Optimum scaling for structural optimization in low-energy electron diffraction

N. C. Kothari, H. Over,* and D. K. Saldin
Department of Physics and Laboratory for Surface Studies, University of Wisconsin-Milwaukee, P.O. Box 413, Milwaukee, Wisconsin 53201
(Received 9 August 1993; revised manuscript received 13 December 1993)

Although the number of dynamical low-energy-electron-diffraction (LEED) calculations that need to be performed in the linear LEED technique scales linearly with the number N of structural parameters, the number of combinations of the calculated intensities that need to be compared with experiment scales exponentially with N, and could be a serious limitation when N is large. In this paper, we suggest an algorithm for overcoming this residual exponential scaling, thus allowing linear LEED to achieve its true potential for linear scaling. We demonstrate our algorithm with model calculations on the Si(111)$(\sqrt{3}\times\sqrt{3})R30°$-Al surface.

Multiple-scattering calculations in low-energy electron diffraction (LEED) (Refs. 1 and 2) are capable of reproducing experimental intensity vs energy (I-E) profiles for all measured beams quite faithfully, once the true crystal surface structure is known. The problem still lies in devising an efficient search algorithm to find the optimum structure among a range of varied structural parameters. All current methods in LEED use some form of objective criterion, known as a reliability (or R) factor[2] for comparing a set of measured I-E curves with ones calculated from various model structures. In an exhaustive search technique, as practiced originally in LEED, the R factors resulting from all combinations of a few selected parameters were varied independently, and the structure that gives the best fit (or minimum R factor) was taken to be the solution of the structural problem. The sheer number of different model calculations that need to be performed in an exhaustive search of, say, N structural parameters grows exponentially with N. Thus, even a few orders of magnitude of improvement in the speed of computer hardware and software can make little impact on the complexity of the surface structures that may be solved. As Pendry, Heinz, and Oed[3] have pointed out, a problem of this sort is known to be of NP- (or nonpolynomial-) complete type, and represents a particularly intractable one in optimization theory.

In the last few years, considerable effort has been devoted towards finding methods of circumventing the NP-complete problem. The simulated annealing technique recently studied by Rous[4] appears to scale as N^6. Another class of techniques employs directed search schemes based on information about the local topography of the R-factor hypersurface to find the location of a minimum.[5] Such methods begin with a given set of values of the N parameters as a starting configuration. Any particular directed search is found to have the much more desirable computer-time scaling proportional to N^{2}.[4,5] However, the general existence of multiple minima of the R factor results in different starting configurations leading to different local minima. Therefore, the searches need to be repeated with different starting configurations, separated by twice the convergence radius $\mathcal{R}$ (in the N-dimensional parameter space) of the optimization scheme [typically about 0.2 Å (Ref. 4)]. Thus, the number of starting configurations also scales exponentially with N, and, in effect, the problem remains of NP-complete type. However, if some rapid method can be found to locate the *region* of the global minimum, to within a radius $\mathcal{R}$, that configuration could be used as the starting point for a single directed search, guaranteed to find the exact global minimum, still scaling as N^2, but much more rapid due to the much smaller parameter space left to be explored.

Much of the recent interest in holographic methods in electron diffraction[6] stem from the desire to find such a method. In this paper, we report on an alternative, but somewhat related, technique for finding such a starting configuration. It combines the ideas of linear LEED (LLEED), recently developed by Wander, Pendry, and Van Hove,[7] with the linear programming based iterative optimization scheme APRIORI recently proposed by Saldin *et al.*[9] for photoelectron diffraction.

We first review the basic ideas of linear LEED.[7,8] LLEED requires a fully dynamical calculation to be performed for a reference structure as a starting point. Suppose that the amplitude of the gth Bragg beam from this reference structure is $A_g^{(0)}(E)$ for an electron of energy E (this is analogous to the reference wave in holography). Now imagine a true structure that differs from the reference structure by displacements $\delta\mathbf{r}_i$ of the structural parameters i, which make nearly *independent* contributions to the resulting diffracted beam amplitudes $A_g(E)$. Under such conditions, the following approximate relations must hold:

$$A_g(E)\simeq A_g^{(0)}(E)+\sum_i \delta A_{g,i}(E,\delta\mathbf{r}_i)\ , \tag{1}$$

where

$$\delta A_{g,i}(E,\delta\mathbf{r}_i)=A_{g,i}(E,\delta\mathbf{r}_i)-A_g^{(0)}(E) \tag{2}$$

is the change in the amplitude of the gth beam at energy E due to the displacements $\delta\mathbf{r}_i$. (The amplitudes $\delta A_{g,i}$ may be thought of as analogous to object waves in holography.) The condition of *independence* restricts somewhat

the sets of displacements $\delta\mathbf{r}_i$ which may be treated by LLEED. For example, displacements of different Cartesian coordinates of a single atom do not make independent contributions to the LEED amplitudes. Strong multiple scattering between two displaced atoms may also diminish the reliability of this approximation.

