The generalized simulated annealing algorithm in the low energy electron diffraction search problem

2005 J. Phys.: Condens. Matter 17 1

(http://iopscience.iop.org/0953-8984/17/1/001)

INSTITUTE OF PHYSICS PUBLISHING JOURNAL OF PHYSICS: CONDENSED MATTER

J. Phys.: Condens. Matter **17** (2005) 1–16 doi:10.1088/0953-8984/17/1/001

The generalized simulated annealing algorithm in the low energy electron diffraction search problem

Eduardo dos R Correia[1], Von B Nascimento[1,3,4], Caio M C de Castilho[1,5], Antonio S C Esperidião[1], Edmar A Soares[2] and Vagner E de Carvalho[2]

[1] Grupo de Física de Superfícies e Materiais, Instituto de Física, Universidade Federal da Bahia, Campus Universitário da Federação, 40210-340, Salvador, BA, Brazil
[2] Departamento de Física, Universidade Federal de Minas Gerais, 30123-970, Belo Horizonte, MG, Brazil

E-mail: caio@ufba.br

Received 8 October 2004, in final form 25 November 2004
Published 10 December 2004
Online at stacks.iop.org/JPhysCM/17/1

Abstract

We present in this work results concerning the application of the generalized simulated annealing (GSA) algorithm to the LEED search problem. The influence of the visiting distribution function (defined by the so-called q_V parameter) in the effectiveness of the method was investigated by the application of the algorithm to structural searches for optimization of two to ten parameters in a theory–theory comparison for the CdTe(110) system. Results, obtained with the scaling relation and probability of convergence as a function of the number of parameters to be varied, indicate the fast simulated annealing (FSA) ($q_V = 2.0$) approach as the best search machine.

1. Introduction

The use of electron diffraction as a tool for probing surfaces has been the most effective approach for surface structural determination [1]. Due to the large electron cross section and multiple-scattering process that occurs in the first atomic layers, the surface structural determination using the low energy electron diffraction (LEED) technique is performed using an indirect methodology [2, 3]. A set of theoretically calculated curves of the intensity of the elastically scattered beam versus the electric potential used to produce the incident electron beam ($I(V)$ curves) is compared with a set of experimentally collected curves. This comparison is performed in a quantitative way, by using the so-called *reliability factor* or *R-factor* [2].

To perform the theoretical calculations, a set of values for a series of structural and non-structural parameters has to be assumed, corresponding to a theoretical model conceived in

[3] Present address: Department of Physics and Astronomy, The University of Tennessee, Knoxville, TN 37996, USA.
[4] Present address: Solid State Division, Oak Ridge National Laboratory (ORNL), Oak Ridge, TN 37831, USA.
[5] Author to whom any correspondence should be addressed.

0953-8984/05/010001+16$30.00

order to mimic the experimental conditions. As the calculated curves are highly sensitive to the parameter values, these must be varied in the search for the best coincidence between theory and experiment. A perfect coincidence among the two sets of curves is, in fact, impossible, as a result of experimental errors and approximations taken in the conception of the theoretical models. It is necessary to assume a minimum of tolerance as corresponding to an effective coincidence among the two sets of curves. While searching this, the set of parameters associated with the calculations is varied in the proposed models in an attempt to find a minimum for the R-factor. This converts the structural determination via LEED into a minimization or search problem.

The strong interaction between the impinging electrons and the surface atoms results in calculated $I(V)$ curves that are extremely sensitive to the values of the parameters. This leads us to infer that the eventual coincidence between calculated and experimental curves can be interpreted as an indication that the assumed values for the parameters in the calculations do correspond to the ones experimentally probed. For carefully obtained experimental curves, the search for agreement is theoretically done in two ways: (a) by perfecting and improving the theoretical model which gives support to the calculations and (b) by varying the structural and non-structural parameters for the adopted theoretical model. Case '(b)' constitutes the route more systematically used for structural determination. In this case, as the number of possible structures is high, it is essential to establish a route for varying the parameters. The values for the R-factor at each point of the parameter-space form a hyper-surface with several maxima and minima. In the search process, the structural determination is considered as achieved when it is possible to identify the global minimum among the several local minima for the R-factor. This factor is then the cost-function to be minimized in the search process. Despite being necessary to establish a route for determining a local minimum [4], a critical aspect of the search is how to be sure that a global minimum has been achieved and not just a local one.

Since the first days of LEED surface structural determination, several methods of identifying the aimed for global minimum have been used, which can be classified in two broad categories: (I) methods based on trial and error search and (II) methods based on some systematic routes for minimization of the cost function. The first structural determinations by LEED involved a trial and error process [3, 5] with the values for the parameters being assumed from indirect hints and guesses in the process of trying to fit the curves. However, the parameters are correlated in such a way that adjusting one of them implies a readjustment of the others. This limitation restricted application of LEED to simple structures, and forced us to establish a more systematic route for the adjustment process, i.e., an automated search. The first structural determination using an automated process was performed by Cowell and de Carvalho [6]. That work used a variation of the steepest descent method, i.e., the partial derivatives of the R-factor with respect to the N parameters were used as a guide in the search for the minimum R-factor.

A few months before the work by Cowell and de Carvalho, the theoretical work by Rous *et al* [7], the tensor LEED approach, was published. This method, although not a proper search method, is useful for structure refinement with the adjustment of parameters being done in a scale of a fraction of an ångström. The method has the ability to calculate the partial derivatives of the scattered intensities, reducing the number of guessed structures previously explored with trial and error procedures. After that, several approaches, sometimes a combination of methods, have been used for LEED structural determinations: tensor LEED with gradient methods [7, 8], the so-called direct methods using the approach adopted by x-ray crystallography [9–11] and methods using the least-squares procedures [12]. Each of these methods has its own deficiencies, either by being too time consuming for computational calculation or by not being able to distinguish among local and global minima. In order to avoid

the possible misidentification of a local minimum as a global one, Rous [13] applied the global optimization simulated annealing (SA) algorithm to the search problem in LEED. His results, from a theory–theory comparison, present, for the cost function, a scaling relation given by N^6, with N being the number of parameters. Motivated by this first tentative attempt at establishing a global search method for the LEED problem, Nascimento *et al* [14, 15] has investigated an alternative approach to the SA, the so-called fast simulated annealing (FSA) [16], where the random step distribution function is a Cauchy–Lorentz function, instead of a Gaussian (or uniform) one, as used in the SA approach. This modification has proven extremely useful, since the scaling factor becomes linear with the number of parameters (N^1), in a CdTe(110) theory–theory comparison. The FSA method was applied to real structural determinations (theory–experiment comparison) for the systems Ag(111), Ag(110) and CdTe(110) [14, 15, 17].

Another approach for identifying a global minimum in a LEED search process was proposed by Kottcke and Heinz [18], with optimization for the theory–experiment fitting process using a random sampling algorithm. In contrast to the SA or FSA approaches, the random sampling algorithm only takes into account downhill moves, with a multiple launching process (several starting structures), becoming a compromise between global and local search methods. This procedure turned out to have an $N^{2.5}$ scaling with the number of parameters.

