Perturbation theory of diffuse RHEED applied to rough surfaces: Comparison with supercell calculations

U. Korte,* J. M. McCoy,† and P. A. Maksym
Department of Physics and Astronomy, University of Leicester, University Road, Leicester LE1 7RH, United Kingdom

G. Meyer-Ehmsen
Fachbereich Physik, Universität Osnabrück, D-49069 Osnabrück, Germany
(Received 1 November 1995; revised manuscript received 2 April 1996)

We present a thorough investigation of the conditions under which the perturbation theory of diffuse RHEED [Phys. Rev. B **48**, 8345 (1993)] (diffuse scattering treated as the transition between states in the periodic part of the potential) can be used for the evaluation of experimental diffuse scattering data from occupational disorder (e.g., steps). Such an investigation is desirable because this method solves the configuration problem, i.e., the presence of many statistically varying disorder configurations, within the scope of a dynamical theory. We have carried out comparisons with rigorous supercell calculations for streak profiles due to bilayer steps upon a Si(100) surface. The general trend of the results obtained is supported by some simple analytical considerations. An analytical expression is derived that predicts the approximate error made in the perturbation approach compared to an exact treatment. Besides general features, such as defect concentration and the strength of the atomic potential, the important structural quantity that determines the quality of the perturbation approach is the correlation length of the disorder along the incident beam azimuth. If this length is sufficiently small, perturbation theory works well and is independent of the diffraction condition. Otherwise, the applicability of perturbation theory depends on the diffraction condition. The physics behind conditions suitable for perturbation theory to work, as well as their experimental realization, is discussed. [S0163-1829(96)05427-6]

I. INTRODUCTION

The calculation of elastic reflection high-energy electron diffraction (RHEED) from ordered surfaces can currently be carried out routinely and accurately within the framework of dynamical diffraction theories, i.e., theories where multiple scattering is fully included. The application of RHEED theory, used in conjunction with proper fitting procedures, enables crystallographic data for various surfaces to be obtained by analysis of experimentally recorded rocking curves (e.g., Refs. 1–4). In spite of this progress, the question of how to treat RHEED from surfaces that are not monatomically flat has still not completely been answered, although the major application of RHEED concerns the control of epitaxial growth. Here, the surface is always in a more or less rough state such that the lateral translation symmetry of the flat surface is broken, giving rise to diffuse scattering. A further interesting aspect of diffuse scattering in RHEED concerns its high sensitivity with regard to surface disorder. Due to the grazing incidence scattering geometry at high electron energy small lateral momentum transfer along the incident beam azimuth is connected with a large momentum transfer normal to the surface (the latter being typically one order of magnitude larger than the former one). For this reason the resolving power of RHEED with regard to the distance over which deviations from the perfect periodicity can be detected is considerable (some 1000 Å) and usually higher than in the case of the complementary technique, low-energy electron diffraction (LEED).

Besides treatments based on the kinematical theory,[5] there are at present three theoretical approaches that treat diffuse scattering in RHEED within the scope of a dynamical diffraction theory. One is adopted from the multislice method for transmission electron diffraction and has been used to calculate the scattering from a single step.[6] In the second method the disorder is simulated within a large surface unit cell (supercell), which is repeated periodically.[7] In this way the calculation can be carried out using established computation techniques that rely on periodicity parallel to the surface.[8,9] This supercell approach has the advantage that the diffraction from the configuration created in this manner is treated in principle exactly. A theory for the simulation of reflection electron microscopy (REM) images that uses the supercell technique has been developed recently.[7]

In the case of both the supercell and the multislice approaches the full dynamical calculation has to be carried out for each new configuration considered. On the other hand, for an analysis of experimental data from real surfaces, calculations for a large number of statistically varying configurations would be necessary. This configuration problem, in conjunction with the large computational effort required for even a single structure, presents a problem for the interpretation of diffuse scattering data from real surfaces.

Recently, a third approach and a corresponding computer program based on perturbation theory have been introduced.[10] Here, the dynamical diffuse scattering is regarded as a transition between states (two-dimensional Bloch waves) in the periodic part of the scattering potential. The nonperiodic part is treated as a perturbation. This formalism solves the configuration problem because the approach can be reduced to the use of conventional structure factors multiplied by ''dynamical'' form factors calculated with the per-

turbation theory. With this theory of diffuse RHEED, the problem of a surface with a single adsorbed atom could be solved. Also experimentally observed thermal diffuse scattering distributions[10,11] (Kikuchi scattering) and streak profiles corresponding to antiphase domains[2,12] could be reproduced rather well.

It was not evident until now whether such a formalism is applicable to the general case of occupational disorder, for example, surface steps and related defects. For instance, it can be shown that for a stepped monolayer of half coverage the sum of all the diffuse Fourier coefficients of the potential is of the same order of magnitude as the coefficients due to the periodic part of the potential (see Sec. II). But at present the perturbation scheme is the only approach within the scope of a dynamical scattering theory that can handle the configuration problem and is, thus, potentially applicable to the interpretation of experimental data in order to obtain surface parameters characterizing the disorder such as typical terrace widths and island sizes. It has also been pointed out that it is probably only with such an approach that formal averaging over many configurations will be achievable.[13] It is therefore important to determine the conditions under which perturbation theory can be used. This is the purpose of the present work.

To that end we will carry out a comparison of diffuse scattering data from the perturbation approach and the supercell technique. Because the latter treats the scattering problem (for a given structure) rigorously, the corresponding computations serve very well as reliable reference calculations. The model system we have chosen is the Si(100) surface with bilayer steps having their edges along [011]. As the comparisons are mainly done for a special type of disorder configuration, a simple analytical treatment of the problem is given as well in order to enable a proper interpretation of the comparison and to transfer the results to realistic surfaces and other materials.

II. SOME INTRODUCTORY CONSIDERATIONS

The structure factor associated with the diffuse scattering from a layer with occupational disorder is defined as

$$S(\mathbf{s})=\sum_n \exp(-i\mathbf{s}\cdot\mathbf{r}_{n\|}), \quad \mathbf{s}\neq\mathbf{g}, \tag{1}$$

where the $\mathbf{r}_{n\|}$ denote the atomic positions and $\mathbf{s}$ denotes the two-dimensional reciprocal vector of the nonperiodic part of the potential; i.e., $\mathbf{s}$ does not coincide with a reciprocal lattice vector $\mathbf{g}$ of the periodic surface lattice. The vectors $\mathbf{s}$ are distributed continuously in k space and are responsible for the diffuse scattering.

The squared modulus of the structure factor is related to the Fourier transform of the pair correlation function $P(\mathbf{r}_\|)$ with respect to the surface coordinate $\mathbf{r}_\|$. $P(\mathbf{r}_\|)$ is the probability of finding two atoms separated by the real-space vector $\mathbf{r}_\|$. For occupational disorder, the pair correlation function of a disordered layer with coverage θ has the general features sketched in Fig. 1. $P(0)$ is the probability of finding an atom in the disordered layer, therefore, $P(0)=\theta$. For $r_\|>0$ the pair correlation function decreases with increasing $r_\|$. The limiting value for $r_\|\to\infty$ is $P(\infty)=\theta^2$. $P(\infty)$ denotes the correlation over infinitely large distances and thus characterizes the long-range order. The full width at half maximum of the decrease with $r_\|$ along a particular azimuth ϕ can be identified with the typical correlation length of the short-range order L_ϕ along that direction.

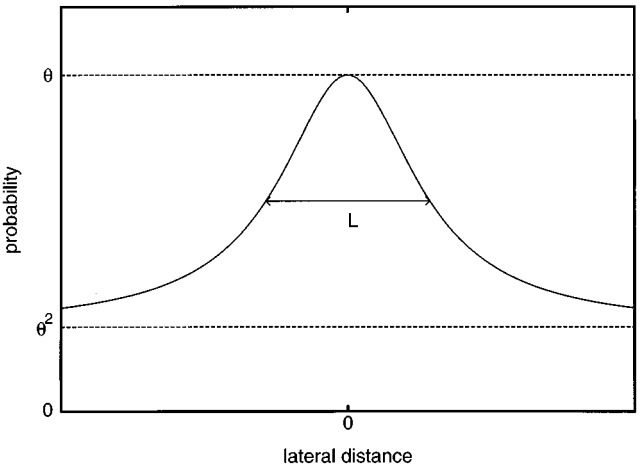

FIG. 1. Typical form of the pair correlation function of a disordered layer in the case of occupational disorder. θ is the coverage and L the typical correlation length involved in the disorder.

Fourier transformation of $P(\mathbf{r}_\|)$ shows that the part P_{diff} of the pair correlation function that is responsible for the diffuse scattering is given by subtraction of the long-range order contribution from the total pair correlation function:

$$P_{\text{diff}}(\mathbf{r}_\|)=P(\mathbf{r}_\|)-P(\infty). \tag{2}$$

$P_{\text{diff}}(\mathbf{r}_\|)$ is proportional to the Fourier transform of $|S(\mathbf{s})|^2$ and it follows that $S(\mathbf{s})$ obeys the sum rule

$$\frac{\Omega}{N4\pi^2}\int_{\text{BZ}} d\mathbf{s}|S(\mathbf{s})|^2=P_{\text{diff}}(0)=\theta(1-\theta), \tag{3}$$

where the integral is over the (two-dimensional) Brillouin zone (BZ), N is the number of surface lattice sites of the system, and Ω is the area of the surface unit cell (see also Ref. 14). This result can also be obtained by direct integration of $|S(\mathbf{s})|^2$ as defined by Eq. (1). The integral in Eq. (3) depends on the coverage only and suggests that, even if the individual diffuse Fourier coefficients are small, the contribution of a whole Brillouin zone is not necessarily negligible. For instance, if $\theta=0.5$, expression (3) reaches its maximum and has the same value as that which would be produced by the Fourier coefficients of a perfectly ordered layer of half scattering strength. This is not really small and it is therefore at first glance not obvious that the diffuse scattering can be treated as a perturbation at all.

In Fig. 2 we give a preview of some results of the numerical comparison (see Sec. VI; the details of the model used will be given in Sec. IV). This demonstrates that the problem is somewhat complicated and the situation is generally not as pessimistic as suggested by Eq. (3). The figure compares diffuse streak profiles for three different situations, calculated with the supercell and with the perturbation approach, respectively. In all cases the supercell consisted of 400 lattice units. The stepped disordered layers were assumed to be of

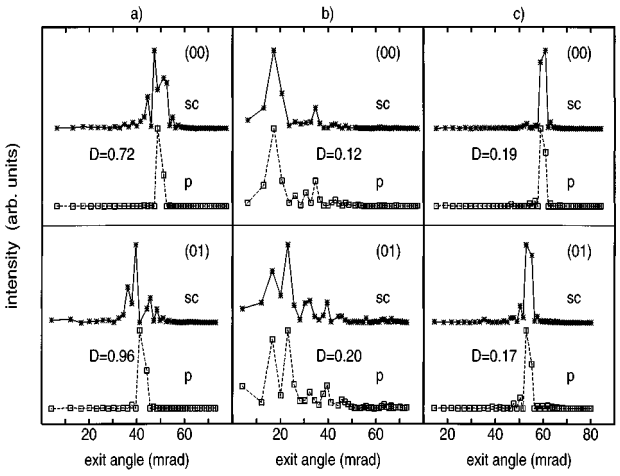

FIG. 2. Streak profiles normal to the shadow edge through the zeroth Laue zone reflections (00) and (01) from rigorous supercell (sc) calculations and perturbation theory (p) for Si(100) bilayer steps and coverage 0.5. The electron energy is 15 keV and the incident azimuth $[0\bar{1}1]$ is perpendicular to the step edges. (a) Regular step array with terrace length of 200 lattice units, incident angle 50 mrad. The curves sc and p do not agree. (b) Random distribution of step units and 50-mrad incident angle. sc and p agree quite well. (c) As in (a) but for 60-mrad incident angle. sc and p agree quite well.

half coverage (which is the most unfavorable case with regard to the coverage). The incident azimuth is perpendicular to the step edges and the lateral diffuse Fourier components are along this direction. The corresponding diffuse scattering profiles are plots of the diffusely reflected intensity versus the polar exit angle. In Fig. 2(a) a regular step array with a terrace length of 200 lattice units and an incident angle of 50 mrad was used. The agreement between supercell and perturbation calculation is indeed bad. The metric distance [the measure used for the comparison, between two curves, defined in Eq. (11)] between perturbation and supercell profiles, averaged over the (00) and the (01) streak, is $D=0.84$, which means virtually no correlation between both profiles. Figure 2(b) shows the comparison for exactly the same diffraction condition and the same coverage. However, this time the step units (corresponding to steps with terrace lengths of one lattice unit) were randomly distributed within the supercell. In this case, the perturbation approach reproduces the rigorous calculation very well. The corresponding metric distance confirms excellent agreement, $D=0.16$. Finally, Fig. 2(c) compares profiles again from the regular step array as used in (a), but at a slightly higher incident angle (60 mrad). Again, the profile obtained from the perturbation calculation represents the exact one quite well ($D=0.17$). These results show that the applicability of the perturbation approach, besides depending on the defect concentration as suggested by Eq. (3), also strongly depends on the statistics of the disorder and the diffraction condition.