The independence of the contributions of such structural displacements has great implications for the computer-time scaling of LLEED. If each of N such structural parameters are discretely varied over say M values, the number of independent full dynamical calculations that need to be performed would be MN, implying an optimal linear scaling with N. However, even after the MN dynamical calculations are performed, the amplitudes would need to be combined into M^N sets of I-E curves, still an exponential scaling with N, and so the nightmare of the NP-complete problem would not have been totally exorcised. In fact, we have found that the computing requirements for forming and comparing the M^N I-E curves with experiment can still be a daunting task in practice, and could form a bottleneck in the LLEED process. In this paper, we demonstrate how this difficulty may be overcome.

Suppose that in our exploration of the displacement parameter space we consider discrete displacements $\delta\mathbf{r}_{i,k}$ of the ith parameter (where k is the displacement index). Following the example of the APRIORI scheme,[9] we let $p(\delta\mathbf{r}_{i,k})$ represent the weight of the contribution of the displacement $\delta\mathbf{r}_{i,k}$ to the changes in the LEED amplitudes due to that displacement. In the LLEED regime then, the LEED amplitudes due to the simultaneous displacements of all N structural parameters may be approximated by

$$A_g(E)\simeq A_g^{(0)}(E)+\sum_{i,k}p(\delta\mathbf{r}_{i,k})\delta A_{g,i}(E,\delta\mathbf{r}_{i,k})\ . \tag{3}$$

Our aim will be to efficiently determine the distributions $p(\delta\mathbf{r}_{i,k})$, which optimize the fit of our calculated LEED intensities to measured ones, assuming that the MN dynamical calculations have been performed to find all the calculated LEED amplitudes on the right-hand side of (3).

First, it follows from (3), that our theoretical estimate of the measured intensities may be written

$$I_g^{\rm th}(E)=|A_g(E)|^2=|A_g^{(0)}(E)|^2+\sum_{i,k}p(\delta\mathbf{r}_{i,k})\left[\{A_g^{(0)*}(E)\delta A_{g,i}(E,\delta\mathbf{r}_{i,k})+\text{c.c.}\}+\delta A_{g,i}^{*}(E,\delta\mathbf{r}_{i,k})\sum_{j,l}p(\delta\mathbf{r}_{j,l})\delta A_{g,j}(E,\delta\mathbf{r}_{j,l})\right]\ , \tag{4}$$

where j and l are extra dummy indices which independently vary over the same range as i and k, respectively. We wish to find the distributions $p(\delta\mathbf{r}_{i,k})$ which optimize the fit of $I^{\rm th}$ to $I^{\rm exp}$. Due to the nonlinear dependence of $I^{\rm th}$ on the sought distribution $p(\delta\mathbf{r}_{i,k})$ above, this does not seem possible in one step. We thus resort to the process of iterative optimization, recently proposed for photoelectron diffraction.[9] On this scheme we start with some initial distribution $p^{(0)}(\delta\mathbf{r}_{i,k})$ and find better estimates of this distribution by the following iterative procedure. We define the error in our fit of theoretical and experimental intensities at the nth iteration by

$$\epsilon_{g,E}^{(n)}=\left[|A_g^{(0)}(E)|^2+\sum_{i,k}p^{(n)}(\delta\mathbf{r}_{i,k})M_{g,i}^{(n)}(E,\delta\mathbf{r}_{i,k})-\mu^{(n)}I_g^{\rm exp}(E)\right]\ , \tag{5}$$

where

$$M_{g,i}^{(n)}(E,\delta\mathbf{r}_{i,k})=[A_g^{(0)*}(E)\delta A_{g,i}(E,\delta\mathbf{r}_{i,k})+\text{c.c.}]+\delta A_{g,i}^{*}(E,\delta\mathbf{r}_{i,k})\sum_{j,l}p^{(n-1)}(\delta\mathbf{r}_{j,l})\times\delta A_{g,j}(E,\delta\mathbf{r}_{j,l})\ . \tag{6}$$

Note that if the distributions $p^{(n-1)}(\delta\mathbf{r}_{j,l})$ at the previous iteration are given, the quantities $M_{g,i}^{(n)}(E,\delta\mathbf{r}_{i,k})$ are known, and the estimates $p^{(n)}(\delta\mathbf{r}_{i,k})$ of the same distribution at the current iteration, and of the normalization constants $\mu^{(n)}$ may be obtained by a linear programming scheme. As in the work on photoelectron diffraction,[9] we employ the simplex method.