Motivated by the relative success of the SA algorithm, as employed by Rous [13], which requires a high number of structures to be tested, Döll and Van Hove [19] have proposed another global search algorithm, called a genetic algorithm (GA) or evolutionary algorithm [20], a method that mimics the natural evolution of living organisms. Despite this first suggestion of applying the GA to the LEED search, this approach still requires a more systematic investigation to determine a scaling relation within the perspective of its possible transformation into an effective and adequate search method for LEED structural determination.

In order to be effective, a global search method has to present two main features: (1) a high probability of locating the global minimum of the R-factor among all other local minima; (2) a favourable scaling relation of number of parameters to be varied during the search process.

In this work we present the results obtained with the investigation of the influence of the distribution function used in the generation of the random steps on the performance of the simulated annealing (SA) algorithm, when applied to the LEED search problem. Aiming for this, we have adopted the *generalized simulated annealing* (GSA) approach proposed by Tsallis and Stariolo [21]. In section 2 we discuss the general characteristics of the GSA. The results of its application to a specific LEED problem are presented and discussed in section 3, while section 4 contains the conclusions.

2. The generalized simulated annealing algorithm (GSA)

The mechanism of the simulated annealing method is based on the process by which molten metals are gradually cooled and annealed, commonly used in metallurgy. In such a process, the temperature of the molten metal is slowly reduced, in order to produce a decreasing mobility of the atoms. If this cooling scheme is sufficiently slow, hopefully the atoms will be able to line themselves up, leading to an organized geometry, reaching the crystalline state, which corresponds to the global minimum of the thermodynamical energy. In the simulated annealing method, an artificial 'temperature' is introduced as a source of stochasticity, in an attempt to create an 'artificial dynamics', able to prevent the search process (associated with an energy or a cost function minimization) from getting trapped in a local minimum. At the final steps of the optimization process, the 'temperature' is near zero, and the search process hopefully results inside the basin of the global minimum (or unfortunately in one of the local minima) and the search procedure behaves asymptotically as a gradient descent local search method.

The primary criticism of the simulated annealing algorithm is that it is a slow convergence method. The main challenge then, in improving the simulated annealing method, consists in cooling the search process as fast as possible, without increasing the probability of getting trapped in any of the local minima. In other words, there is a need for the quickest cooling scheme which assures, 100%, the probability of locating the global minimum.

Kirkpatrick and co-workers [22] proposed a first approach for this problem, following the quasi-equilibrium Boltzmann–Gibbs statistics. In this approach, the random steps of the search procedure were taken according to a Gaussian visiting distribution in the neighbourhood of the actual search point. Downhill movements, associated with a decrease in the energy or other cost functions, are always accepted. If the step is uphill, it will be accepted or not according to an acceptance probability, which is assumed to be the Boltzmann–Gibbs one. A necessary and sufficient condition for assuring a probability of locating the global minimum is that, during the cooling scheme, the 'temperature decreases' logarithmically with time, i.e, simulated annealing steps [23]. This approach is usually called *classical simulated annealing* (CSA) or *Boltzmann machine*.

Another approach was proposed by Szu and Hartley [16], which used a Cauchy–Lorentz visiting distribution in the generation of the random SA steps. With this semi-local distribution, the random steps are frequently local, but with the occurrence of occasional long ones. The acceptance probability, as in the *Boltzmann machine*, is assumed to be the Boltzmann–Gibbs one. The cooling scheme can be much faster, with the temperature decreasing with the inverse of time. The algorithm is commonly called *fast simulated annealing* (FSA) or the *Cauchy machine*.

More recently, Tsallis and Stariolo [21] proposed a generalization of both annealing schemes (CSA and FSA) inspired in the non-extensive Tsallis statistics [24, 25], with an additional advantage of being able to provide annealing schemes quicker than the one proposed by Szu. In this scheme, called *generalized simulated annealing*, the visiting distribution used in the definition of the random steps ($x \to x + \Delta x$) is given by

$$g_{q_{\mathrm{V}}}(\Delta x) = \left(\frac{q_{\mathrm{V}}-1}{\pi}\right)^{\frac{D}{2}} \frac{\Gamma\left(\frac{1}{q_{\mathrm{V}}-1}+\frac{D-1}{2}\right)}{\Gamma\left(\frac{1}{q_{\mathrm{V}}-1}-\frac{1}{2}\right)} \frac{[T_{q_{\mathrm{V}}}(t)]^{-D/(3-q_{\mathrm{V}})}}{\left(1+(q_{\mathrm{V}}-1)\frac{(\Delta x)^2}{[T_{q_{\mathrm{V}}}(t)]^{2(3-q_{\mathrm{V}})}}\right)^{1/(q_{\mathrm{V}}-1)+(D-1)/2}}, \tag{1}$$

with q_{V} being a parameter defining the curvature for the distribution function (the larger the q_{V}, the more open is the distribution function, as can be seen in figure 1), Γ being the gamma function, t the discrete time corresponding to the search step and D the number of parameters being optimized.

With such a visiting distribution, an annealing scheme can be defined to assure a probability of convergence equal to unity:

$$T_{q_{\mathrm{V}}}(t) = T_{q_{\mathrm{V}}}(1)\frac{2^{q_{\mathrm{V}}-1}-1}{(1+t)^{q_{\mathrm{V}}-1}-1} \sim T_{q_{\mathrm{V}}}(1)\frac{2^{q_{\mathrm{V}}-1}-1}{t^{q_{\mathrm{V}}-1}} \qquad (t \to \infty), \tag{2}$$

with $T_{q_{\mathrm{V}}}(1)$ being the initial 'temperature' value.

A generalized acceptance probability for the step $x_l \to x_{l+1}$ was also proposed by Tsallis and Stariolo [21]:

$$P_{q_{\mathrm{A}}}(x_l \to x_{l+1}) = \begin{cases} 1 & \text{if } E(x_{l+1}) < E(x_l) \\ [[1+(q_{\mathrm{A}}-1)(E(x_{l+1})-E(x_l))/T_{q_{\mathrm{A}}}]^{1/(q_{\mathrm{A}}-1)}]^{-1} & \text{if } E(x_{l+1}) \geqslant E(x_l), \end{cases} \tag{3}$$

with q_{A} standing for a control parameter for the acceptance probability $P_{q_{\mathrm{A}}}(x_l \to x_{l+1})$ and $T_{q_{\mathrm{A}}}$ for the 'temperature'.