III. PERTURBATION AND SUPERCELL APPROACH

Our investigation is essentially based on a comparison of numerical data from the perturbation and the supercell method. In what follows, we will therefore describe briefly the basic elements and assumptions employed in the perturbation and the supercell programs. In order to elucidate the differences between the two approaches, a common formulation of the scattering problem is used. We note, however, that the actual numerical solution is obtained in quite different ways. For these details we refer to the corresponding references, Ref. 10 (perturbation) and Refs. 7 and 15 (supercell).

We start with the usual two-dimensional Fourier expansion of the scattering potential $V(\mathbf{r})$ and the wave function $\psi(\mathbf{r})$.[8] The potential and the wave function are then split up into their periodic (p) and diffuse (d) parts. For the potential we have

$$V(\mathbf{r})=V_p(\mathbf{r})+V_d(\mathbf{r}),$$

$$V_p(\mathbf{r})=(\hbar^2/2m)\sum_{\mathbf{g}} V_p(\mathbf{g},z)\exp(i\mathbf{g}\cdot\mathbf{r}_\parallel),$$

$$V_d(\mathbf{r})=(\hbar^2/2m)\int d\mathbf{s}\, V_d(\mathbf{s},z)\exp(i\mathbf{s}\cdot\mathbf{r}_\parallel). \tag{4}$$

Here, $\mathbf{r}_\parallel$ is the two-dimensional (parallel to the surface) component of the real-space vector $\mathbf{r}$ and z is the coordinate normal to the surface. The discrete set $\{\mathbf{g}\}$ denotes the reciprocal surface lattice vectors corresponding to the periodic part of the potential of the surface. $\mathbf{s}$ stands for the two-dimensional reciprocal vectors of the nonperiodic potential part, which are distributed continuously in k space and are responsible for the diffuse scattering.

The diffuse Fourier coefficients $V_d(\mathbf{s},z)$ can be written as

$$V_d(\mathbf{s},z)=(1/2\pi)^2 v(\mathbf{s},z)S(\mathbf{s}), \tag{5}$$

where

$$v(\mathbf{s},z)=2m/\hbar^2\int d\mathbf{r}_\parallel V_{\text{atom}}(\mathbf{r})\exp(-i\mathbf{s}\cdot\mathbf{r}_\parallel) \tag{6}$$

is the (two-dimensional) Fourier transform of the atomic potential V_{atom} and $S(\mathbf{s})$ is the structure factor (1) associated with the diffuse potential part of the disordered layer.

Analogously, the decomposition of the wave function yields

$$\psi(\mathbf{r})=\psi_p(\mathbf{r})+\psi_d(\mathbf{r}),$$

$$\psi_p(\mathbf{r})=\exp(i\mathbf{K}_\parallel\cdot\mathbf{r}_\parallel)\sum_{\mathbf{g}} w(\mathbf{g},z)\exp(i\mathbf{g}\cdot\mathbf{r}_\parallel),$$

$$\psi_d(\mathbf{r})=\exp(i\mathbf{K}_\parallel\cdot\mathbf{r}_\parallel)\int d\mathbf{s}\, y(\mathbf{s},z)\exp(i\mathbf{s}\cdot\mathbf{r}_\parallel), \tag{7}$$

where $\mathbf{K}_\parallel$ is the parallel component of the incident wave vector. ψ_p consists of the strong waves scattered into directions corresponding to the reciprocal surface lattice vectors $\mathbf{g}$ whereas ψ_d consists of the diffusely scattered waves, which are assumed to be weak.

The z-dependent Fourier coefficients of the diffuse waves are given by the solution of the following set of coupled differential equations [obtained by inserting Eq. (7) into the Schrödinger equation]:

$$\frac{d^2}{dz^2}y(\mathbf{s},z)+k_{\mathbf{s}}^2 y(\mathbf{s},z)-S_0(\mathbf{s},z)-D_f(\mathbf{s},z)-C_2(\mathbf{s},z)=0. \tag{8}$$

Here, $k_{\mathbf{s}}$ is the z component of the (diffuse) wave vector associated with the scattering vector $\mathbf{s}$. The terms $S_0(\mathbf{s},z)$, $D_f(\mathbf{s},z)$, and $C_2(\mathbf{s},z)$ describe the couplings of the wave $\mathbf{s}$ with all other waves. In particular, these couplings are

$$S_0(\mathbf{s},z)=(1/2\pi)^2\sum_{\mathbf{g}} v(\mathbf{s}-\mathbf{g},z)S(\mathbf{s}-\mathbf{g})w(\mathbf{g},z),$$

$$D_f(\mathbf{s},z)=\sum_{\mathbf{g}} V_p(\mathbf{g},z)y(\mathbf{s}-\mathbf{g},z), \tag{9}$$

$$C_2(\mathbf{s},z)=(1/2\pi)^2\int d\mathbf{s}_2 v(\mathbf{s}-\mathbf{s}_2,z)S(\mathbf{s}-\mathbf{s}_2)y(\mathbf{s}_2,z).$$

S_0 (source term) denotes the coupling of $\mathbf{s}$ with the set of strong waves $\{\mathbf{g}\}$ via the nonperiodic part of the potential V_d. In perturbation theory this term is responsible for the excitation of the diffuse waves by the field of the strong waves $\mathbf{g}$ and can be regarded as the source term for the diffuse scattering. D_f (diffraction term) is the contribution arising from the interaction between the wave $\mathbf{s}$ and the diffuse set $\{\mathbf{s}+\mathbf{g}\}$, established by the periodic potential V_p. D_f describes the diffraction of the diffuse waves by V_p. C_2 (second-order coupling term) contains the coupling of $\mathbf{s}$ with the diffuse set $\{\mathbf{s}_2,\mathbf{s}_2-\mathbf{s}\neq\mathbf{g}\}$ via the nonperiodic potential V_d.

An expression analogous to Eq. (9) (without source term) holds for the coefficients $w(\mathbf{g},z)$ of the strong waves. In that case the second-order couplings describe the feedback of the diffuse waves into the strong waves.

A. The perturbation approach

The computer program based on perturbation theory, developed at the University of Osnabrück, employs approximations that enable a very effective handling of the scattering problem for a large number of statistically varying configurations. However, certain dynamical wave couplings have to be neglected in order to achieve this advantage. The periodic part ψ_p of the wave function ψ [see Eq. (7)] corresponding to the strong waves $\{\mathbf{g}\}$ is calculated fully dynamically for the periodic part of the potential V_p. In the case of the diffuse wave function ψ_d the diffraction term D_f, the couplings between a diffuse wave $\mathbf{s}$ and the set of diffuse waves $\{\mathbf{s}+\mathbf{g}\}$, is included exactly. D_f describes fully dynamically the propagation of the diffusely scattered waves in the periodic potential V_p. The contributions of S_0, i.e., the interactions between the diffuse and the strong waves $\mathbf{g}$, are included to first order in perturbation theory with respect to the diffuse potential part V_d. This means that the diffuse wave function ψ_d is assumed to be excited by the strong wave function ψ_p while the second-order recoupling of the diffuse waves to the strong ones is neglected. Finally, the coupling contribution C_2 is neglected as well because it is of second order with regard to the perturbation (coupling to *diffuse* waves by the *diffuse* potential). Formally, this approach is equivalent to the distorted wave Born approximation (e.g., Ref. 16) with $[(-\hbar^2/2m)\nabla^2+V_p(\mathbf{r})]$ as the unperturbed Hamiltonian and V_d as the perturbation. In practice, the numerical solution is obtained using the reflectivity matrix method[17] adapted to diffuse scattering.[10] This method and its numerical aspects are described in detail in Ref. 10.

The great advantage of the perturbation approach is that for occupational disorder, where the sites of the atoms in the disordered layer only differ by parallel translation vectors $\mathbf{r}_{n\|}$, which are compatible with the periodicity of V_p, the formalism can be reduced to a ''pseudokinematical'' scattering formula; i.e., the diffusely scattered amplitude, $F(\mathbf{s})$, of each disordered layer can be expressed as a product of the conventional structure factor with a ''dynamical'' form factor:

$$F(\mathbf{s})=f_{\mathrm{dyn}}(\mathbf{s})\sum_n \exp(-i\mathbf{s}\cdot\mathbf{r}_{n\|}), \quad \mathbf{s}\neq\mathbf{g}. \tag{10}$$

The dynamical form factor $f_{\mathrm{dyn}}(\mathbf{s})$ represents the scattering amplitude for a single atom, situated in the periodic potential part and is *independent* of the detailed statistics involved in the disorder; i.e., independent of the concrete configuration considered. The modulus of f_{dyn} depends strongly on both azimuthal and polar exit angle. The corresponding modulations are due to the diffraction of the diffusely scattered waves by the periodic potential part.

The information about the statistics is contained in the structure factor. Therefore, once $f_{\mathrm{dyn}}(\mathbf{s})$ is known, the scattering amplitude for an arbitrarily large number of different configurations can be calculated simply by the use of structure factors. In particular, the effective methods of dealing with the configuration problem developed for the kinematical scattering theory (e.g., Patterson analysis) can be transferred. The cpu time for the calculation of $f_{\mathrm{dyn}}(\mathbf{s})$ for one $\mathbf{s}$ point is approximately three times as high as the cpu time needed for one point of the rocking curve from the periodic part the same structure.

B. The supercell approach

The computer program based on the supercell approach, developed at the University of Leicester, does not neglect specific wave couplings and the scattering problem for a given configuration is solved exactly. In the case of the supercell method the second-order couplings C_2 as well as the full couplings S_0 with the set $\{\mathbf{g}\}$ of strong waves are fully included. Generally, the diffuse potential V_d (and thus the diffuse wave function ψ_d) has a continuous (two-dimensional) Fourier spectrum. In order to numerically handle a finite number of diffuse waves, a proper discretization of the Fourier transform of V_d (and ψ_d) into a Fourier series has to be carried out. Physically, this means that the considered configuration is repeated periodically with a period corresponding to the degree of discretization. In this way the diffuse scattering problem is reduced to the treatment of a periodic surface but with a very large surface unit cell (supercell). Therefore, the calculation can be undertaken using techniques developed for periodic structures.[8,9] As the computational effort is roughly proportional to the third power of the number of waves included, the calculation time for a reasonable degree of discretization is considerable. In

this work we used a (1×400) supercell and a total of 405 waves. Use of a highly optimized program, based on the approach of Ref. 15 enabled the calculations to be done without excessive use of cpu time.

IV. SURFACE MODEL AND CALCULATIONAL DETAILS

The surface model used for the investigation is based on the unreconstructed Si(100)(1×1) surface. On the ordered surface a bilayer with disorder along the $[0\bar{1}1]$ direction was introduced, consisting of bilayer terraces, with step edges along $[011]$. The basic ''bilayer unit'' consists of an upper terrace of extension 1 LU (lattice unit) in the $[01\bar{1}]$ direction ($1\text{LU}=a_0/\sqrt{2}=3.84$ Å) and a height corresponding to 2 ML ($a_0/2=2.715$ Å). The ''dynamical'' form factor for the perturbation treatment is then obtained using such a unit as perturbation. The configurations of the disordered bilayer were always constructed of such units within a supercell of length $L_{\text{cell}}=400$ LU. Usually the configurations chosen for the comparison were regular step arrays of various coverages θ. Such structures consist of one upper and one lower terrace within the supercell. The length of the ''up'' terrace is then θL_{cell} and the length of the ''down'' terrace is $(1-\theta)L_{\text{cell}}$.

The reason for this choice of these artificial configurations is to keep the computational effort of the rigorous calculations within a tractable time scale. Simulation of realistic disorder would in principle require many statistically varying configurations or at least a very large cell in order to introduce statistics within the supercell. However, as will be shown in Sec. V, the key structural parameters that influence the quality of the perturbation approach are the coverage θ of the disordered layer(s) and the correlation length L involved in the disorder along the incident beam azimuth. Regular step arrays now offer the unique advantage that a given correlation length can be easily realized by a *single configuration*. The correlation length L of the disorder simulated in that way is simply given by the length of the ''up'' terrace for coverages $\theta<0.5$ and by the length of the ''down'' terrace if $\theta\geqslant0.5$.

We believe that a comparison of data from these configurations provides a worst case estimation concerning the applicability of perturbation theory. Although the surface is not monatomically flat, these are nevertheless very regular structures (even within the supercell) such that the possible cancellation of the disturbing multiple scattering processes due to statistical variations in the disorder is excluded. For some comparisons, random configurations within the supercell were used as well (Sec. VI C). In all these cases the structures consisted of the bilayer units described above.

The calculations were carried out for a 15-keV electron energy. For the atomic scattering potentials the original values from the tables of Doyle and Turner[18] without any corrections for the inner potential were used. It is important to use those corrections for the analysis of experimental data.[19] They are, however, irrelevant with regard to the comparison of the two approaches as long as in both cases the same potential is used. The crystal temperature was assumed to be 300 K. The ratio of imaginary absorption potential to real potential was assumed to be 0.19 (see Ref. 20).

