gaussian, Cauchy, fuzzy) to generate perturbing changes in the vectors during the mutation step. DE will use the present population to generate the vectors increments in magnitude and orientation.

The differences between DE and GA algorithms will be better understood by comparing our description of DE to be following presented with references on genetic algorithms [3,5]. However we will present a brief discussion about Genetic Algorithms in order to allow a better understanding of the differences between this class of algorithms and the Differential Evolution method.

Genetic algorithms keep closely to the metaphor of genetic reproduction. Even their language is mostly the same of genetic reproduction. In both cases there is the use of terms like chromosomes, genes, (genes are distinct alphabets) and crossover (crossover is fairly close to a low-level understanding of genetic reproduction). An initial population of individuals is randomly created as a starting generation. The information about the parameters to be optimized is stored in chromosomes (genes) that are coded as bit strings or float numbers depending on the algorithm implementation. An evaluation of fitness is calculated for every individual of this first population. The fitness parameter is defined as a constant real number minus the cost function to be minimized, thus turning the minimization problem of the cost function into a maximization process of the fitness parameter. After this initial population is created and characterized a loop is performed over generations until a certain stop (convergence) criterium is fulfilled: *i*) Pairs of individuals are selected as parents for the next generation based on their fitness. *ii*) Two offspring are created during a crossover process for every pair previously selected. *iii*) Mutation is applied to every new individual and fitness is calculated. *iv*) All the generated offspring will consist on the next population and the parents will die. *v*) A mechanism called elitism may be applied depending on the algorithm implementation. Elitism process takes part of the top best individuals, based on fitness, to the next generation.

Although DE algorithm apparently presents the same basic operational steps (mutation, crossover and selection) present in genetic algorithm implementations, the order and way these steps are performed

Table 1
Summary of results obtained with the application of global search methods to the LEED problem. Applied search method, explored surface system (theory–theory comparison) and scaling relation are presented. Only results that included a scaling behavior have been included in this table.

Method	System	Scaling
SA(Rous)	Ir(110)(2 × 1)	$N^{6.0}$
FSA/GSA	CdTe(110)	$N^{1.0}$
Random sampling	Ir(110)(2 × 1)	$N^{2.5}$
SO	Ir(110)(2 × 1)	$N^{4.1}$
GA + local optimization	Ni(111)($\sqrt{3}x\sqrt{3}$)–R30^{0}Sn	$N^{1.5}$
DE	Ultra-thin films of $BaTiO_3$(001)	$N^{(1.47 \pm 0.08)}$