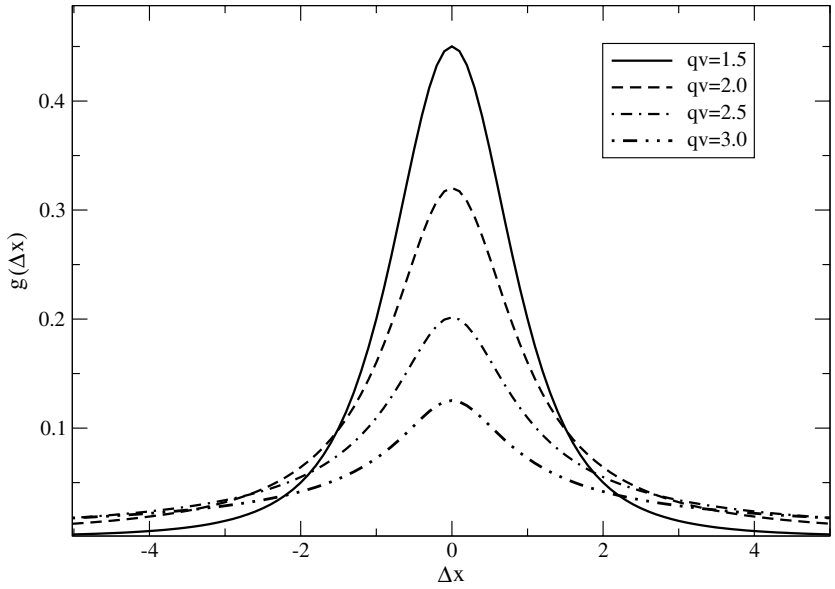

Figure 1. Generalized simulated annealing visiting distribution functions for different q_{V} values. The curves are normalized with respect to the area between the curve and the horizontal axis.

It can be shown that, for $q_{\mathrm{A}} = 1$, the acceptance function reduces to the well known Boltzman–Gibbs one [21]. In addition to this, assuming q_{V} values equal to 1.0 and 2.0, it is able to reduce the distribution g_{q_v} to the Gaussian and Cauchy–Lorentz distributions, respectively. By adjusting the parameters q_{A} and q_{V} to adequate values, it is able to obtain, through the generalized simulated annealing, both Boltzmann and Cauchy search machines [21]. Besides this, by varying the q_{V} and q_{A} parameters, associated with the distribution function and acceptance probability, respectively, it is possible to define many search machines other than the Boltzmann and Cauchy ones, which can be suitable for specific optimization problems. The influence of the q_v parameter on the slope of the distribution functions can be seen in figure 1, where distribution functions for several q_v values are presented.

As discussed before, the simulated annealing has been applied to the search problem associated with the surface structural determination by LEED [13–15]. With this minimization problem, it is hopefully possible to locate the global minimum of the so-called reliability factor (R-factor) among all other local minima. The promising results obtained with the application of the simulated annealing (more specifically with the FSA approach [14, 15]) to the LEED search problem justify the search for optimized SA search machines, which can be defined with the generalized simulated annealing variation. Aiming to identifying more efficient search machines, an investigation has been performed, using a theory–theory comparison search problem for the CdTe(110) system. In the next two subsections, descriptions of the dynamical LEED calculations, for the GSA implementation, as well as of the methodology employed in the GSA algorithm investigation, are presented.

2.1. Implementation of a GSA code for LEED application

The theoretical LEED calculations, concerning a theory–theory comparison for the CdTe(110) system, were performed using the muffin-tin model for the potential of the Cd and Te atoms, with muffin-tin radius values of 1.40 Å being assumed for both. The numerical integration of

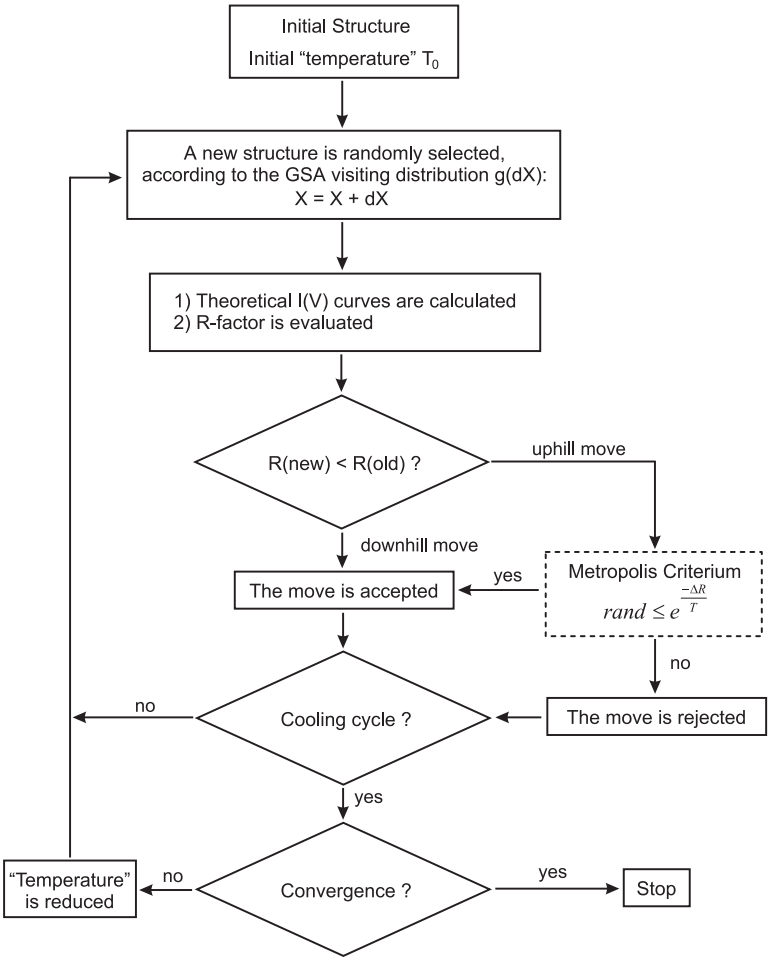

Figure 2. Flowchart for the generalized simulated annealing implementation in LEED.

the radial part of the Schrödinger equation in the muffin-tin spheres was used in the calculation of the phase shifts. A modified version of the Van Hove–Tong conventional LEED code [3] was used in the dynamical calculation of the $I(V)$ curves, using the reverse forward scattering approximation. The calculations were performed in a Linux PC cluster (IF-UFBA) composed by 16 dual-processor Pentium III-1.0 GHz PCs. Eight phase shifts were used in the calculations, with an inner potential of $V_0 = (-10+5\mathrm{i})$ eV, optimizing its real part during the search process. A Debye temperature of 140 K was adopted for the Cd and Te atoms in the surface and bulk layers.

The scheme previously proposed by Rous [13] was adopted in the implementation of the simulated annealing algorithm within the modified Van Hove–Tong LEED code, while the necessary changes for the implementation of the generalized simulated annealing were based on the initial proposal of Tsallis and Stariolo [21]. A flowchart for the LEED implementation of the GSA search algorithm is presented in figure 2. A key point in the implementation of the GSA approach is the effective way of generating the random numbers with the visiting distribution $g(\Delta x)$, given by equation (1). In this work, the multidimensional random steps

were taken by updating one parameter (dimension) at a time, i.e., using a unidimensional visiting distribution ($D = 1$). This procedure [21] is simpler than updating the whole state vector (all parameters) at the same time.

2.2. GSA performance analysis—methodology

A global search method, considered as an adequate one, to be applied to the LEED structural problem must present a high probability of locating the global minimum of the R-factor among all the other local minima, and a favourable scaling relation of the number of parameters. The scaling relation represents how the computational effort, necessary for one search algorithm to locate the global minimum of the R-factor, increases with the number of structural parameters (N) being optimized, i.e., its efficiency as a function of the dimensionality of the parameter space to be explored.