The incident beam azimuth was $[011]$, i.e., perpendicular to the step edges. The wave set for the rigorous calculations consisted of the 5 integral-order waves $(0,h)$ $(h=\overline{2},\ldots,2)$ in the zeroth Laue zone and the fractional (diffuse) waves $(m/400,h)$ $(m=\overline{40},\ldots,40, m\neq0)$ such that a total of 405 waves was included. It has been checked by the calculation of rocking curves for the flat surface that inclusion of more than 5 integral-order waves led to only small changes in the shape of the curves. Memory constraints made the restriction to 5 integral-order waves necessary. It has, however, been checked that an increase in the number of fractional-order waves did not change the results significantly. If we therefore use the same set of integral-order waves for the perturbation calculation, comparability is guaranteed and also the conclusions concerning the quality of the perturbation treatment should hold generally. The consistency of the ''Leicester'' and ''Osnabrück'' programs with regard to the periodic potential part was checked by the comparison of the rocking curves from the flat surface. The reflected *absolute intensities* produced by the two programs were found to agree within 0.5%. The metric distance [defined below in Eq. (11)] between rocking curves from the two programs was $D=0.005$.

Both the supercell and perturbation program profiles of the (00) and (01) streak were calculated for the 6 incident angles 34, 40, 45, 50, 55, and 60 mrad. The computations were performed on a HP 735 workstation. In the case of the rigorous supercell calculation the computation time needed for one incident angle was about 1.5 h utilizing the mirror symmetry of the scattering problem.[21] The corresponding calculation time in the case of the perturbation program to calculate the ''dynamical'' form factor was about 3 min (without utilizing the symmetry) on the same computer.

In all the comparisons only the diffuse intensity of a profile was used; i.e., the single point corresponding to the integral-order wave in the zeroth Laue zone was excluded. The results from both approaches were compared by means of the (root mean squared) metric distance D (as used already in earlier work[2]) to quantify the deviations between calculated and experimental data:

$$D(f,g)=\frac{1}{\sqrt{2}}\left[\sum_i\left(\frac{f_i}{[\Sigma_i f_i^2]^{1/2}}-\frac{g_i}{[\Sigma_i g_i^2]^{1/2}}\right)^2\right]^{1/2}. \quad (11)$$

The summations over i are over the exit angles of the profiles f and g. $D(f,g)=0$ if $f\propto g$ whereas $D(f,g)=1$ for the case that f and g are completely uncorrelated. This measure concerns only the shape of the curves and not their absolute intensities. This is on the one hand reasonable because in experiments usually only relative intensities are measured. On the other hand, first-order perturbation theory itself tends to produce higher intensities than the rigorous theory. It can be shown that the effect of second-order pertubations on the propagation of the incoming and the diffusely scattered waves is equivalent to the effect of an additional (in general nonlocal) absorptive potential.[22–24] For this reason, it is expected that prominent profile features that already arise from first order diffuse scattering appear with reduced intensity in a more rigorous treatment. Our numerical results have confirmed this trend. We have found discrepancies in absolute intensity between about $5-20\%$ for cases with the best agreement (as given by the D value) up to an order of mag-

nitude for the cases of worst agreement between the two approaches. Moderate changes of the absorption potential have a strong influence upon absolute intensities but have usually little influence upon the shape of curves (i.e., peak positions and their relative heights). Hence, in the figures that illustrate the comparison between both kinds of calculations, the curves are normalized to their maximum and plotted on a linear scale using arbitrary units.

V. ANALYTICAL CONSIDERATIONS

Before discussing the actual comparison between the supercell and perturbation approaches for the specific case of the Si(100) bilayer system, we shall use some simple analytical arguments to anticipate which physical quantities influence the quality of the perturbation approach. Although some strong simplifications will be used, the essential features important for high-energy scattering from (disordered) surfaces will be retained. The analytical considerations will in particular help us to clearly understand the numerical comparison and thus to transfer the results, obtained mainly for special configurations, to realistic surfaces.

In what follows an approximate error will be derived, which characterizes the strength of higher-order diffuse scattering processes and, thus, the quality of a perturbational treatment.

A. Derivation of the approximate error

For simplicity, we assume a monatomic material and a single disordered top layer upon an otherwise perfect bulk. All atoms in the disordered layer occupy lattice sites compatible with the structure of the ordered system (occupational disorder). The origin of the z coordinate (normal to the surface) is the center of the disordered layer. Furthermore, we consider diffuse scattering from the surface Brillouin zones associated with the zeroth-order Laue zone.

In what follows we will estimate for the system of equations (8) the magnitude of the second-order coupling term C_2, which is neglected in the perturbation treatment, compared to the source term S_0. To that end we estimate first the magnitude of the diffuse coefficients $y(\mathbf{s},z)$. This is performed by the integration of Eq. (8) retaining, of all the coupling terms, only the source term S_0 (the couplings with the strong waves). By approximating the exact Green's function for Eq. (8) with the corresponding free-electron Green's function we obtain

$$y(\mathbf{s},z)=\frac{1}{2ik_{\mathbf{s}}}S(\mathbf{s})\sum_{\mathbf{g}} I(\mathbf{g},\mathbf{s},z), \tag{12}$$

where $I(\mathbf{g},\mathbf{s},z)$ denotes the integral

$$I(\mathbf{g},\mathbf{s},z)=\frac{1}{(2\pi)^2}\int dz_2 \exp(ik_{\mathbf{s}}|z-z_2|)v(\mathbf{s}-\mathbf{g},z_2)w(\mathbf{g},z_2). \tag{13}$$

Both equations hold for propagating ($k_{\mathbf{s}}=\sqrt{k_{\mathbf{s}}^2}$) as well as for evanescent waves ($k_{\mathbf{s}}=i\sqrt{-k_{\mathbf{s}}^2}$).

Physically, the Green's-function approximation used to obtain Eq. (12) means that the diffuse scattering is excited by the strong wave field associated with the periodic potential part and that the diffusely scattered waves propagate as in the vacuum. The latter behavior is due to the neglect of the diffraction term D_f and means in principle a quite drastic simplification because the propagation does not contain any diffraction by the strong periodic potential V_p. The propagation within V_p leads, besides the general trends derived here, indeed to important additional effects depending on the specific diffraction condition (see Sec. VII). Note that the actual perturbation theory used for the calculations of course includes all the couplings D_f. For the general estimation of the order of magnitude of the amplitude of the diffuse waves, however, the simplified expression should be sufficient because D_f represents no source of diffuse scattering but rather redistributes among the diffuse waves that flux which is scattered into the diffuse waves. The simplified expression used here in particular retains the important property that states with low $|k_{\mathbf{s}}|$ tend to be excited strongly (e.g., for surface resonances[25]), and it allows the existence of evanescent waves. We note further that the simplified propagation does not include absorption, which would lead in general to smaller values for $y(\mathbf{s},z)$. In particular, the singularity at $k_{\mathbf{s}}=0$ would disappear in that case so the expression obtained here tends to overestimate the magnitude of the diffuse scattering.

The couplings C_2, which are neglected in the perturbation scheme, are now estimated by insertion of (12) into the third equation of (9). Exploiting the fact that the two-dimensional Fourier coefficients (6) of the atomic potential are only slowly varying within a Brillouin zone, one obtains

$$C_2(\mathbf{s},z)=\sum_{\mathbf{g}} \Delta C(\mathbf{s}-\mathbf{g},z), \tag{14}$$

such that each BZ that belongs to the reciprocal lattice vector $\mathbf{g}$ contributes to the second-order diffuse coupling with the amount

$$\Delta C(\mathbf{s}-\mathbf{g},z)=\frac{1}{(2\pi)^2}v(\mathbf{s}-\mathbf{g},z)\int_{\mathrm{BZ}_{\mathbf{g}}}\frac{d\mathbf{s}_2}{2ik_{\mathbf{s}_2}} \times S(\mathbf{s}-\mathbf{s}_2)S(\mathbf{s}_2)\sum_{\mathbf{g}_2} I(\mathbf{g}_2,\mathbf{s}_2,z). \tag{15}$$

Here, the sum $\mathbf{g}_2$ is over the surface reciprocal lattice vectors and $\int_{\mathrm{BZ}\ \mathbf{g}} d\mathbf{s}_2$ means integration over the BZ associated with $\mathbf{g}$.

Analogously, for the source term S_0 (the coupling with the strong waves), each BZ $\mathbf{g}$ contributes the amount

$$C(\mathbf{s}-\mathbf{g},z)=(1/2\pi)^2 v(\mathbf{s}-\mathbf{g},z)S(\mathbf{s})w(\mathbf{g},z). \tag{16}$$

We now estimate the contribution $\Delta C(\mathbf{s}-\mathbf{g},z)$ of the second-order couplings compared to the first-order contribution $C(\mathbf{s}-\mathbf{g},z)$. The evaluation is carried out for the center of the disordered layer ($z=0$), the potential at this particular z coordinate being representative for the strength of the potential.

Quantitative evaluation of Eq. (15) for arbitrary $\mathbf{s}$ and in particular the derivation of an approximate analytical expression is practically not possible. The reason for this is that the product of the structure factors $S(\mathbf{s}-\mathbf{s}_2)$ and $S(\mathbf{s}_2)$ usually oscillates strongly with $\mathbf{s}_2$ and thus prevents a simple evalu-

ation of the integral over $\mathbf{s}_2$. A further complication arises from the fact that the second-order diffuse scattering contribution from single scattering on two different atoms does not have the same θ dependence as the second-order contribution from double scattering on the same atom. Thus $C(\mathbf{s}-\mathbf{g},z)$ is generally a complicated function of both $\mathbf{s}$ and θ but for practical purposes we would like to have an error measure that is independent of $\mathbf{s}$.

The problems addressed above can be avoided if we evaluate $\Delta C(\mathbf{s}-\mathbf{g},z)$ at the center of the Brillouin zone associated with the reciprocal surface vector $\mathbf{h}$. In that case the product of the structure factors $S(\mathbf{s}-\mathbf{s}_2)$ and $S(\mathbf{s}_2)$ reduces to the positive expression $|S(\mathbf{s}_2)|^2$ and a comparably simple discussion of the behavior of the BZ integral over $\mathbf{s}_2$ in Eq. (15) is possible. In particular, an approximate explicit expression with a physically sensible coverage dependence can be found. Strictly, however, $\Delta C(\mathbf{h}-\mathbf{g},z)$ refers to the neglected second-order couplings for the integer wave field corresponding to the reciprocal lattice vectors $\mathbf{h}$ and not to the diffuse wave field.[26] It is on the other hand physically clear that this quantity can be used for our purposes because it represents a measure for higher-order diffuse scattering, which is generally neglected in the perturbation approach. Further, we have found that the error measure based on this approach gives a good description of our numerical results.

In order to define a relative error that compares the neglected diffuse couplings from BZ $\mathbf{g}$ with the couplings included in perturbation theory, a suitable normalization of $\Delta C(\mathbf{h}-\mathbf{g},z)$ has to be carried out. Among the various possibilities for a normalization we have chosen a comparison of $\Delta C(\mathbf{h}-\mathbf{g},z)$ with the source terms $C(\mathbf{s}-\mathbf{g},0)$ of the diffuse scattering [Eq. (16)] and defined a relative error by

$$\epsilon_{\mathbf{g}}=\sqrt{2}\left[\frac{|\Delta C(\mathbf{h}-\mathbf{g},0)|^2}{(N\Omega/4\pi^2)\int_{\mathrm{BZ}\ \mathbf{h}} d\mathbf{s}|C(\mathbf{s}-\mathbf{g},0)|^2}\right]^{1/2}. \tag{17}$$

The normalization to the sum of all the source couplings in BZ $\mathbf{h}$ [i.e., the integral in the denominator of Eq. (17)] is necessary in order to account for the different scaling behavior of $\Delta C(\mathbf{h}-\mathbf{g},0)$ and $C(\mathbf{s}-\mathbf{g},0)$ with the number of available sites N.

The dimensionless error $\epsilon_{\mathbf{g}}$ defined in this way has the advantage that it is related to the order of magnitude of the metric distance D [see Eq. (11)] used to quantify the degree of agreement between diffuse profiles calculated with the perturbation and supercell approaches. As can be seen from Eq. (8) in its integrated form, $\epsilon_{\mathbf{g}}$ is related to the relative error in amplitude $\sqrt{2}\Delta A/A$. The metric distance D roughly measures the relative error in intensity $\Delta I/(I\sqrt{2})$, which is in turn about $\sqrt{2}\Delta A/A$. This correspondence should, of course, not be exaggerated particularly because the second-order term in the numerator of $\epsilon_{\mathbf{g}}$ refers to the neglected couplings affecting the integer instead of the diffuse waves. Nevertheless, as will be demonstrated in Sec. VI, $\epsilon_{\mathbf{g}}$ is able to predict the results of the numerical comparison even semiquantitatively quite well.