The procedure is to minimize the "objective function" consisting of the sum over all beams g and all energies E, of the absolute errors in our fit between theory and experiment, defined by

$$\sum_{g,E}|\epsilon_{g,E}^{(n)}| \tag{7}$$

subject to the constraints

$$p^{(n)}(\delta\mathbf{r}_{i,k})\geq 0\ \text{ for all } i \text{ and } k \tag{8}$$

and

$$\sum_{k}p^{(n)}(\delta\mathbf{r}_{i,k})=1\ \text{ for } \textit{all}\ i\ . \tag{9}$$

In analogy with our earlier work, we take

$$p^{(0)}(\delta\mathbf{r}_{j,l})=\delta_{i,j}\delta_{k,l} \tag{10}$$

to start the iterations, although we found that after convergence, the final distributions were quite insensitive to the precise choice of a starting distribution. In the work reported below, satisfactory convergence was found after about 20 iterations. By this process of iterative optimization, a linear programming technique was able to solve a nonlinear problem.

As a test of our proposed scheme, we first performed a full dynamical calculation, using the computer program of Moritz,[10] of the LEED intensities from the Si(111)

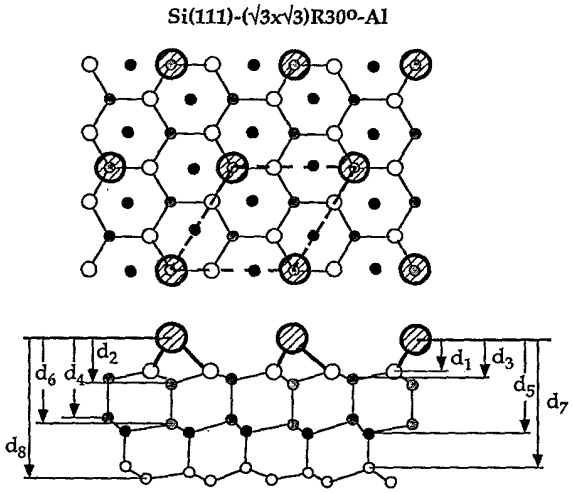

FIG. 1. Normal and side views of the Si(111)$(\sqrt{3}\times\sqrt{3})R30°$-Al surface. The latter shows the layer displacements d_1 to d_8, which were varied in our calculations.

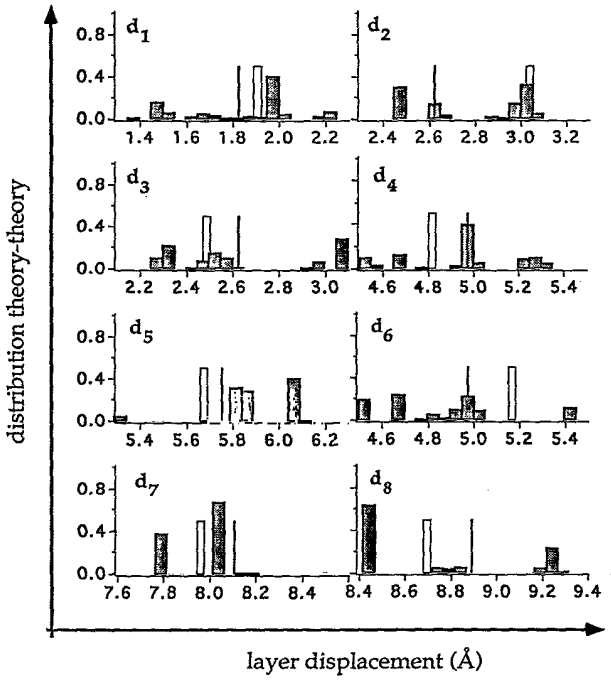

FIG. 2. The distributions $p(\delta r_{i,k})$ for layers $i=1,\ldots,8$, as calculated by our procedure from the smaller set of 197 data points. The ranges over which the layer displacements d_i were varied are shown on the abscissas of the separate panels. The vertical dark lines on each panel indicate the values of d_i in the reference structure, and the white columns their values in the "true" structure. The correlation of the maxima of the histogram values (gray columns) with the white columns is not too convincing, except perhaps for the first two layers.