The exhaustive search method (based on a trial and error procedure) scales exponentially with N [26], representing a non-polynomial (NP) problem, while the directed-search algorithms present approximately an N^2 scaling [27]. The first application of the simulated annealing method to the LEED problem performed by Rous [13] has pointed out a not very favourable scaling (N^6). Döll and Van Hove have investigated the application of the genetic algorithm to the LEED search problem [19], but a scaling behaviour was not obtained. The global search method proposed by Kottcke and Heinz [18] resulted in a more favourable scaling ($N^{2.5}$). The recent work of Nascimento and co-workers [14, 15], concerning the fast simulated annealing approach, has obtained a very favourable N^1 scaling. So, it seems interesting to apply the GSA method [21] to the LEED optimization problem since, with this algorithm, it is possible to explore other parameters (q_A and q_V) that may be important to adjust in order to get the best search engine for the LEED analysis.

In this work, the GSA method was applied to the structural search for the CdTe(110), in a theory–theory comparison. This choice was mainly motivated by the fact that the CdTe(110) surface presents a relatively complex reconstruction [6, 28], in which the first atomic layer presents a bond-length-conserving rotation and a contraction towards the substrate. Due to the characteristics of this reconstruction, the structure determination is a complex one, requiring the optimization of a relatively large number of parameters: ten parameters, if the first three atomic layers were allowed to vary, as shown in figure 3. Also the use of the CdTe(110) system in this investigation would enable a direct comparison with the previous results obtained by Nascimento *et al* [14, 15].

The methodology used here consisted of repeated GSA structural searches performed for the minimization of the Pendry R-factor (R_P) [3, 29], using the following values for the q_V parameter: 1.0, 1.5, 1.7, 2.0, 2.3, 2.5 and 2.7. For all the investigated cases, we have adopted a value of 1.0 for the q_A parameter, that corresponds to the classical Boltzmann–Gibbs acceptance probability. Initial 'temperatures' ($T_{q_V}(1)$) equal to 10 were adopted for every explored q_V [14, 15]. Typical results of these structural searches can be seen in figure 4, for the case of $q_V = 2.3$, where curves for three different starting points are shown. For each explored value of the q_V parameter, 30 structural searches (starting from different points in parameter space) have been executed for optimization of two, four, six, eight and ten structural parameters. The first four structural parameters, associated with the first layer, were varied in a range of 1.0 Å around their 'bulk terminated' values. The other six parameters, associated with the second and third layers, were varied in a total range of 0.5 Å, centred at 'bulk terminated' values.

The convergence criterion adopted was as follows: a search process was considered as having achieved convergence to the global minimum if the R_P factor reached a value lower

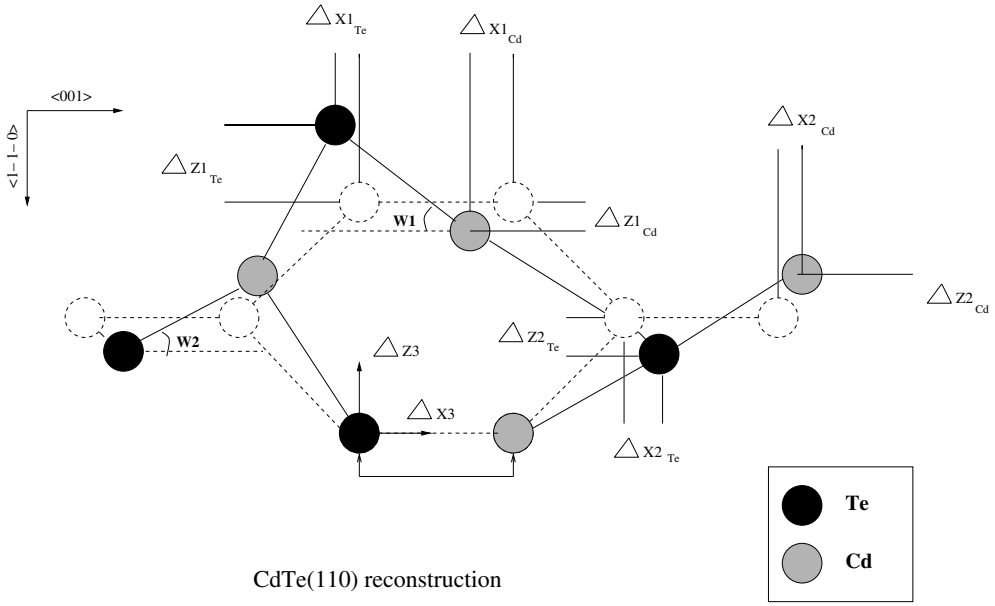

Figure 3. Representation of the complex reconstruction presented by the (110) surface of CdTe. The first four structural parameters, $\Delta X1_{\mathrm{Te}}$, $\Delta X1_{\mathrm{Cd}}$, $\Delta Z1_{\mathrm{Te}}$, $\Delta Z1_{\mathrm{Cd}}$, define the first layer rumple as well as the distance between the first and second layers. The $\Delta X2_{\mathrm{Te}}$, $\Delta X2_{\mathrm{Cd}}$, $\Delta Z2_{\mathrm{Te}}$, $\Delta Z2_{\mathrm{Cd}}$ parameters define the second layer rumple and the spacing between the second and third layers. Third layer related parameters $\Delta X3$, $\Delta Z3$ are assigned to displacements of the entire layer along $\langle 001 \rangle$ and $\langle \bar{1}\bar{1}0 \rangle$ directions, without any rumpling. The X and Z axes are defined along $\langle 001 \rangle$ and $\langle \bar{1}\bar{1}0 \rangle$ directions respectively. The total hypervolume explored in this investigation was $(2.0 \times 2.0 \times 2.0 \times 2.0 \times 0.5 \times 0.5 \times 0.5 \times 0.5 \times 0.5 \times 0.5)$ Å^{10}.

than 0.10. In a theory–theory comparison the global R-factor minimum presents a very deep and narrow basin and, for R_{P} values of less than 0.10, the optimized structural parameters present values very close to their optimum ones, with differences around just 0.01 Å.

The mean number of structures accepted by the Metropolis criterion, before convergence to the global minimum, for the optimization of two to ten parameters—and for each explored q_{V} value—was obtained. However, during the search process, several structures were rejected by the Metropolis criterion, and were not included in the accepted ones. In fact, theoretical calculations have been also performed for the rejected structures as well, and this also requires some computational effort. So, it would be more realistic [14] to obtain a scaling behaviour using the total number of trial investigated structures before convergence, which includes structures accepted and rejected during the search process, and not only the accepted ones. Therefore, in order to do this, the mean number of trial structures examined during the search process before convergence were also obtained for the optimization of two to ten parameters, for each explored q_{V} value.