In order to elucidate which physical parameters essentially influence the relative error, we estimate the typical order of magnitudes of the quantities under the integral in the expression for the second-order couplings, Eq. (15). For the z component $k_{\mathbf{s}}$ of the diffuse wave vector corresponding to the lateral scattering vector $\mathbf{s}=(s_x,s_y)$, the relation

$$k_{\mathbf{s}}^2=k_0^2-s_y^2-2K_0 s_x \tag{18}$$

holds to a very good approximation for the RHEED case (high electron energy, low glancing angles and strong forward scattering). The indices x and y refer to the vector components along (x) and perpendicular (y) to the incident beam azimuth, respectively, and $\mathbf{K}_0$ is the wave vector of the incident electrons of energy $E=(\hbar K_0)^2/2m$. Within the BZ, $k_{\mathbf{s}}$ varies only slowly as a function of s_y, but varies strongly as a function of s_x because K_0 is large. The relevant s_x are roughly between $-\pi/L$ and π/L where L denotes the correlation length of the disorder along the incident azimuth. Therefore, the correlation length L has a strong influence on the magnitude of the integral in Eq. (15) over the second-order couplings. For sufficiently short L, the modulus of the relevant $k_{\mathbf{s}}$ is typically between 0 and $\sqrt{\Lambda^{-1}}=\sqrt{2\pi K_0/L}$. We introduce $k_{\mathbf{s}}'=\sqrt{\Lambda}k_{\mathbf{s}}$ as a useful dimensionless quantity. We have found that the quantity,

$$\Lambda=L/2\pi K_0 \tag{19}$$

is an important parameter. Because the parallel component of the electron momentum is large, Λ^{-1} is approximately equal to the change in the square of the perpendicular momentum, that is $\Lambda^{-1}\simeq k_{\mathbf{0}}^2-k_{\mathbf{s}}^2$, and in Sec. V B we will show that the corresponding energy,

$$\Delta E_{\perp}=\frac{\hbar^2}{2m}\Lambda^{-1}, \tag{20}$$

has a strong influence on the quality of the perturbation approach.

The following further dimensionless quantities (indicated by the prime symbol) are introduced under the integral in Eq. (15). $\mathbf{s}$ is normalized to reciprocal surface lattice units such that $d\mathbf{s}'=\Omega/(4\pi^2)d\mathbf{s}$. For the structure factor S we write $S'=S/\sqrt{N\theta(1-\theta)}$, which yields $\int_{\mathrm{BZ}}d\mathbf{s}'|S'|^2=1$ [Eq. (3)], and the atomic potential coefficients v are normalized to $v(\mathbf{s}=0,z=0)$. Finally, we introduce $z'=z/\sqrt{\Lambda}$. Now the relative error can be expressed as

$$\epsilon_{\mathbf{g}}=\frac{1}{\sqrt{2}}\frac{2m}{\hbar^2}V_{\mathrm{av}}(0)\sqrt{\theta(1-\theta)}\Lambda \times\left|\frac{1}{w(\mathbf{g},0)}\int_{\mathrm{BZ}_{\mathbf{g}k_{\mathbf{s}'}'}}\frac{d\mathbf{s}'}{}|S'(\mathbf{s}')|^2\sum_{\mathbf{g}_2'}I'(\mathbf{g}_2',\mathbf{s}',0)\right|, \tag{21}$$

with

$$I'(\mathbf{g}_2',\mathbf{s}',0)=\int dz'\ \exp(ik_{\mathbf{s}'}'|z'|) \times v'(\mathbf{s}'-\mathbf{g}_2',z'\sqrt{\Lambda})w(\mathbf{g}_2',z'\sqrt{\Lambda}). \tag{22}$$

$V_{\mathrm{av}}(0)$ is the laterally averaged atomic potential at $z=0$. Because v' as a function of $\mathbf{s}'$ varies only slowly within the

BZ, the integration with regard to $\mathbf{s}'$ is essentially an integral over the product of $|S'(\mathbf{s}')|^2$ and a function depending on $k'_{\mathbf{s}'}$.

The expression (21) reveals that except for a dimensionless integral that depends in a complicated manner on the structure factor profile, the shape of the atomic potential and of the exciting wave field, most generally, very low and very high coverages θ lead to a small error $\epsilon_{\mathbf{g}}$ because in both cases the concentration of defects (with regard to the periodic potential part of the disordered layer) is very small. The coverage dependence reaches a maximum for $\theta=0.5$, where the largest concentration of defects appears. A further general feature is the importance of the strength of the atomic potential in the disordered layer $V_{\mathrm{av}}(0)$. These two dependences on the defect concentration and the potential are also physically evident as both determine most generally the magnitude of the perturbation.

B. Discussion of limiting cases

As the magnitude of the integral in Eq. (21) can depend on Λ, the error is not generally proportional to Λ as one might infer from the term in front of the modulus signs. Actually, the kind of dependence of $\epsilon_{\mathbf{g}}$ on Λ (which means for fixed electron energy the dependence on the correlation length L) depends on the magnitude of Λ compared with the magnitude of the z component $k_{\mathbf{g}}$ of the wave vector associated with the strong wave $\mathbf{g}$.

1. The case $\Lambda \ll |k_{\mathbf{g}}|^{-2}$

We first consider the case of small Λ ($\ll |k_{\mathbf{g}}|^{-2}$). Then the definition of $k'_{\mathbf{s}'}$ implies that for those $\mathbf{s}'$ where $|S'(\mathbf{s}')|^2$ is strong, $k'_{\mathbf{s}'}$ always covers about the same range ($|k'_{\mathbf{s}'}| \leqslant 1$), independent of Λ. Hence, the $\mathbf{s}'$ integral only depends weakly on Λ. Also the z' integral I' depends only weakly on Λ in this case: Then z' varies strongly over the relevant range μ' of the potential coefficient v'. If $k'_{\mathbf{s}'}$ corresponds to a propagating wave the exponential has a large number of oscillations within the range μ' but if $k'_{\mathbf{s}'}$ corresponds to an evanescent wave, the exponential decreases rapidly. In both cases the magnitude of I' is determined mainly by the magnitude of $k'_{\mathbf{s}'}$ and to a lesser exent by the Λ-dependent μ'. To summarize, for small correlation length the whole expression between the modulus signs is more or less independent of Λ and the relative error $\epsilon_{\mathbf{g}}$ is thus essentially proportional to Λ. Alternatively, using definition (20), one can say that in this case the error follows the law

$$\epsilon_{\mathbf{g}} \propto \sqrt{\theta(1-\theta)}\frac{V_{\mathrm{av}}(0)}{\Delta E_{\perp}}; \tag{23}$$

i.e., the error is proportional to the strength of the perturbation potential compared with the energy parameter, $\Delta E_{\perp}$, associated with the relevant diffuse waves.

Consequently, a small Λ generally leads to a small error such that for sufficiently short correlation length and/or high electron energy perturbation theory can work even for high defect concentrations and strong scattering potentials. In this case the strong forward scattering at high energies corresponds to an averaging over the disorder. We note that this explains the good agreement between rigorous and perturbation calculation in the case of the random configuration in Fig. 2(b). Besides the above rather formal arguments, this behavior can also be physically rationalized. The scattering at high electron energy and low glancing angles implies that small momentum transfer *along* the incident azimuth is connected with a large momentum transfer *normal* to the surface. Therefore, most of the diffuse waves will have a comparatively large momentum component normal to the surface and will be dynamically scattered in the ordered bulk (where the perturbation approach describes the multiple scattering of the diffuse waves exactly) rather than in the disordered layer. In the case of the evanescent diffuse waves the large perpendicular (imaginary) momentum leads to a rapid exponential decay of the corresponding part of the wave function. This also decreases the probability of multiple diffuse scattering of these waves within the disordered layer.

2. The case $\Lambda \gg |k_{\mathbf{g}}|^{-2}$

If Λ is large ($\gg |k_{\mathbf{g}}|^{-2}$), the structure factor related function $|S'(\mathbf{s}')|^2$ is sharply peaked around the reciprocal surface lattice vector $\mathbf{g}'$ and thus the relevant $k'_{\mathbf{s}'}$ will be of order $\sqrt{\Lambda}k_{\mathbf{g}}$. In that case the integral I' will be proportional to $\sqrt{\Lambda^{-1}}$. Together with the $k'_{\mathbf{s}'}$ in the denominator of (21) the whole expression between the modulus signs of Eq. (21) will be proportional to Λ^{-1} and thus cancel the factor Λ on the left-hand side of the modulus. Therefore, for large Λ, the dependence of the error on this parameter disappears. Moreover, the range of relevant $k_{\mathbf{s}}$ lies in a small region around $k_{\mathbf{g}}$ in that case and the error $\epsilon_{\mathbf{g}}$ therefore tends to increase inversely to $|k_{\mathbf{g}}|$. (The simplified error $\epsilon_{\mathbf{g}}$ [Eq. (29) in Sec. V C] is proportional to $|k_{\mathbf{g}}|^{-2}$ when $|k_{\mathbf{g}}|$ is large. As in the case of small correlation lengths, the error is then essentially determined by the perturbation potential compared with the perpendicular energy of the diffuse waves.) This means that for very large correlation length $\epsilon_{\mathbf{g}}$ approaches a singularity if $k_{\mathbf{g}} \to 0$. Then, the neglected coupling contributions from BZ $\mathbf{g}$ inevitably become very strong and perturbation theory is likely to fail. Physically, such a situation means that all the diffuse waves of the BZ move parallel to the disordered layer and can thus undergo a large number of multiple scattering events by the nonperiodic potential part. These arguments are slightly weakened for the real situation because absorption (which was not taken into account for this analytical estimation) will reduce the number of multiple scattering events. In particular the singularity for $k_{\mathbf{g}}=0$ would disappear.

On increasing $|k_{\mathbf{g}}|$ the error $\epsilon_{\mathbf{g}}$ tends to decrease so that for sufficiently large $|k_{\mathbf{g}}|$ the perturbation approach tends to become adequate again in spite of a large correlation length. Physically, the diffuse waves of the BZ $\mathbf{g}$ are then traveling more and more normally to the disordered layer. In that case the multiple scattering of the diffuse waves is strongly reduced in the defect layer and will essentially be due to the ordered bulk underneath. The dynamical scattering within the ordered layers, however, is treated rigorously in the perturbation approach. This favorable situation can be approached experimentally for high glancing angles if the majority of those integral-order waves that are important to

describe the multiple scattering have emerged from the shadow edge.

3. The case $\Lambda = |k_{\mathbf{g}}|^{-2}$

We finally note that a particularly unfavorable situation appears for $\Lambda = \Lambda_{\max} = |k_{\mathbf{g}}|^{-2}$ if $\epsilon_{\mathbf{g}}$ is considered for fixed material, coverage, and diffraction conditions as a function of the correlation length L. According to Eq. (21) the contribution of a wave $\mathbf{s}$ to the error is weighted with $|k_{\mathbf{s}}|^{-1}$, which reflects the general fact that second-order scattering events are more likely when the diffuse waves in the disordered layer propagate parallel to it. With regard to this aspect a profile $|S'|^2$ that is ''smeared out'' along the x direction (incident azimuth) causes a larger error than a sharply peaked one as long as the corresponding relevant diffuse waves in the BZ are either *all propagating* or are *all evanescent*. The reason for this is that $|k_{\mathbf{s}}|^{-1}$ is a concave function of s_x. The most unfavorable situation appears if the diffuse streak of BZ $\mathbf{g}$ has either just completely emerged from the shadow edge or just completely vanished, i.e., for $\Lambda_{\max} = |k_{\mathbf{g}}|^{-2}$. Therefore a decrease of the error with the correlation length is to be expected if $\Lambda > |k_{\mathbf{g}}|^{-2}$. This may of course be partly weakened for certain diffraction conditions by the structure of the exciting wave field and/or the specific potential shape. However, the above arguments hold quite generally for scattering from surfaces at glancing angles and the predicted behavior will be confirmed by the numerical results in Sec. VI.

C. Simplified expression

In order to obtain a simple explicit expression for $\epsilon_{\mathbf{g}}$ that may be used as a ''rule of thumb,'' the magnitude of the integral in Eq. (21) has to be evaluated. To that end a number of somewhat crude assumptions are employed where, however, the relevant orders of magnitude are retained. The main assumptions are a rectangular structure factor profile, an exponential z dependence $\sim \exp(-|z|/\mu)$ for the atomic potential coefficients, and the neglect of the z dependence for the wave coefficients $w(\mathbf{g},z)$. This latter assumption appears, at first glance, extremely crude. It can, however, be verified by comparison of this assumption with the more realistic form $w(\mathbf{g},z) \sim \exp(\pm k_{\mathbf{g}} z)$ that due to the neglect of additional oscillations (or an exponential decrease in the case of evanescent waves) the simplification generally tends to overestimate the magnitude of the z integral (22). Only if both $k_{\mathbf{g}}$ and $k_{\mathbf{s}}$ correspond to propagating waves does the simplification lead to a significant underestimation in the region $k_{\mathbf{s}} \approx k_{\mathbf{g}}$.