$(\sqrt{3}\times\sqrt{3})R30°$-Al surface (see Fig. 1). For this calculation we assumed the structure recently proposed by Huang *et al.*,[11] in which the Al atoms are assumed to be adsorbed on the $T4$ sites and to induce far-reaching relaxations in the first few Si substrate layers of up to 0.4 Å. For the purpose of our test, we may regard this as being the "true" structure, and we may identify the calculated LEED intensities with $I_g^{\exp}(E)$ in our notation above, i.e., we imagine them to be a set of "experimental" intensities measured in a LEED experiment on the true surface structure. We will attempt to determine the true structure by the combination of LLEED and our optimization scheme above, assuming a knowledge of just the experimental intensities.

The first step is to choose a reference structure. Given a lack of any knowledge of the surface relaxations, a convenient reference structure would be the bulk truncated one for the Si substrate, with a height of the Al atoms above the top substrate layer determined by the conventional atomic radii of Al and Si. This implies layer displacements of the first eight substrate Si layers from the adsorbate layer of $d_i=1.82$, $d_2=d_3=2.62$, $d_4=d_6=4.97$, $d_5=5.76$, $d_7=8.11$, and $d_8=8.89$, all in Å units. The LEED amplitudes $A_g^{(0)}(E)$ from this reference structure were also calculated by Moritz's[10] computer program. The values of the layer displacements for the true structure were up to 0.4 Å different from those for the reference structure, and thus unreachable in practice by optimization schemes based on the information about the local topography of the R-factor hypersurface.[5] The amplitude changes $\delta A_{g,i}(E,\delta r_{i,k})$ corresponding to *independent* changes of these layer displacements (assuming all other layer displacements fixed at that for the reference structure) were then calculated, and the optimization scheme described above applied to determine the final self-consistent distributions $p^{(\infty)}(\delta r_{i,k})$.

The results of the calculations for an experimental data set consisting of 197 LEED intensities (with an energy interval between successive intensities of 5 eV) are shown on Fig. 2. The histograms (dark columns) represent the values of $p^{(\infty)}(\delta r_{i,k})$ as functions of the depths of the first eight Si layers from the adsorbate. The white columns represent the layer positions in the true structure, and the dark vertical lines those of the reference structure, which represents the starting point in the search. The maximum histogram values are indeed found very close to the true positions of the first two layer displacements d_1 and d_2, which of course affect the measured intensities to the greatest extent. In the case of d_2, even the rather large displacement of $\simeq 0.4$ Å from the reference structure is found very accurately, but the true displacements of the deeper layers are not located convincingly.

However if our scheme is repeated for a larger experimental data set, consisting of 346 LEED intensities, also with the same 5-eV energy intervals, the results improve quite dramatically, as seen on Fig. 3. Here the maxima of the histogram values are found within $\simeq 0.05$ Å of the true structure for *all eight* of the layer displacements. Alternatively, treating the weights $p(\delta r_{i,k})$ as a statistical distribution, the mean values of d_k determined for each of the eight layers were found to be within a standard deviation of the true values. Even taking the average of the standard deviations of the eight distributions as an error estimate, we get a value of 0.14 Å. This is an impressive

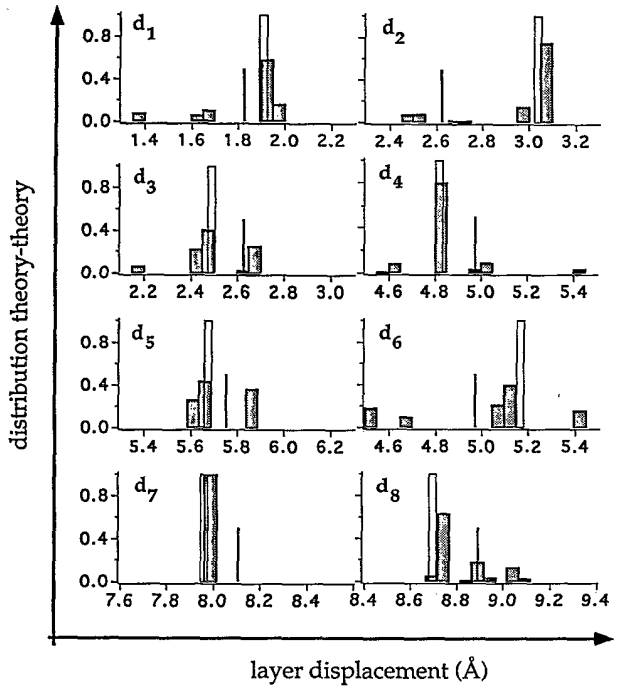

FIG. 3. Same as Fig. 2, except that a larger set of 346 data points were used. Here the correlation of the histogram maxima and the "true" layer displacements (white columns) is almost perfect.

result when compared with the fact that the directed search method of Kleinle, Moritz, and Ertl[12] and Over *et al.*[13] was quite unable to get even close to the true structure, starting from the same reference structure, since it became entrapped in a nearby local minimum.