Another important practical feature of a global search method consists of the probability of convergence to the global minimum, as a function of the number of explored parameters. A probability of convergence that decreases slowly with an increasing number of explored parameters plays a key role in the effectiveness of a global search algorithm. In the second part of this study, the probability of convergence to the global minimum as a function of the number of optimized parameters has also been obtained for each examined value of q_{V}. This

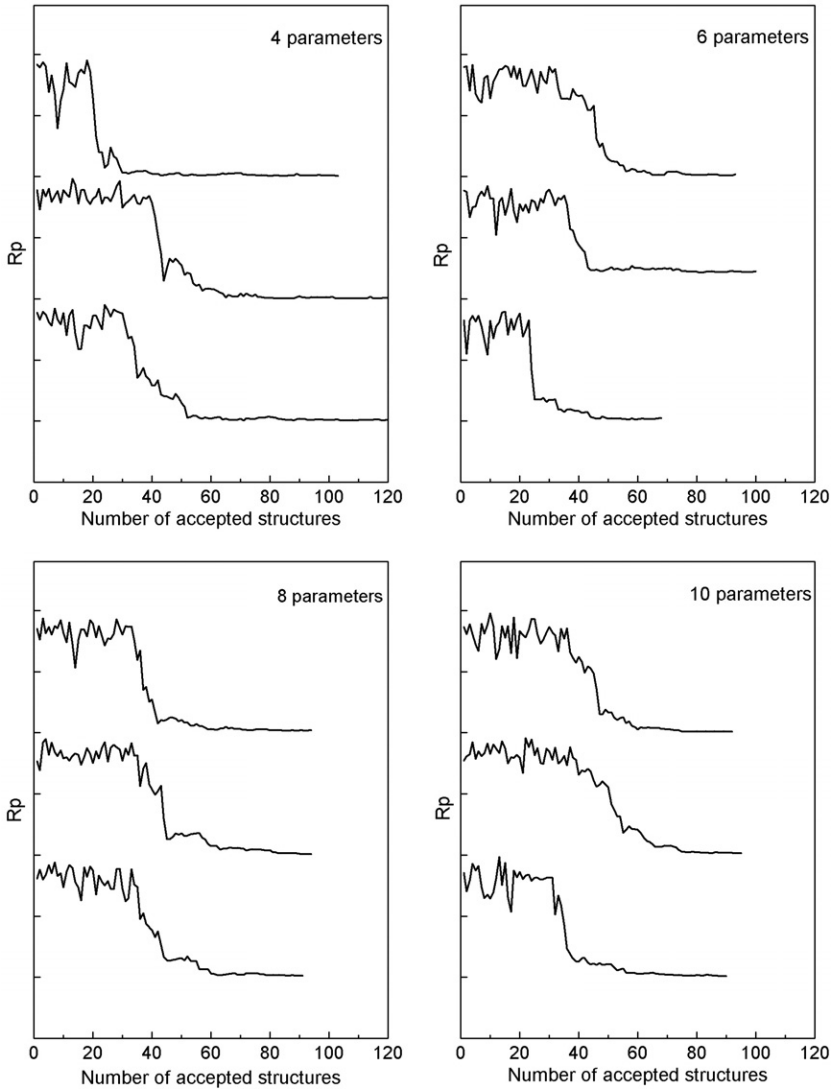

Figure 4. Typical GSA structural searches ($q_V = 2.3$) performed for the optimization of four, six, eight and ten parameters for the CdTe(110) system in a theory–theory comparison, obtained from three different initial points, for each number of parameters. Each structural search has started from a different initial point. The scaling relations, for the number of accepted and trial structures, as well as the probability of convergence have been obtained from the analysis of these structural searches.

probability of convergence was obtained as the ratio between the number of structural searches that has reached the global R_P minimum and the total number of structural searches performed.

3. Results and discussion

Scaling relations have been obtained for the number of accepted structures as well as for the number of trial structures. The final results are presented in figures 5 and 6 respectively. In order to get information about the total computational effort, only the results concerning the

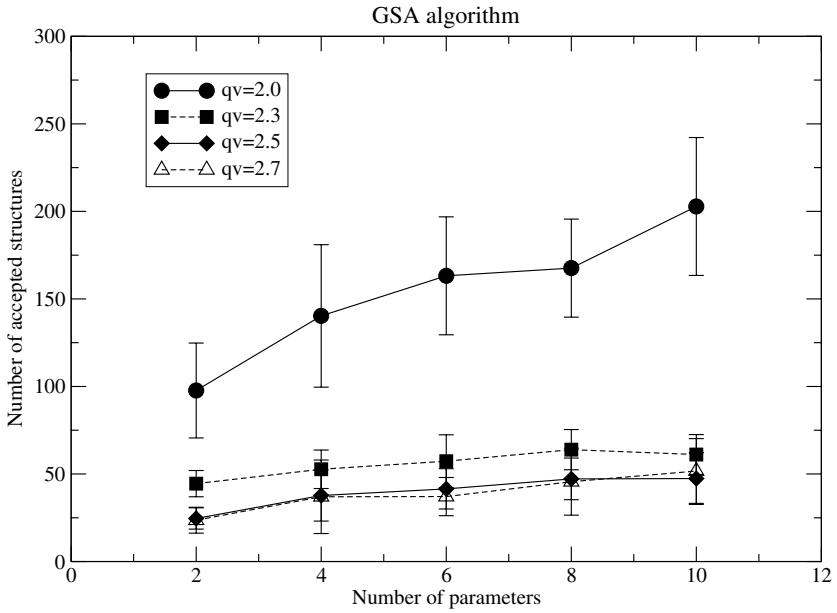

Figure 5. Mean number of accepted structures as a function of the number of varied structural parameters, for various values for the q_V parameter. The error bars were obtained from a standard mean deviation from the mean values.

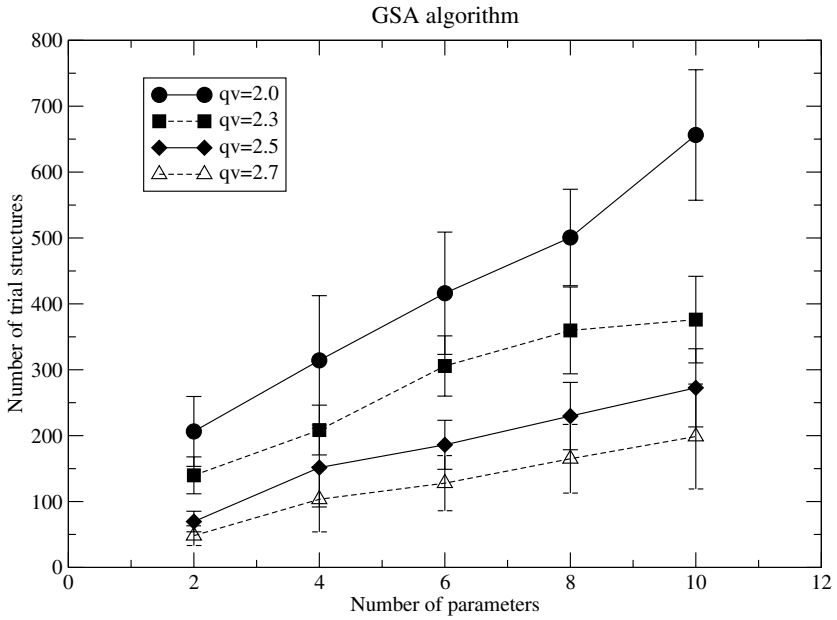

Figure 6. Mean number of trial structures (accepted and rejected) as a function of the number of varied structural parameters, for various values for the q_V parameter. The error bars were obtained from a standard mean deviation from the mean values.

total number of trial structures examined in the process will be discussed here. So, as can be seen in figure 6, the trial structure scaling behaviours, obtained for all investigated q_V parameter