The main steps of this somewhat lengthy evaluation can be found in the Appendix. The result for the error $\epsilon_{\mathbf{g}}$ is

$$\epsilon_{\mathbf{g}} = \frac{1}{\sqrt{2}} \sqrt{\theta(1-\theta)} \frac{d}{\mu} \frac{2m}{\hbar^2} V_0 \Lambda U(k'_{\mathbf{g}}, \mu'). \tag{24}$$

Here, θ is the coverage of the disordered layer, d is the distance between neighboring layers, μ the range of the atomic potential, and V_0 the mean potential (average over a corresponding bulk unit cell) of the atoms in the disordered layer. The form of the function U in Eq. (24) depends on whether diffuse propagating and evanescent waves appear simultaneously within BZ $\mathbf{g}$ or if only one wave type is present:

$$U = \begin{cases} U_1 & \text{if } |k'_{\mathbf{g}}|^{-2} \leqslant 1 \\ U_2 & \text{otherwise.} \end{cases} \tag{25}$$

U_1 applies if in BZ $\mathbf{g}$ both propagating and evanescent waves appear. This means that only a part of the corresponding diffuse streak has emerged from the shadow edge. U_2 is used if all diffuse waves are either propagating or evanescent. With the abbreviations $\beta_{\pm}^2 = \mu'^2(|k'_{\mathbf{g}}|^2 \pm 1)$, U_1 and U_2 are given by

$$U_1(\beta_+, \beta_-) = \left[\arctan^2 \beta_+ + \ln^2 \frac{\sqrt{1+\beta_+^2}}{1+|\beta_-|} \right]^{1/2}, \tag{26}$$

and

$$U_2(\beta_+, \beta_-) = \left[(\arctan\beta_+ - \arctan\beta_-)^2 + \frac{1}{4} \ln^2 \frac{1+\beta_+^2}{1+\beta_-^2} \right]^{1/2}. \tag{27}$$

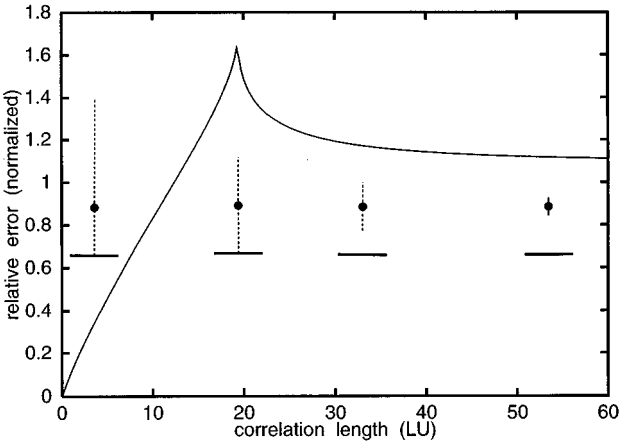

FIG. 3. Plot of the relative error $\epsilon_{\mathbf{g}}$ vs correlation length L. $\epsilon_{\mathbf{g}}$ measures the order of magnitude of the diffuse couplings from Brillouin zone $\mathbf{g}$, which are neglected in perturbation theory. The ordinate is normalized to the coverage θ-dependent factor $\sqrt{\theta(1-\theta)}$. The plot is for the Si(100) surface and a 15-keV electron energy. The perpendicular wave vector $|k_{\mathbf{g}}|$ of the integral-order wave is 2.3 Å^{-1}. The insets sketch the extension of the diffuse profile (dashed line) with regard to the shadow edge and the position of the sharp beam (dot) for various L.

The explicit result Eq. (24) reveals the same features as already discussed qualitatively by means of the more general expression (21). Figure 3 shows how, using Eq. (24), the error $\epsilon_{\mathbf{g}}$ depends on the correlation length L of the disorder along the incident azimuth for $|k_{\mathbf{g}}| = 2.3$ Å^{-1}. The other parameters are the same as used in the numerical investigation [Si(100) surface, $E = 15$ keV, $d = 1.36$ Å, $V_0 = 13.92$ eV, $\mu = 0.43$ Å (Ref. 27).] The ordinate is normalized to the coverage-dependent factor $\sqrt{\theta(1-\theta)}$.

For small L the error increases linearly with the correlation length and is independent of $k_{\mathbf{g}}$. In that case Eq. (24) can be simplified to

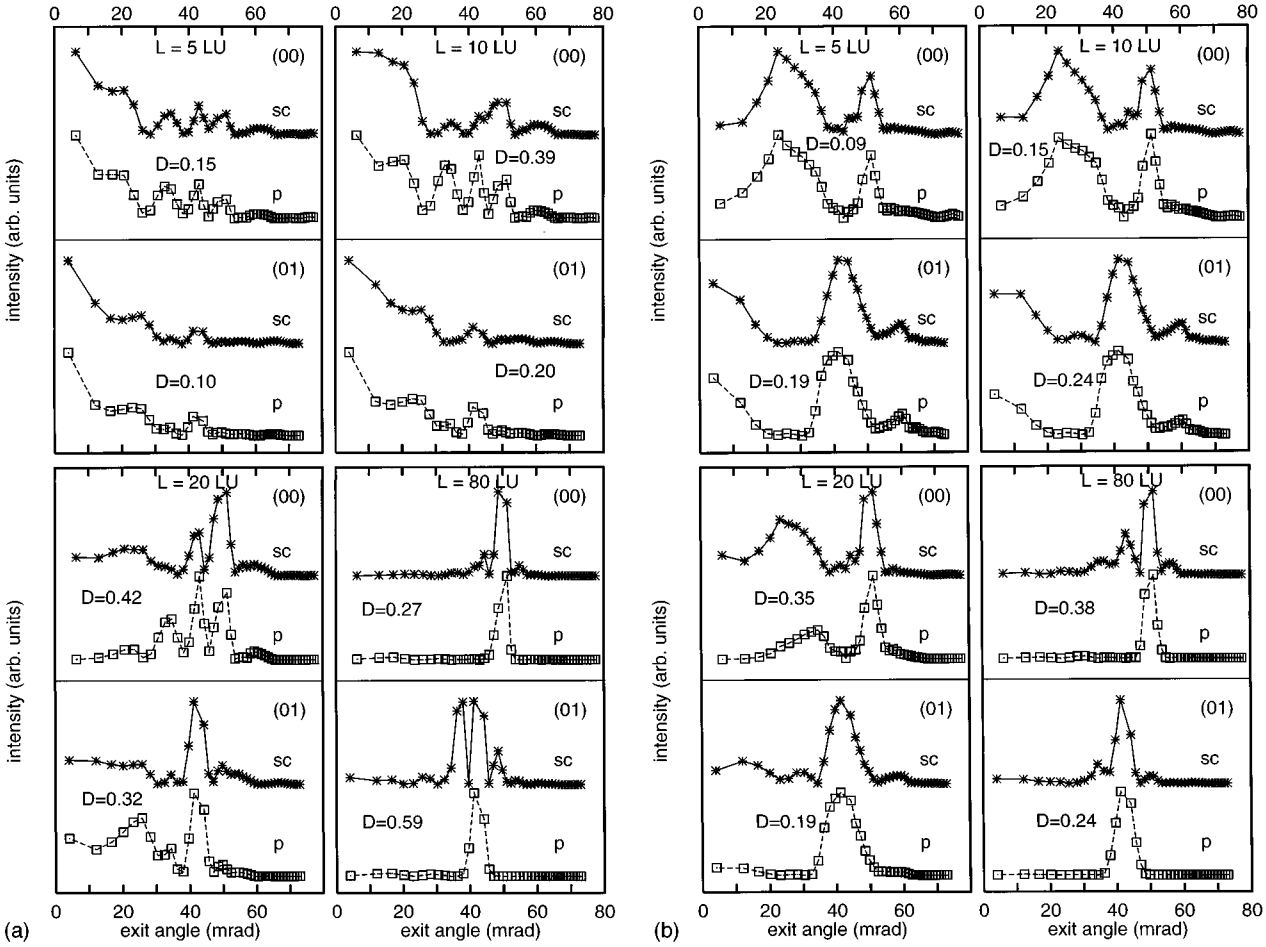

FIG. 4. Profiles from rigorous supercell (sc) calculations and perturbation theory (p) for regular bilayer step arrays on Si(100). The electron energy is 15 keV, the azimuth is [011] and the incident angle is 50 mrad. The single point corresponding to the integral-order wave in the zeroth Laue zone is excluded. (a) For coverages $\theta<0.5$. (b) For coverages $\theta>0.5$. L denotes the corresponding correlation length of the simulated disorder and D the metric distance between sc and p.

$$\epsilon=\frac{\pi}{2\sqrt{2}}\sqrt{\theta(1-\theta)}\frac{d}{\mu}\frac{2m}{\hbar^2}V_0\Lambda \tag{28}$$

such that, for fixed coverage and material parameters, only Λ determines the quality of the perturbation approach. $\epsilon_{\mathbf{g}}$ reaches a maximum for the unfavorable situation $\Lambda_{\max}=|k_{\mathbf{g}}|^{-2}$ where the diffuse streak of BZ $\mathbf{g}$ has either just completely emerged from the shadow edge (if $|k_{\mathbf{g}}|^2>0$) or just vanished (if $|k_{\mathbf{g}}|^2<0$). On increasing the correlation length beyond the corresponding $\Lambda_{\max}$, $\epsilon_{\mathbf{g}}$ decreases again and reaches finally a constant value for $\Lambda\to\infty$. The limiting value for very large correlation lengths is

$$\epsilon_{\mathbf{g}}^{\infty}=\frac{1}{\sqrt{2}}\sqrt{\theta(1-\theta)}\frac{d}{\mu}\frac{2m}{\hbar^2}V_0\frac{1}{|k_{\mathbf{g}}|}(|k_{\mathbf{g}}|^2+\mu^{-2})^{-1/2}. \tag{29}$$

If the correlation length is large, the diffuse streak of BZ $\mathbf{g}$ shrinks to a small intensity region around the sharp principal beam associated with the reciprocal lattice vector $\mathbf{g}$. In this case $\epsilon_{\mathbf{g}}$ depends essentially on $|k_{\mathbf{g}}|$ and decreases monotonically if this quantity is increased.

Besides giving a quality measure for the use of perturbation theory, the above considerations further reveal the important physical quantities to be considered for possible second-order corrections to types of approaches related to the distorted wave Born approximation, e.g., by simulating the influence of the neglected couplings C_2 through an additional part of the absorptive potential. Our result shows in particular that for rough surfaces any realistic correction should include, not only the potential strength and defect concentration, but also the disorder correlation length along the incident azimuth.

VI. NUMERICAL RESULTS

In this section we will present the comparison between the numerical results from the perturbation theory and supercell approach for the Si(100) model surface described in Sec. IV.

A. Correlation length and defect concentration

In Fig. 4 profiles along the (00) and (01) streaks are plotted for the (1×400) supercell with regular step arrays of various terrace width, obtained from the rigorous calculation and the perturbation approach. The correlation length L of the simulated short-range order in the case of regular step

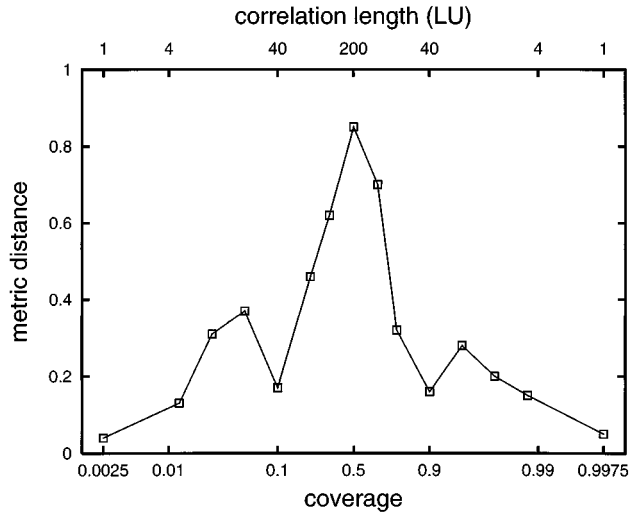

FIG. 5. Metric distance between profiles from perturbation theory and supercell approach, calculated for regular step arrays, vs coverage and correlation length. The incident angle is 50 mrad. The correlation length axis has a logarithmic scale. Note the equivalence of the cases $\theta<0.5$ and $\theta>0.5$.

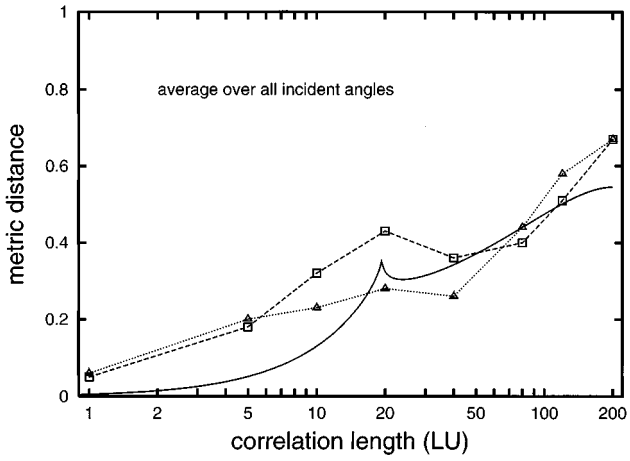

FIG. 6. Metric distance between profiles from perturbation theory and the supercell approach, calculated for regular step arrays, vs correlation length. The metric distance is averaged over all incident angles used for the investigation. Dashed curve: coverage $\theta\leqslant 0.5$. Dotted curve: $\theta\geqslant 0.5$. Solid curve: error from the analytical expression. For further explanation, see text.

arrays is directly related to the coverage and the length L_{cell} of the supercell. For fixed size of the supercell a given correlation length is realized by two different coverages, namely, L/L_{cell} and $1-L/L_{\mathrm{cell}}$.