We make a few comments regarding the computer-time scaling of our optimization scheme. At the first step of our procedure, MN dynamical LEED calculations are performed, as in the usual LLEED process. It is anticipated that most of the computing time will be spent at this stage, which has the very desirable linear scaling with N. At the second stage of usual LLEED procedure, these MN sets of LEED amplitudes are combined in M^N ways, unfortunately still an NP-complete process. In contrast, linear programming practice indicates a time scaling of each optimization step of our scheme proportional to $(MN+2D)(D+N)^2$, where D is the number of experimental data points. Thus, when as is usual, $D \gg MN$, the scaling with the number of parameters has the extraordinary feature of being almost *independent* of N. [Even in the worst (and unrealistic) case of $N \gg D$ the scaling is no worse than N^3.] Thus, in practice, our scheme allows LLEED to realize its true potential as a process linearly proportional to N, with a fixed computer-time overhead for the final search for the optimum structure, almost *independent* of the number of structural parameters, and relatively weakly dependent ($\sim D^3$) on the number of LEED amplitudes measured.

On comparing Figs. 2 and 3, we notice an important feature of our scheme, and that is that there appears to be a minimum to the size of an experimental data set for the reliable reconstruction of the true values of a given number of varied parameters. Calculations like the ones reported here could thus serve to suggest the number of data points that would need to be measured in a real experiment. If further accuracy were required than the $\simeq 0.05$–0.14 Å estimated above, a conventional directed search scheme may be further applied for a final structural refinement. That refinement would be expected to scale as N^2, but, due to the much smaller volume of parameter space left to be explored, would be expected to be both reliable and very rapid.

A powerful by-product of our optimization scheme is that it does not force a particular structural parameter to have just a single value. In fact, at any finite temperature, atoms vibrate, and a diffraction pattern reflects the thermal distributions of the vibrating atoms. In common with the direct methods proposed by Pendry and Heinz,[3,14] our algorithm should be capable of recovering this distribution from experimental data. Other applications could include the determination of the distributions of different surface domains contributing to a diffraction pattern, as frequently encountered with LEED data taken with nonnormal electron incidence.

We thank Minnie Patel and Xiang Chen for very helpful discussions. N.C.K. acknowledges financial support from the Laboratory for Surface Studies of the University of Wisconsin-Milwaukee (UWM), and H.O. acknowledges a grant from the "Deutsche Forschungsgemeinschaft." D.K.S. thanks John Pendry for drawing his attention to the linear LEED technique, and the Donors of the Petroleum Research Fund, administered by the American Chemical Society, for support of this research.

*On leave from the Fritz Haber Institut der MPG, Faradayweg 4-6, D-14195 Berlin Dahlem, Germany.

[1]J. B. Pendry, *Low Energy Electron Diffraction* (Academic, London, 1974).

[2]M. A. Van Hove, W. H. Weinberg, and C. T. Chan, *Low Energy Electron Diffraction* (Springer, Berlin, 1986).

[3]J. B. Pendry, K. Heinz, and W. Oed, Phys. Rev. Lett. **61**, 2953 (1988).

[4]P. J. Rous, Surf. Sci. **296**, 358 (1993).

[5]M. A. Van Hove, W. Moritz, H. Over, P. J. Rous, A. Wander, A. Barbieri, N. Materer, U. Starke, and G. A. Somorjai, Surf. Sci. Rep. **19**, 191 (1993).

[6]See, e.g., A. Szöke, Phys. Rev. B **47**, 14 044 (1993).

[7]A. Wander, J. B. Pendry, and M. A. Van Hove, Phys. Rev. B **46**, 9897 (1992).

[8]J. B. Pendry, Surf. Sci. Rep. **19**, 87 (1993).

[9]D. K. Saldin, X. Chen, N. C. Kothari, and M. H. Patel, Phys. Rev. Lett. **70**, 1112 (1993).

[10]W. Moritz, J. Phys. C **17**, 353 (1984).
[11]H. Huang, S. Y. Tong, W. S. Yang, H. D. Shih, and F. Jona, Phys. Rev. B **42**, 7483 (1990).
[12]G. Kleinle, W. Moritz, and G. Ertl, Surf. Sci. **238**, 119 (1990).
[13]H. Over, U. Ketterl, W. Moritz, and G. Ertl, Phys. Rev. B **46**, 15 438 (1992).
[14]J. B. Pendry and K. Heinz, Surf. Sci. **230**, 137 (1990).