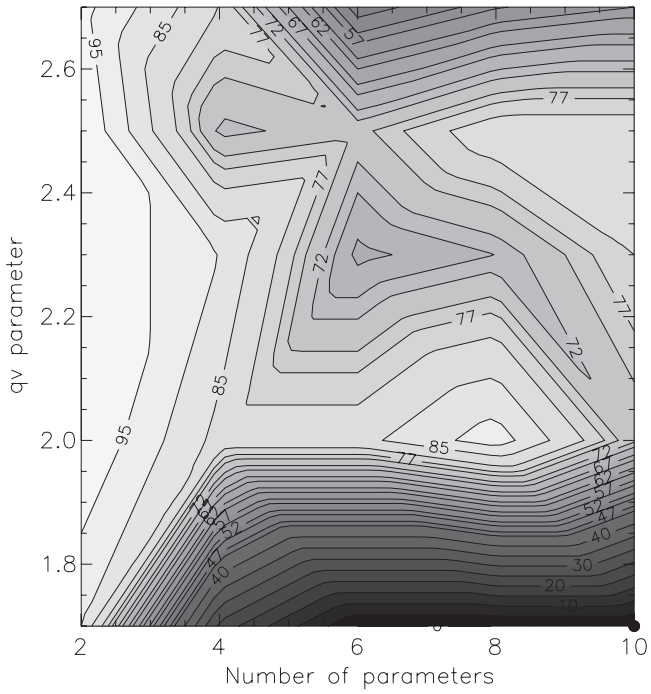

Figure 7. Contour plot of the probability of convergence (%) as a function of the number of varied structural parameters and of the q_V parameter. The grey scale varies from white (100%) to black (0%).

values (2.0, 2.3, 2.5 and 2.7), seem to be (within the associated error bars) linear (N^1). The results for q_V equal to 2.0 (corresponding to the FSA approach) agree with those obtained in the previous work by Nascimento and co-workers [15], with the advantage that they have been obtained in a larger range of investigated parameters (from two to ten). Despite the fact that all q_V value results indicates this favourable linear scaling, as can be inferred only by the trial structure results (figure 6), the higher the q_V parameter, the lower the computational effort seems to be. Based only on these results, it would be expected that the larger values of q_V would be the most effective ones. However, as will be discussed in the following, the probability of convergence results do not indicate the same conclusion.

The probabilities of convergence as a function of the number of parameters have been obtained for the explored values of q_V. For a less wide distribution function ($q_V = 1.0$ (Gaussian), 1.5 and 1.7), the obtained probabilities of convergence presented a fast decrease with the number of explored parameters. Such behaviour has made it impossible to investigate the scaling relation for these three q_V values, since the number of structural searches that had reached convergence reduced to zero for a number of parameters equal to four (for q_V values of 1.0 and 1.5) and to six ($q_V = 1.7$). The final results obtained for q_V values of 1.7, 2.0, 2.3, 2.5 and 2.7 are presented as a grey scale contour plot in figure 7. As can be seen there, for q_V equal to 1.7 the probability presents a high value for two parameters (90.0%), but decreases very fast and reaches zero (0%) after only six parameters. The $q_V = 2.0$ (FSA) results present a slow decrease rate with the number of parameters. The results obtained for the $q_V = 2.3$, 2.5 and 2.7 values present some complexity.

The associated probabilities of convergence for the 2.3, 2.5 and 2.7 q_V values do not present themselves as an almost monotonically decreasing curve as in the case of the FSA

(and also $q_V = 1.7$), each presenting a shallow minimum for a number of parameters equal to six, four and six respectively, as can be seen in figure 7. As an example, for $q_V = 2.5$, the probability of convergence presents a value of 100% at two parameters, a minimum of 71% at four parameters, with its value rising to 83% at ten parameters. A monotonic decrease in the probability of convergence with the number of parameters, as in the $q_V = 2.0$ and 1.7 cases, would be, at first, expected for all cases.

The existence of these shallow minima may be attributed to limitations in our statistical evaluation, where a total number of 30 structural searches have been performed. Perhaps, with a more representative statistics, from a larger number of structural searches ($\approx$100), these minima would not be observed. However, this more representative statistics would consist in a very tough task, that would require a considerable computational effort, making this approach not a practical one. Therefore, we believe that the reason for the existence of these minima is the large number of long random steps taken during the search process (due to the wider distribution functions). These long steps would then be responsible for the reduction of the probability of the search process escaping from local minima, in the case of high q_V values.

At this point, in order to get some enlightenment, an investigation has been performed of all the structural search processes that have not been able to achieve convergence to global minimum, for the number of structural parameters corresponding to the probability minima. The results of this analysis indicates that for all the three q_V cases (2.3, 2.5 and 2.7) in the non-converging searches, the GSA algorithm got trapped in a well defined local minimum, not being able to escape. So it can be suggested that the Cauchy–Lorentz distribution, with a reduced number of long random steps (in comparison with larger q_V distributions), seems to be the most effective in escaping from local minima.

Another interesting aspect is that the position of the minimum, related to the number of parameters, varies with the adopted value for the q_V parameter, indicating a dependence on the dimensionality of the explored parameter space. This dependence, in our view, is strongly influenced by two main aspects: the complexity of the R_P hypersurface, and the lower sensitivity of the $I(V)$ curves, and consequently of the R_P, to the second and third layer parameters. However, a more detailed investigation must be performed in order to elucidate this.

In another attempt to understand this complex dependence of the probability of convergence as a function of the number of explored parameters on the q_V parameter value, another investigation was performed. The purpose was to try to understand the effect of changes in the cooling scheme, given by equation (2) for a determined q_V value, on the scaling behaviour and probability of convergence. A better understanding could enable the proposal of new practical search engines that would be able to overcome the problems associated with undesirable behaviours of the probability of convergence with the number of parameters previously discussed. For this investigation, two q_V values have been chosen:

(i) the less wide 1.7 distribution, that presented a fast decrease of the probability of convergence with the number of parameters;
(ii) the wide 2.5 distribution, that presented a minimum in the probability for a number of four parameters.

In the case of q_V equal to 1.7, a $T_{q_V}(t) \approx T_{q_V}(1)/t$ cooling scheme was adopted, instead of the slow $T_{q_V}(t) \approx T_{q_V}(1)/t^{0.7}$ given by equation (2). As previously discussed, $T_{q_V}(1)$ and t correspond to the initial 'temperature' value and the number of cooling steps performed, respectively. This new $q_V = 1.7$ cooling procedure consists in a 'quenching' scheme, a situation in which a fast temperature schedule is used, unable to establish a true ergodic search, i.e., to assure a probability of convergence equal to one (100%). For the $q_V = 2.5$ case, the