The figure shows how the agreement between the two types of calculations changes with varying coverage and/or correlation length L. L is given in lattice units (1 LU=3.84 Å) along the incident azimuth. All profiles correspond to the same incident beam direction and were all obtained using a 50-mrad glancing angle. The case $\theta<0.5$ is shown in Fig. 4(a) and the case $\theta>0.5$ in Fig. 4(b). For small coverage θ the agreement is excellent but clearly decreases with increasing θ (increasing L). For coverages $\theta>0.5$ the quality of the perturbation approach improves again when θ increases (L decreases) and virtually perfect agreement with the rigorous calculation is reached for very high coverages. This behavior is illustrated in a more quantitative manner in Fig. 5. Here the metric distance D [see Eq. (11)], averaged over the (00) and (01) profile, between both types of calculations for the profiles of Fig. 4 is plotted versus coverage and the corresponding correlation length. The symmetrical dependence of D on θ about $\theta=0.5$ is due to the equivalence of the coverages L/L_{cell} and $1-L/L_{\mathrm{cell}}$. Both cases correspond to the same defect concentration with regard to the periodic potential part and to the same correlation length of the short-range order.

Figure 6 illustrates the general trend for the agreement between perturbation and supercell approaches versus the correlation length L for the regular step arrays. The degree of agreement is again expressed in terms of the metric distance, this time averaged over the (00) and (01) diffuse profile and over all 6 incident angles (see Sec. IV) used for the investigation. In this way possibly misleading effects due to particularly favorable or unfavorable diffraction conditions are largely avoided and thus the general trend should be represented quite well. The graph is the result of a comparison of a total of 192 different diffuse profiles. The dashed curve was obtained for coverages $\theta\leqslant 0.5$ ($L=L_{\mathrm{cell}}\theta$); the dotted curve holds for the case $\theta\geqslant 0.5$ [$L=L_{\mathrm{cell}}(1-\theta)$]. Both curves develop very similarly due to the equivalence of the two coverage cases. The general increase of the metric distance with the correlation length L is clearly seen. Because regular step arrays were considered, a given L is always connected with one of the two coverages L/L_{cell} and $1-L/L_{\mathrm{cell}}$ (both coverages correspond to the same defect concentration with regard to the periodic potential part). Therefore, the graph includes the influences of both the dependence on the correlation length as well as the dependence on the defect concentration. A notable feature is that the general increase of the metric distance with L ceases at about 20 LU and even becomes a slight decrease in spite of the increase of the defect concentration. Above about 50 LU an increase is seen again.

This apparently peculiar behavior can be easily rationalized in terms of the analytical considerations of Sec. V. There, we considered the error $\epsilon_{\mathbf{g}}$, which measures for the surface Brillouin Zone $\mathbf{g}$ the influence of the neglect of the second-order diffuse scattering processes in perturbation theory [Eq. (24)]. It was shown that the dependence of $\epsilon_{\mathbf{g}}$ on the correlation length L has a maximum that is related to the wave vector K_0 of the incident beam and to the z component of the wave vector associated with $\mathbf{g}$ by $L_{\max}=2\pi K_0/|k_{\mathbf{g}}|^2$. The typical magnitude of the $|k_{\mathbf{g}}|$ involved in the calculations is about 2.3 Å^{-1} (being the average over all $|k_{\mathbf{g}}|$ corresponding to the 5 integral-order waves and 6 incident angles between 34 and 60 mrad). This value corresponds to an $L_{\max}$ of 20 LU, which is identical to the correlation length in Fig. 6 at which the increase of the metric distance with L discontinues. According to Fig. 3 the dependence of the error $\epsilon_{\mathbf{g}}$ on the correlation length virtually disappears soon after $L_{\max}$ has been reached. This means that in the case of the graph in Fig. 6, which represents the overview for the regular step arrays, the increase of the metric distance with L beyond about 50 LU is essentially deter-

mined by the dependence on the coverage (i.e., the dependence on the defect concentration).

The solid line in Fig. 6 shows for the regular step array case the relative error $\epsilon_{\mathbf{g}}$ as a function of L for the average value of $|k_{\mathbf{g}}| = 2.3$ Å^{-1} (see above). This curve reproduces the general trend of the above results astonishingly well, even in a semiquantitative sense. We note that the over optimistic values from the analytical expression for low metric distances ($D < 0.2$) are not relevant because in this region the agreement between perturbation and supercell approach can be still considered as good. The expression predicts in particular the conditions in which problems with perturbation theory are expected. This is a quite useful result because with the simple expression Eq. (24) for $\epsilon_{\mathbf{g}}$ it is now possible to assess the quality of the perturbation approach for a given material and surface without carrying out time consuming reference calculations. Such a quantitative estimation of the possible error is regarded as more appropriate than quoting a fixed limiting condition since the accuracy desired for a theory normally depends on the application. For instance, in cases where a more qualitative understanding of the diffuse diffraction features is sufficient, the requirements concerning the accuracy are usually lower compared to the requirements for an exact calculation.

B. The single defect case

Of particular interest is the comparison for the case of a single defect because it is closely related to the scattering problem of one atom (or one molecule) situated on a perfect surface. For our (1×400) supercell model, this corresponds to the coverages 1/400 and $1 - 1/400$. In the first case only a single bilayer unit within the supercell is situated on the ordered substrate, whereas the latter coverage corresponds to a single bilayer unit missing from the otherwise perfectly ordered surface of the supercell.

Figure 7 illustrates the comparison for the single defect case by means of a series of profiles at various incident angles. The agreement between perturbation and supercell approach is practically perfect. The metric distance, averaged over all profiles and all incident angles, is $D = 0.05$. The worst distance obtained was only $D = 0.12$ [for the high coverage, incident angle 60 mrad, (00) profile], which still means a very small discrepancy as is evident from the figure.

The excellent consistency for the single defect case proves that here the perturbation approach contains all important multiple scattering events and can certainly be used as a computational method for the quantitative analysis of diffuse scattering data. It has already been shown, using this approach, that the diffuse scattering caused by an adatom adsorbed on a surface depends strongly on the adsorption site, as well as on the height of the adsorbate above the topmost substrate layer.[10] Therefore, a quantitative structure analysis of randomly distributed adatoms (or molecules) based on diffuse RHEED data should be possible. It is evident from the above that a reliable and efficient RHEED theory for such purposes is available. This result should also be of relevance for the simulation of corresponding high-resolution REM images of single absorbed species on surfaces, however, note that the present investigation concerns reflected *intensities* while in the REM case the phase of the reflected amplitude is also important.

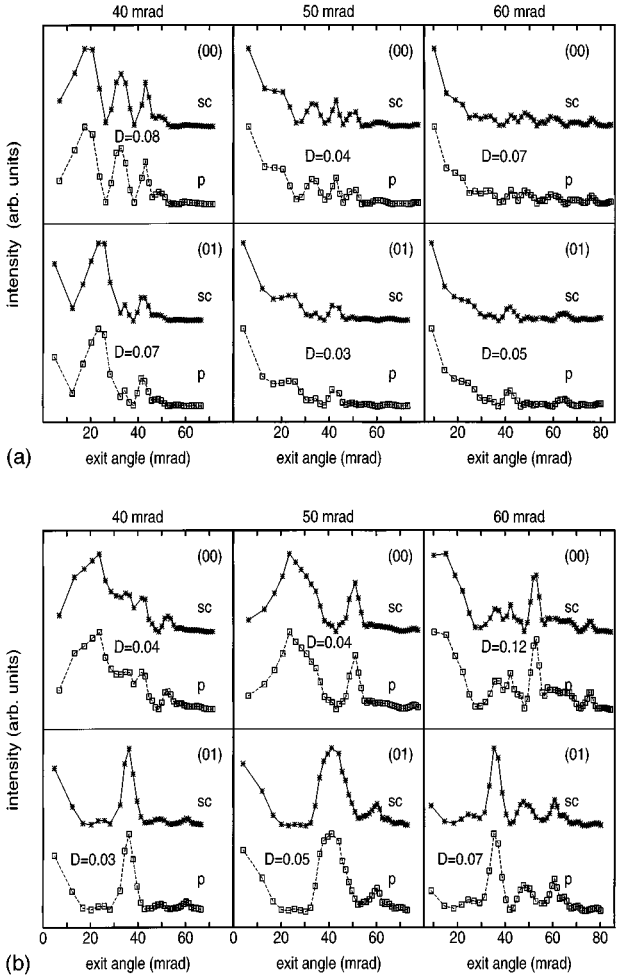

FIG. 7. Profiles from rigorous supercell (sc) calculations and perturbation theory (p) related to the single defect case for Si(100) and various incident angles. The electron energy is 15 keV and the azimuth is $[0\bar{1}1]$. (a) One bilayer unit (per supercell) present upon the otherwise perfect surface. (b) One bilayer unit missing from the otherwise perfect surface. D denotes the metric distance between sc and p.

We attribute the remaining small differences in the results of both approaches to the artificial periodicity caused by the supercell. Although a length of 400 lattice units represents a quite large cell, there is still a small effect of the successive multiple scattering by the periodically arranged defects that is not considered in the perturbation treatment. In particular the periodicity may enhance these effects due to a corresponding artificial coherence in the diffuse scattering. For this reason the perturbation calculation is probably even more suitable for the description the diffuse scattering caused by a single defect or statistically independent defects than the supercell approach.

C. Random configuration at high coverages

In the foregoing section the potential of the perturbation approach to obtain structural data for adatoms by a quantitative evaluation of the corresponding diffuse RHEED data was demonstrated. At low coverages an evaluation based on calculations for the single defect case should be realistic as

long as the adatoms are randomly distributed (but on equivalent geometric sites). In this case the experimental data can be assumed to be the incoherent superposition of the diffuse scattering caused by the individual adatoms such that the shape of the diffuse scattering distribution is the same as for a single atom situated on the periodic substrate. For higher coverages, this is no longer valid for two reasons. Firstly, the distribution of the atoms on the surface would result in further structure in the diffuse scattering because of long-range correlations between the adatoms. Secondly, at high coverages there is increasing probability of successive multiple scattering by the adatoms. The first of the two problems results effectively in an unknown structure factor. In analysis of diffuse LEED data this factor is eliminated by evaluating the logarithmic derivative of the intensities with respect of the energy instead of the original data.[28] In RHEED, the corresponding data manipulations would refer to the incident glancing angle instead of the electron energy. In contrast to the LEED case, the second problem of the multiple scattering by the adatoms is easily solved in RHEED as long as the adatoms are randomly distributed, i.e., the correlation lengths involved are small. As analytically discussed in Sec. V and demonstrated numerically in Fig. 2(b) the second-order diffuse scattering is then negligible [see also Eq. (28)] and the multiple diffuse scattering within the disordered layer is then essentially determined only by its periodic potential part, which is independent of the detailed distribution of the adatoms. The corresponding numerical treatment requires only the inclusion of a further *ordered* layer whose potential is multiplied by the coverage θ. If θ is unknown, it can easily be treated as a fitting parameter and thus may even be measured.

It follows from the above argument that structure information for a disordered adlayer can even be obtained *without* evaluating the diffuse scattering. If the multiple scattering processes within the disordered layer are governed by its periodic potential part, then the rocking curves of the sharp reflections on the Laue circle should be determined by this potential as well. We have checked this by calculating rocking curves of the sharp reflections using the supercell program and the same random configuration of half coverage as in Fig. 2(b) where the bilayer units were distributed within the supercell by means of a random number generator. These curves were compared with rocking curves calculated with the periodic potential part of this structure, i.e., the potential of an ordered bilayer multiplied by the coverage 0.5. The comparison of this periodic potential approximation with the supercell calculation is illustrated in Fig. 8 by means of the rocking curves of the (00) and (01) reflection. The metric distance between the curves from the two calculations is 0.02 for both reflections. This degree of agreement has to be regarded as excellent. Furthermore, the corresponding rocking curves for the monatomically flat surface (coverage 1.0) are plotted as well. They differ substantially from the curves for the disordered surface of coverage 0.5. This demonstrates clearly the sensitivity of the rocking curves with regard to the disordered layers. The above comparison shows that the scattering of such a complicated structure can be described in an astonishingly simple way. This result is especially important from the experimental point of view. Compared to the accurate measurement of diffuse scattering, accurate rocking curve data can be obtained very easily. In fact, the experimental procedure is the same as for an ordered surface. It is finally noted that these results are transferable to materials with stronger scattering potentials. In connection with a recent experimental work, a similar comparison between rocking curves from supercell calculations and the periodic potential approximation has been carried out for the strong scatterer Pt.[29] There, it is demonstrated furthermore that the periodic potential approximation even retains the sensitivity of the curves with regard to small variations of structural parameters like relaxations of the top interlayer distance.