$T_{q_V}(t) \approx T_{q_V}(1)/t$ cooling scheme was again adopted, and not the fast $T_{q_V}(t) \approx T_{q_V}(1)/t^{1.5}$ one given by equation (2). In other words, we have tried to impose a faster cooling scheme for the less wide and slow cooling $q_V = 1.7$ distribution and a slower cooling scheme for the wide and fast cooling $q_V = 2.5$ distribution. The FSA ($q_V = 2.0$) cooling scheme ($T_{q_V}(t) \approx T_{q_V}(1)/t$) consists in a relatively small change in both $q_V = 1.7$ and 2.5 original cooling schemes, and was adopted in both cases. The results obtained in the first investigation are presented in figure 8, with a comparison between the $q_V = 1.7$ 'quenching' and SA cases. As can be seen in figure 8(b), the change in the cooling scheme ('quenching') has improved the probability of convergence, that presented a slower decreasing. However, the scaling relation was found to be exponential, as seen in figure 8(a). As previously mentioned, a scaling relation could not be obtained for the $q_V = 1.7$ SA case, due to the null probability of convergence after just six parameters, as can be seen in figure 8(b). For the $q_V = 2.5$ slower cooling scheme investigation case, the probability of convergence obtained presents a slightly shallower minimum, if compared to the SA situation, as shown in figure 8(d). However, the scaling behaviour for the slow cooling scheme presents a exponential scaling relation (figure 8(c)). In both investigated cases, the changes in the cooling scheme have led to a computational effort that increases exponentially with the number of parameters. Within the limitations of the investigation, the results seem to indicate that changes in the cooling scheme for a given q_V cannot result in a more effective search engine.

Therefore, considering the results obtained so far, this study has provided information about the scaling relation and probability of convergence, with the FSA approach ($q_V = 2$) seeming to be the most effective search machine. The FSA scheme presents, at least within the limitations of this study, a very favourable linear scaling (N^1) as well as a slow decreasing probability of convergence.

Another interesting aspect is that it is possible to perform a detailed comparison of the results obtained in this study with the previous work of Nascimento *et al* [15] for the case of $q_V = 2.0$, i.e., the fast simulated annealing search engine. Both studies have employed a theory–theory comparison for the CdTe(110) system. A comparison between the results obtained for the number of trial structures is presented in figure 9(a). As can be seen, the scaling relation is linear (N^1) for both cases. However, the least squares fitting results indicate different slopes for each data set. This can be explained because, although both studies have employed a theory–theory comparison for the CdTe(110) system, this comparison has been performed in different ways. This work has adopted a normal incidence geometry for the theory–theory comparison in order to achieve a better use of symmetry in the multiple-scattering calculations and consequent reduction of the computational effort. In the previous work [15], an off-normal symmetry was adopted in the theory–theory comparison in order to mimic the situation for an experimental data set then collected. For each situation (normal and off-normal incidence) a different hypersurface is associated, for the R_P factor. Each hypersurface presents different topographic features, and different results should be expected for the number of trial structures as a function of the number of explored parameters. So, as can be inferred from this comparison, the slope for the trial structure curve seems to be case dependent, but the scaling relation does not. The probability of convergence as a function of the number of parameters is presented in figure 9(b) for both studies. There are some changes in the behaviour of the probability of convergence, with the curve for the off-normal situation presenting a faster decrease with the number of parameters. These changes can also be explained as a consequence of the different features in the R-factor hypersurfaces for both cases, that will alter the probability of locating the global minimum. So, the probability of convergence also seems to be case dependent. However, for both situations, the FSA algorithm has presented not only the very favourable linear scaling, but also a slow decreasing probability of convergence with the

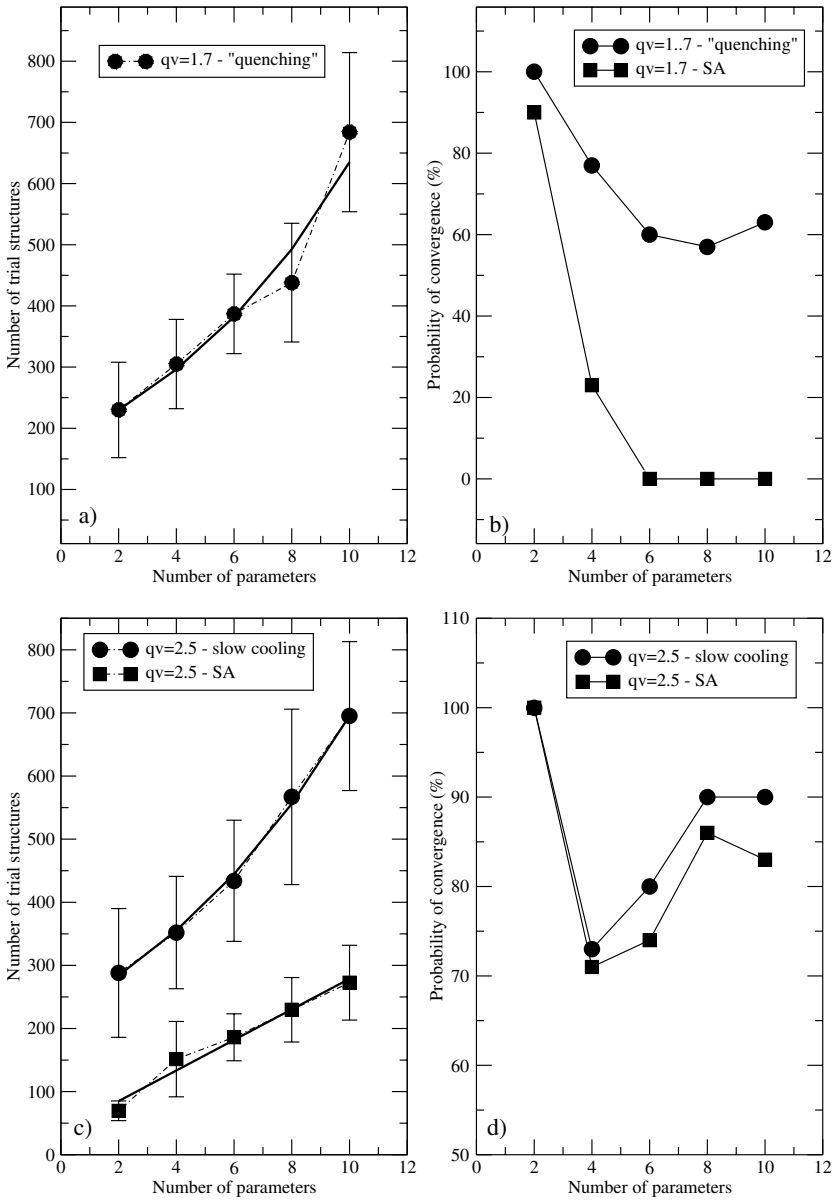

Figure 8. Number of trial structures (a) and (c), and probability of convergence (b) and (d), as a function of the number of parameters, for q_{V} equal to 1.7 and 2.5 respectively. For the $q_{\mathrm{V}} = 1.7$ case, besides the standard simulated annealing scheme ($T_{q_{\mathrm{V}}}(t) \approx T_{q_{\mathrm{V}}}(1)/t^{0.7}$), a faster cooling scheme ($T_{q_{\mathrm{V}}}(t) \approx T_{q_{\mathrm{V}}}(1)/t$) was investigated, in a 'quenching' situation. The well defined curve in (a) corresponds to a least-squares exponential fitting, which presented a correlation equal to 0.983. In the $q_{\mathrm{V}} = 2.5$ case, the standard simulated annealing scheme ($T_{q_{\mathrm{V}}}(t) \approx T_{q_{\mathrm{V}}}(1)/t^{1.5}$) and also a slower cooling scheme ($T_{q_{\mathrm{V}}}(t) \approx T_{q_{\mathrm{V}}}(1)/t$) were investigated. The well defined curves in (c) correspond to exponential (correlation equal to 0.999) and linear least-squares fittings (correlation equal to 0.987), respectively for the slower cooling and the SA scheme data.

number of investigated parameters. These two features indicate the FSA search machine as a very effective one.