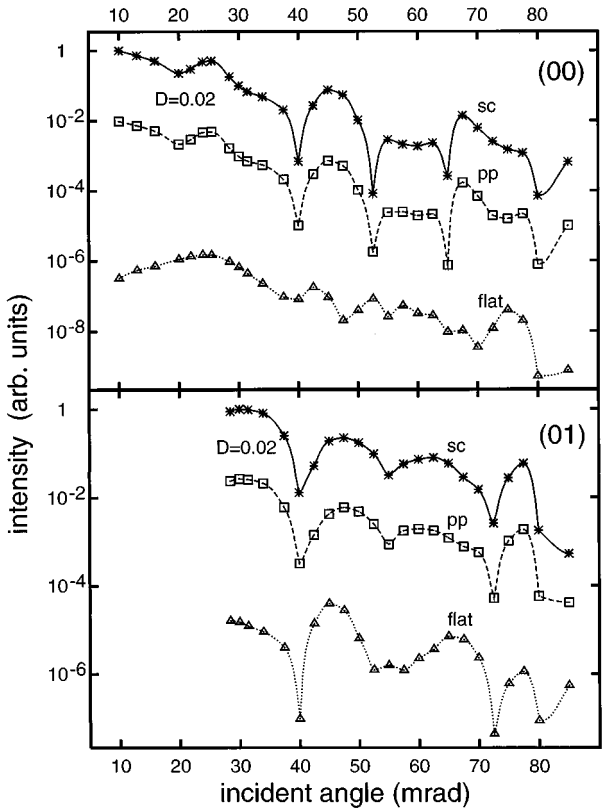

FIG. 8. Rocking curves of the sharp reflections (00) and (01) from rigorous supercell (sc) calculations and periodic potential approximation (pp) in case of a random configuration of steps at coverage 0.5 for Si(100) [same configuration as in Fig. 2(b)]. The electron energy is 15 keV and the azimuth is [011]. D denotes the metric distance between sc and pp. For the purpose of comparison the curves for the perfect periodic surface (flat) are also plotted.

VII. DIFFRACTION CONDITIONS

For large correlation lengths L (and high defect concentrations) perturbation theory cannot generally be expected to work properly and its applicability depends on the diffraction condition. For the Si(100) step arrays considered here, this occurs when the correlation length is greater than about 20 LU (see Fig. 3), the threshold value above which the dependence of the relative error on L disappears. Additional supercell-perturbation calculations for incident angles up to 155 mrad have been carried out for a step array with $L=80$ LU and coverage $\theta=0.8$. This length is, on the one

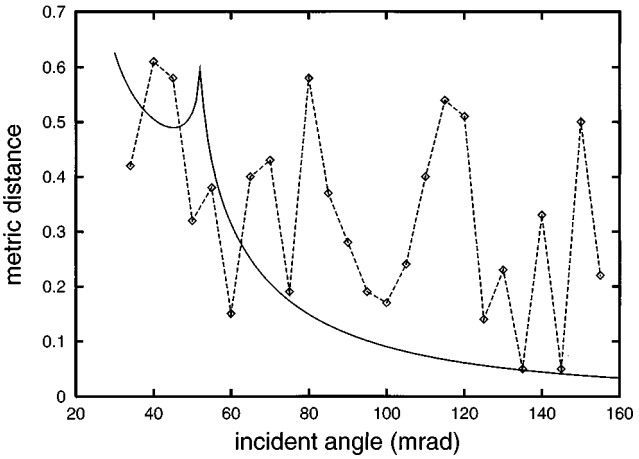

FIG. 9. Dashed curve: Metric distance between profiles from perturbation theory and supercell approach, calculated for a regular step array with coverage $\theta=0.8$, vs incidence angle. Solid curve: error from the analytical expression. The general trend of a decrease of the metric distance with increasing angle is superimposed by strong fluctuations. For further explanation, see text.

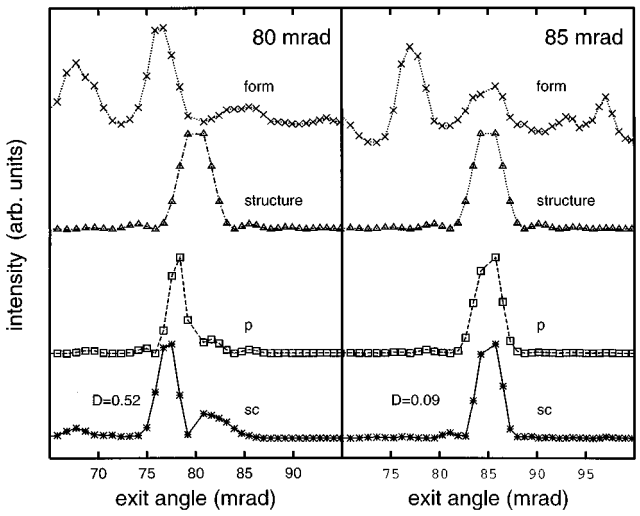

FIG. 10. Example for the correlation of the quality of perturbation theory with the overlap of the ''dynamical'' form factor and the structure factor. The illustration is for a regular bilayer step array on Si(100) with the coverage 0.8 and an upper terrace length of 80 lattice units. The (00) diffuse streak is shown. The electron energy is 15 keV and the azimuth is [011]. Left panel: For 80-mrad incident angle the structure factor profile (structure) coincides with a minimum of the ''dynamical'' form factor profile (form) such that the overlap is low. The agreement between supercell (sc) and perturbation calculation (p) is only moderate. Right panel: For 80-mrad incident angle the structure factor profile coincides with a maximum of the ''dynamical'' form factor profile. such that the overlap is high. Here, the agreement between supercell and perturbation theory is very good. D denotes the metric distance between sc and p.

hand, large enough to lie beyond the above-threshold value and, on the other hand, still low enough to produce reasonably large diffuse profile widths (3–12 mrad, dependent on the incident angle for these additional calculations).

Figure 9 shows how the metric distance D, averaged over the (00) and (01) profiles, between supercell and perturbation profiles, depends on the incident angle. Furthermore, the angular dependence of the error $\epsilon_{\mathbf{g}}$ from the analytical estimation is plotted for the averaged $k_{\mathbf{g}}$ involved. Except for strong fluctuations (see below) the metric distance, as well as $\epsilon_{\mathbf{g}}$, decreases gradually with the incident angle. This general trend is due to the fact that on increasing the incident angle the waves involved propagate more and more obliquely to the disordered layers, thus decreasing the probability of second-order diffuse scattering within the disordered layers. The figure also shows that this general trend is superimposed by strong fluctuations in the angular dependence of D. Indeed, within the whole angular range situations can be found where the metric distance is very low (even at low angles) or very high. A close inspection of the dynamical form factor and the streak profiles [see Eq. (10)] revealed that these cases follow systematic rules. The metric distance for a profile turned out to be low if the strong Fourier coefficients of the corresponding structure factor S (kinematical profile) correspond to exit angles where the ''dynamical'' form factor f_{dyn} exhibits a maximum, i.e., if the overlap of f_{dyn} and S is strong. On the other hand, the agreement with the rigorous approach was bad if the strong coefficients of the structure factor correspond to exit angles near a minimum or a steep shoulder of f_{dyn}, i.e., if the overlap of f_{dyn} and S is weak. This is demonstrated in Fig. 10 for the (00) profile at 80- and 85-mrad incident angle, respectively. In the first case, the structure factor profile coincides with a minimum of the ''dynamical'' form factor and the agreement with the supercell calculation is only moderate ($D=0.52$). For the angle 85 mrad, however, the overlap is strong and the agreement is very good ($D=0.09$).

In Fig. 11 the metric distance D and the inverse of the normalized overlap integral

$$I_{\mathrm{ov}}=\int ds_x|f_{\mathrm{dyn}}(s_x)S(s_x)|^2 \Big/ \int ds_x|f_{\mathrm{dyn}}(s_x)|^2 \quad (30)$$

between the dynamical factor and the structure factor are plotted versus the incident angle for both the (00) and the (01) profiles. The overlap I_{ov} defined in this way is nothing but the (normalized) integrated profile from the perturbation scheme. The metric distance and the inverse overlap I_{ov}^{-1} as functions of the incident angles develop qualitatively in a quite similar manner. Virtually all maxima (minima) in the incident angle dependence of the overlap coincide with maxima (minima) of D. Further examples for streak profiles under the present conditions can be found in Fig. 4. It is noted in particular that for the favorable situations, where the angular dependence of the metric distance has a minimum, the corresponding distances are very low (around 0.15) such that perturbation theory works very well in those cases. It is therefore desirable to understand the physics behind this result.

A large overlap appears under diffraction conditions at which a large portion of the structure factor profile corresponds to exit angles where the dynamical factor f_{dyn} is large compared to the other exit angles. Physically, at such diffraction conditions, the diffuse waves in the crystal correspond-

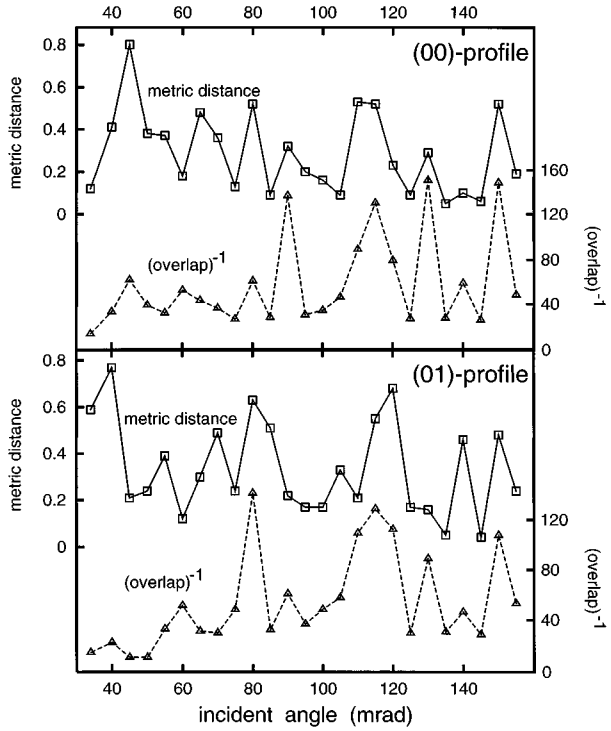

FIG. 11. Correlation of the quality of perturbation theory with the overlap of ''dynamical'' form factor and structure factor. For further explanation, see text.

ing to the profile considered are excited strongly due to scattering processes via the periodic part of the potential (couplings D_f) in Eq. (8). These processes are, however, included exactly in the perturbation approach. In other words, the profile tends to be dominated by the effect of those diffraction processes, which are described accurately by perturbation theory. These are therefore cases where the method works adequately.

This can be also seen more formally from Eq. (8). Integration of this equation yields an outgoing solution of the type

$$y(\mathbf{s},z)=\frac{1}{2ik_{\mathbf{s}}}\int dz_2\ \exp(ik_{\mathbf{s}}|z-z_2|) \times[S_0(\mathbf{s},z_2)+D_f(\mathbf{s},z_2)+C_2(\mathbf{s},z_2)]. \quad (31)$$

If f_{dyn} is large at the exit angle belonging to the scattering vector **s**, the wave couplings S_0 and *especially* D_f *via the periodic potential part* contribute strongly to the corresponding amplitude y. Then, in the above sum of couplings C_2 is small compared to the other terms so the perturbation treatment works well. On the other hand, if f_{dyn} is small, S_0 and D_f contribute only weakly to the amplitude. In this unfavorable case virtually the whole amplitude is determined by the couplings neglected in perturbation theory. Consequently, this approach becomes unrealistic under such conditions.

Fortunately, it should be possible to identify the favorable diffraction conditions experimentally because the modulations of the dynamical factor are by its definition related to the modulations of the diffuse broad background observed in an experimental diffraction pattern [see also Refs. 10–12]. Hence, it should be possible to identify regions of favorable exit angles directly on the fluorescent RHEED screen and thus to select the diffraction conditions at which an effective evaluation can be carried out.

VIII. CONCLUSION

In this work we have carried out a thorough investigation of the conditions in which the perturbation theory of diffuse RHEED (Ref. 10) can be applied to the evaluation of experimental diffuse scattering data from occupational disorder (e.g., steps). The approach treats the diffuse scattering as a transition between two-dimensional Bloch waves in the periodic part of the potential and can be reduced to the calculation of conventional structure factors, which are then multiplied by ''dynamical'' form factors. Such an investigation is desirable because the perturbation method is at present the only approach within the scope of a dynamical theory that solves the configuration problem (the presence of many statistically varying disorder configurations).

To this end a comparison between calculations based on the perturbation approach and rigorous supercell calculations was carried out for bilayer steps upon an unreconstructed Si(100) surface. The general validity of the results is supported by some analytical arguments. An analytical expression [Eq. (24)] was found, which roughly predicts for a given material and structure the order of magnitude of the error made in the perturbation approach compared to an exact treatment. Most generally, the quality of the perturbation approach is determined by the strength of the atomic scattering potential in the disordered layers and the concentration of defects with regard to the periodic part of the potential.

In addition, a key structural parameter that determines the applicability of the perturbation method is the correlation length describing the disorder along the incident beam azimuth. If this length is sufficiently small, the perturbation method works well and is independent of the diffraction condition. For this case, it was further demonstrated by means of a random distribution of defects at half coverage (where the defect concentration is particularly high) that, besides the diffuse scattering, the rocking curves of integer beams from both the supercell and the perturbation approach are very similar. Hence, structural information about the periodic part of a *strongly disordered surface* can be extracted from rocking curves of the principal reflections using the same evaluation methods as for ordered surfaces.

For large correlation lengths (and non-negligible defect concentrations) the applicability of perturbation theory depends on the diffraction condition. As a general trend the applicability is favored if the polar exit angles of those integral-order waves, which are strongly excited, are large. For these cases the corresponding diffuse waves (concentrated around the integral-order ones) travel mainly obliquely to the disordered layers thus decreasing the probability of second-order diffuse scattering, which is not included in the perturbation scheme. Moreover, diffraction conditions could be deduced at which perturbation theory produces reliable results for large correlation lengths, even if the glancing angles involved are low. Under such conditions the diffuse intensity of a profile is dominated by those scattering processes that are included exactly in the perturbation approach.

It should be possible to identify the corresponding favorable exit angles experimentally by means of the modulations in the diffuse background distribution. The above results show that for a quite general range of disordered structures, conditions can be found where perturbation theory can be applied. Finally it is believed that our investigation will contribute to an effective dynamical evaluation of diffuse RHEED data in the future.

ACKNOWLEDGMENTS

The authors would like to thank Professor J.L. Beeby for his support of this work and for stimulating discussions. It is further a pleasure to thank Dr. Z. Mitura for many helpful and stimulating discussions. This work was supported by the UK Engineering and Physical Sciences Research Council, NATO and the European Commission. U.K. thanks the European Commission for financial support.

APPENDIX: DERIVATION OF THE RELATIVE ERROR

In order to obtain an estimate of the error $\epsilon_{\mathbf{g}}$, we need to evaluate the magnitude of the BZ integral,

$$I_{\mathrm{BZ}\ \mathbf{g}}=\int_{\mathrm{BZ}\ \mathbf{g}}\frac{d\mathbf{s}'}{k'_{\mathbf{s}'}}|S'(\mathbf{s}')|^2\sum_{\mathbf{g}'_2} I'(\mathbf{g}'_2,\mathbf{s}',0) \tag{A1}$$

in Eq. (21).

For $|S'(\mathbf{s}')|^2$ we assume the following simplified expression within the Brillouin zone BZ $\mathbf{g}$:

$$|S'(\mathbf{s}')|^2=|S(s'_x,s'_y)|^2=\begin{cases} L_xL_y/\Omega & \text{if } (|s'_x-g'_x|<\sqrt{\Omega}/L_x \text{ and } |s'_y-g'_y|<\sqrt{\Omega}/L_y) \\ 0 & \text{otherwise .}\end{cases} \tag{A2}$$

Here, the indices x and y refer to the vector components along (x) and perpendicular (y) to the incident beam azimuth, respectively. (In the main text L_x is denoted simply as L because the final result will not depend on L_y.) This expression accounts for the fact that the strongest diffuse Fourier coefficients are typically concentrated within regions of size $2\pi/L_i$ around the reciprocal lattice vectors $\mathbf{g}$ and it also obeys the conservation law (here in its normalized form) given by Eq. (3).

For the z' integral $I'(\mathbf{g}'_2,\mathbf{s}',0)$ [Eq. (22)] it is assumed that the coefficients $w(\mathbf{g}'_2,z)$ of the strong waves only vary slowly with $z=z'\sqrt{\Lambda}$ such that we consider $w(\mathbf{g}'_2,z)=w_{\mathbf{g}'_2}$ as constant. This means physically that the diffuse scattering is mainly excited by waves with a small $|k_{\mathbf{g}_2}|$ (which are usually the strongest). This crude assumption tends to overestimate I' due to the neglect of oscillations (propagating) or an exponential decrease (evanescent) under the integral and thus contributes to a pessimistic estimation of the error $\epsilon_{\mathbf{g}}$. The z' dependence of the atomic potential coefficient is represented by $v(\mathbf{s}'-\mathbf{g}'_2,z')=v'(\mathbf{s}'-\mathbf{g}'_2,0)\exp(-|z'|/\mu')$. This corresponds approximately (for the zeroth coefficient exactly) to a screened Coulomb potential with the screening length μ'. These assumptions lead to

$$I'(\mathbf{g}'_2,\mathbf{s}',0)=w_{\mathbf{g}'_2}v(\mathbf{s}'-\mathbf{g}'_2,0)\frac{2}{\mu'^{-1}-ik'_{\mathbf{s}'}}. \tag{A3}$$

It is now possible to easily estimate the BZ integral (A1). Because $k'_{\mathbf{s}'}$ as a function of s_y varies little within the BZ, we neglect this dependence for the integration over s'_y and put $k'_{\mathbf{s}'}(s'_x,s'_y)=k'_{\mathbf{s}'}(s'_x,g'_y)$. In the case of the integration over s'_x the strong variation of $k'_{\mathbf{s}'}$ with s'_x has to be fully included. Depending on whether the nonzero diffuse Fourier coefficients within the BZ [see Eq. (A2)] correspond to evanescent, propagating, or both types of waves, different forms of (A3) have to be used under the integral.

Insertion of $I_{\mathrm{BZ}\ \mathbf{g}}$, obtained in this way, into the general expression (21) yields for the relative error $\epsilon_{\mathbf{g}}$:

$$\epsilon_{\mathbf{g}}=\sqrt{2}\sqrt{\theta(1-\theta)}\frac{2m}{\hbar^2}V_{\mathrm{av}}(0)\Lambda U(k'_{\mathbf{g}},\mu')\times\left|\frac{1}{w_{\mathbf{g}}}\sum_{\mathbf{g}'_2} v'(\mathbf{g}'-\mathbf{g}'_2,0)w_{\mathbf{g}'_2}\right|, \tag{A4}$$

where

$$U=\begin{cases} U_1 & \text{if } |k'_{\mathbf{g}}|^2\leqslant 1 \\ U_2 & \text{if } k'^2_{\mathbf{g}}>1 \\ U_3 & \text{if } -k'^2_{\mathbf{g}}>1. \end{cases} \tag{A5}$$

U_1 applies if in BZ $\mathbf{g}$ both propagating and evanescent waves appear. U_2 holds if all diffuse waves are propagating; U_3 if all diffuse waves in the BZ are evanescent. It can be shown that for equal $|k_{\mathbf{g}}|$, the shapes of U_2 and U_3 as a function of Λ are very similar but U_2 is always slightly greater than U_3. For reasons of simplicity we therefore always use U_2 for the case $\Lambda>|k_{\mathbf{g}}|^{-2}$ (i.e., the more ''pessimistic'' function). We proceed correspondingly in the case of U_1. For different signs of $k^2_{\mathbf{g}}$ and the same $|k_{\mathbf{g}}|$, U_1 as a function of Λ has a similar form, with slightly greater values for the case $k^2_{\mathbf{g}}>0$. Hence, again for simplicity we always use $k^2_{\mathbf{g}}>0$ in U_1 for a given $|k_{\mathbf{g}}|$ such that eventually the expressions (26) and (27) are used in the final result.

It follows from the general relation $|\psi|^2\leqslant 1$ for the wave function that

$$\left|\sum_{\mathbf{g}'_2} v'(\mathbf{g}'-\mathbf{g}'_2,0)w_{\mathbf{g}'_2}\right|\leqslant 1. \tag{A6}$$

We now replace the sum in (A4) by 1; i.e., we assume that the right- and the left-hand sides of (A6) are equal. This

overestimate of the left-hand side (which would hold for the case that $w_{\mathbf{g}}$ only is excited with $|w_{\mathbf{g}}|=1$) is compensated by using $|w_{\mathbf{g}}|=1$ in the denominator as well. Furthermore, the laterally averaged potential $V_{\mathrm{av}}(0)$ at $z=0$ is expressed in terms of the volume average potential V_0 of the material (interlayer distance d) by $V_{\mathrm{av}}(0)\approx(d/2\mu)V_0$. These substitutions into (A4) finally result in the estimate Eq. (24) for the error $\epsilon_{\mathbf{g}}$.

*Present address: Fachbereich Physik, Universität Osnabrück, D-49069 Osnabrück, Germany.

†Present address: Environmental Molecular Sciences Laboratory, Pacific Northwest Laboratory, P. O. Box 999, Richland, Washington 99352.

[1] M. Stock and G. Meyer-Ehmsen, Surf. Sci. **226**, L59 (1990).

[2] U. Korte and G. Meyer-Ehmsen, Surf. Sci. **271**, 616 (1992).

[3] J. M. McCoy, U. Korte, P. A. Maksym, and G. Meyer-Ehmsen, Phys. Rev. B **48**, 4721 (1993).

[4] T. Hanada, S. Ino, and H. Daimon, Surf. Sci. **313**, 143 (1994).

[5] C. S. Lent and P. I. Cohen, Surf. Sci. **139**, 121 (1984); P. R. Pukite, C. S. Lent, and P. I. Cohen, *ibid.* **161**, 39 (1985).

[6] L.-M. Peng and J. M. Cowley, Acta Crystallogr. Sec. A **42**, 545 (1986).

[7] J. M. McCoy and P. A. Maksym, Surf. Sci. **297**, 113 (1993)

[8] P. A. Maksym and J. L. Beeby, Surf. Sci. **110**, 423 (1981).

[9] A. Ichimiya, Jpn. J. Appl. Phys. **22**, 176 (1983).

[10] U. Korte and G. Meyer-Ehmsen, Phys. Rev. B **48**, 8345 (1993).

[11] U. Korte and G. Meyer-Ehmsen, Surf. Sci. **298**, 299 (1993).

[12] U. Korte and G. Meyer-Ehmsen, Surf. Sci. **277**, 616 (1992).

[13] J. L. Beeby, Surf. Sci. **298**, 307 (1993).

[14] P. R. Pukite, P. I. Cohen, and S. Batra, in *Reflection High-Energy Electron Diffraction and Reflection Electron Imaging of Surfaces*, edited by P. K. Larsen and P. J. Dobson (Plenum, New York, 1988), p. 427.

[15] P. A. Maksym, in *Thin Film Growth Techniques for Low Dimensional Structures*, edited by R. F. C Farrow, S. S. P. Parkin, P. J. Dobson, J. H. Neave, and A. S. Arrott (Plenum, New York, 1987), p. 95.

[16] A. Messiah, *Quantum Mechanics* (North-Holland, Amsterdam, 1965), Vol. II, p. 825.

[17] G. Meyer-Ehmsen, Surf. Sci. **219**, 177 (1989).

[18] P. A. Doyle and P. S. Turner, Acta Crystallogr. Sec. A **24**, 390 (1968).

[19] J. M. McCoy, U. Korte, P. A. Maksym, and G. Meyer-Ehmsen, Surf. Sci. **306**, 247 (1994).

[20] J. V. Ashby, N. Norton, and P. A. Maksym, Surf. Sci. **175**, 604 (1986).

[21] Recent program optimizations have resulted in a cpu time saving by a factor of about 4 and a better than cubic scaling of cpu time with number of waves; P. A. Maksym, Surf. Rev. Lett. (to be published).

[22] A. Howie and R. M. Stern, Z. Naturforsch. Teil A **27**, 382 (1972).

[23] P. H. Dederichs, in *Solid State Physics*, edited by H. Ehrenreich, F. Seitz, and D. Turnbull (Academic, New York, 1972), Vol. 27, p. 135.

[24] S. L. Dudarev, L.-M. Peng, and M. J. Whelan, Phys. Rev. B **48**, 13 408 (1993).

[25] H. Marten and G. Meyer-Ehmsen, Surf. Sci. **151**, 570 (1985).

[26] In this sense we here deviate from our original convention that C_2 in Eq. (8) refers only to the diffuse wave field ψ_d. However, the formal split of the wave field into ψ_p and ψ_d is rather used for technical reasons and Eq. (8) can be extended to the integer waves without problems.

[27] The value 0.43 Å for the range μ of the atomic potential was obtained by comparing the first two terms of the Taylor expansion of the atomic scattering amplitude (with regard to the modulus of the scattering vector) of a screened Coulomb potential and the exponential representation by the Doyle-Turner coefficients (Ref. 18).

[28] K. Heinz, K. Müller, W. Popp, and H. Lindner, Surf. Sci. **173**, 366 (1986).

[29] S. Schwegmann, W. Tappe, and U. Korte, Surf. Sci. **334**, 55 (1995).