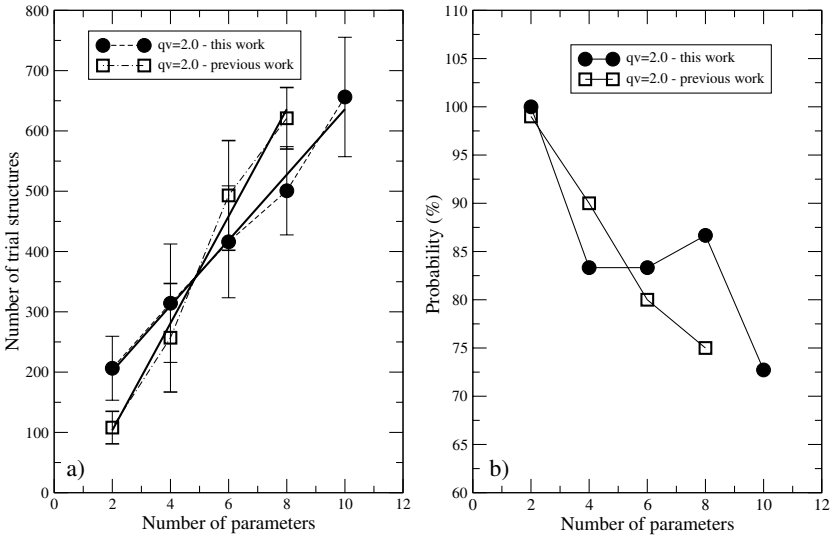

Figure 9. Number of trial structures (a) and probability of convergence (b) as function of the number of varied parameters for the FSA search engine ($q_v = 2.0$) obtained in this study and in the previous work of Nascimento *et al* [15]. The dark straight lines in (a) represent least squares fitting results for the two data sets. Correlations equal to 0.995 and 0.993 have been obtained for the results concerning this and the previous work, respectively.

4. Conclusions

This work has investigated the influence of the visiting distribution function in the behaviour of the generalized simulated annealing (GSA) algorithm when applied to the LEED search problem. Several searches have been performed (for the optimization of two, four, six, eight and ten structural parameters) in an attempt to locate the R_P global minimum in the case of a theory–theory comparison for the CdTe(110) system. The obtained results, concerning the scaling relation and convergence probability, indicate the FSA approach ($q_V = 2.0$) to be the most effective search machine. For higher values of the q_V parameter a complex dependence of the probability of convergence on the dimensionality of the parameters space was observed and a more detailed investigation is necessary.

Two investigations have been performed aiming to examine the effects of changes in the cooling schemes from its defined relation (equation (2)), for the cases of $q_V = 1.7$ and 2.5. A faster ('quenching') and a slower cooling scheme have been employed for the $q_V = 1.7$ and 2.5 cases respectively. It can be inferred, from the results obtained in both cases, that although changes in the cooling scheme can improve the probability of convergence the associated computational effort increases exponentially with the number of parameters investigated.

Acknowledgments

The authors acknowledge the financial support of CNPq, CAPES, FAPESB and FAPEMIG, Brazilian agencies. This work has partial support of FINEP-CTPETRO/FAPEX/UFBA (contract 65.99.0487.00). The authors would like to thank Professors C Tsallis (CBPF, Brazil), K C Mundin (UNB, Brazil) and M A Van Hove (LBL, USA) for helpful discussions. Mrs A C F de Castilho was very helpful in revising the text.

References

[1] Pendry J B 1993 *Surf. Sci. Rep.* **19** 87
[2] Van Hove M A, Weinberg W H and Chan C-M 1986 *Low Energy Electron Diffraction—Experiment, Theory and Surface Structure Determination* (Berlin: Springer)
[3] Van Hove M A and Tong S Y 1979 *Surface Crystallography by LEED* (Berlin: Springer)
[4] Adams D L 2002 *Surf. Sci.* **519** 157
[5] Van Hove M A 1997 *Surf. Rev. Lett.* **4** 479
[6] Cowell P G and de Carvalho V E 1987 *Surf. Sci.* **187** 175
[7] Rous P J, Pendry J B, Saldin D K, Heinz K, Müller K and Bickel N 1986 *Phys. Rev. Lett.* **57** 2951
[8] Van Hove M A, Moritz W, Over H, Rous J P, Wander A, Barbieri A, Materer N, Starke U and Somorjai G A 1993 *Surf. Sci. Rep.* **19** 191
[9] Pendry J B 1974 *Low Energy Electron Diffraction* (London: Academic)
[10] Kleinle G, Moritz W, Adams D L and Ertl G 1989 *Surf. Sci.* **219** L637
[11] Pendry J B and Heinz K 1990 *Surf. Sci.* **230** 137
[12] Kleinle G, Moritz W and Ertl G 1990 *Surf. Sci.* **238** 119
[13] Rous P J 1993 *Surf. Sci.* **296** 358
[14] Nascimento V B, de Carvalho V E, de Castilho C M C, Soares E A, Bittencourt C and Woodruff D P 1999 *Surf. Rev. Lett.* **5** 651
[15] Nascimento V B, de Carvalho V E, de Castilho C M C, Costa B V and Soares E A 2001 *Surf. Sci.* **487** 15
[16] Szu H and Hartley R 1987 *Phys. Lett.* A **122** 157
[17] Nascimento V B, Soares E A, de Carvalho V E, Lopes E L, Paniago R and de Castilho C M C 2003 *Phys. Rev.* B **68** 245408
[18] Kottcke M and Heinz K 1997 *Surf. Sci.* **376** 352
[19] Döll R and Van Hove M A 1996 *Surf. Sci.* **355** L393
[20] Goldberg D E 1989 *Genetic Algorithms in Search, Optimization and Machine Learning* (Reading, MA: Addison-Wesley)
[21] Tsallis C and Stariolo D A 1987 *Physica* A **233** 395
[22] Kirkpatrick S, Gellat C D and Vecchi M P 1983 *Science* **220** 671
[23] Geman S and Geman D 1984 *IEEE Trans. Pattern Anal. Mach. Intell.* **6** 721
[24] Tsallis C 1988 *J. Stat. Phys.* **52** 479
[25] Curado E M F and Tsallis C 1991 *J. Phys. A: Math. Gen.* **24** L69
Tsallis C 1995 *Phys. Lett.* A **206** 389
[26] Pendry J B, Heinz K and Oed K 1990 *Phys. Rev. Lett.* **61** 2953
[27] Rous P J, Van Hove M A and Somorjai G A 1990 *Surf. Sci.* **226** 15
[28] Soares E A, de Carvalho V E and Nascimento V B 1999 *Surf. Sci.* **431** 74
[29] Pendry J B 1980 *J. Phys. C: Solid State Phys.* **13** 937