Statistical error analysis of surface-structure parameters determined by low-energy electron and positron diffraction: Data errors

C. B. Duke, A. Lazarides, A. Paton, and Y. R. Wang
Xerox Wilson Center of Research & Technology, 800 Phillips Road, 0114-38D, Webster, New York 14580
(Received 22 February 1995)

An error-analysis procedure that gives statistically significant error estimates for surface-structure parameters extracted from analyses of measured low-energy electron and positron diffraction (LEED and LEPD) intensities is proposed. This procedure is applied to a surface-structure analysis of Cu(100) in which experimental data are simulated by adding Gaussian-distributed random errors to the calculated intensities for relaxed surface structures. Quantitative expressions for the variances in the surface-structural parameters are given and shown to obey the expected scaling laws for Gaussian errors in the experimental data. The procedure is shown to describe rigorously parameter errors in the limit that the errors in the measured intensities are described by uncorrelated Gaussian statistics. The analysis is valid for structure determinations that are of sufficient quality to admit errors that have magnitudes within the region of convergence of a linear theory that relates perturbations of diffracted intensities to perturbations in structural parameters. It is compared with previously proposed error-estimation techniques used in LEED, LEPD, and x-ray intensity analyses.

I. INTRODUCTION

In this paper, we develop and apply a methodology to obtain statistically significant error estimates for surface-structural parameters extracted from dynamical analyses of low-energy electron or positron diffraction (LEED) and LEPD) intensities. This methodology is based on applying techniques, which are well known in discrete inverse theory,[1] to the determination of surface structures by LEED and LEPD.[2–4] We confine our attention to uncertainties in the structural parameters caused by random errors in the measured intensities, primarily because the analysis is greatly simplified in this case yet affords considerable insight.

Our effort on this problem is motivated by the confluence of three factors. First, in our prior comparisons of LEED and LEPD structure analyses for GaAs(110) and InP(110), it became evident that current approaches to the estimation of errors in the resulting structural parameters fail to account satisfactorily for systematic differences between electron and positron diffraction.[5] Second, other similar cases have occurred in the literature in which LEED yielded structures different from those obtained using other techniques[6,7] or different LEED analyses of the same system[8–10] yielded different structures. Taken together with the current trend toward determining complex structures exhibiting small atomic displacements from a reference geometry,[4] these observations suggest that a closer examination of error estimates in LEED and LEPD intensity analyses would be productive. Third and finally, mathematical techniques adapted for error analyses in inverse problems have been developed in a variety of contexts[1] and offer promise for surface crystallography as well. Therefore, we have both the motivation and the means for another look at this old problem.

We proceed in three steps. In Sec. II, we define the analysis methodology. In Sec. III, this methodology is applied to two model problems on Cu(110) to show how it works in practice. Section IV is devoted to its comparison with other methods described in the literature. We close with a synopsis.

II. ANALYSIS PROCEDURE

Consider the extraction of one or more surface-structural parameters in a case in which the only errors are those due to random statistical fluctuations in the measured intensities (due, e.g., to Poisson statistics of a digital electron or positron detector[5,11]). The geometry of a LEED or LEPD experiment[12,13] is illustrated schematically in Fig. 1. An incident beam with momen-

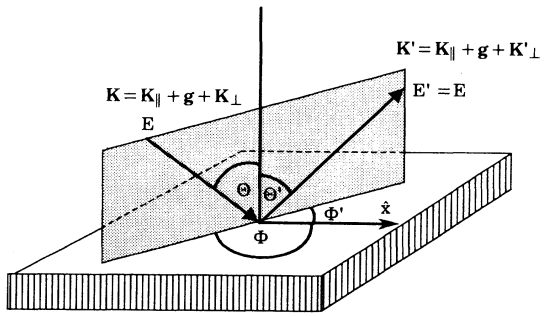

FIG. 1. Schematic diagram of LEED or LEPD experimental in which θ is the incident polar angle, Φ the incident azimuthal angle, E the energy, and $\mathbf{g}(hk)$ the reciprocal vectors of the two-dimensional surface bravais net. In general, the plane defined by the surface normal and the exit beam differs from that defined by the surface normal and the incident beam.

tum parallel to the surface of

$$\mathbf{k}_{\parallel}=k_x\mathbf{i}+k_y\mathbf{j} , \tag{1a}$$

$$k_x=k(E)\sin\theta\cos\Phi , \tag{1b}$$

$$k_y=k(E)\sin\theta\sin\Phi , \tag{1c}$$

$$k^2(E)=2mE/\hbar^2 , \tag{1d}$$

is diffracted into a series of beams

$$\mathbf{k}'_{\parallel}(hk)=\mathbf{k}_{\parallel}+\mathbf{g}(hk) , \tag{2a}$$

in which $\mathbf{g}(hk)$ are the reciprocal-lattice vectors of the surface Bravais net. E designates the energy of the incident electron (positron); m is its mass; and $\hbar$ is Planck's constant. Only beams with real values of

$$|\mathbf{k}'_{\perp}|=[k^2(E)-|\mathbf{k}'_{\parallel}|^2]^{1/2} \tag{2b}$$

emerge from the crystal: a condition which restricts the values of (hk) in Eq. (2a) to those which satisfy

$$2\mathbf{k}_{\parallel}\cdot\mathbf{g}(hk)+g^2(hk)<2mE\cos^2\theta/\hbar^2 . \tag{2c}$$

Typically the diffracted beams either disappear or become sensitive primarily to bulk properties for

$$E>E_m\sim 300\text{ eV} . \tag{3}$$

Thus, the total body of intensity data available for surface-structure determination consists of the intensities of all of the allowed [from Eq. (2c)] (hk) beams as functions of E, θ, and Φ for $E<E_m$, i.e.,

$$\{I_{hk}(E,\theta,\Phi),\theta<\pi,\Phi<2\pi,E<E_m\} . \tag{4}$$

Surface structures are determined by comparing these data with model calculations in which the structural parameters of the uppermost few atomic layers are treated as adjustable parameters to be extracted from fitting a selected sample of the potentially available database defined by (4). This fitting is typically done using a reliability ("R") factor to define an optimal geometry[2,4–6,8–11] (associated with the minimum value of R). Many different R factors have been introduced in the literature.[14–21] Use of these R factors has been motivated by various rationales (e.g., analogy with x-ray-diffraction,[22] economy,[21,22] the nature of the theory of electron diffraction,[17,18] etc.). These R factors have not, however, been selected on the basis of a systematic statistical consideration of the consequences of errors in the input data and models on the values extracted for the structural parameters. Our goal in this section is to utilize discrete inverse theory[1] to define an R factor whose minimum value and curvature about that minimum specify precisely the "best" estimates of the structural parameters and the uncertainties therein obtained, by varying these parameters to achieve the best fit to a selected sample of the potentially available measured intensities.

The basis for our calculation is the analysis by Menke[1] of a model in which the theory of LEED (LEPD) is perfect, all parameters in the model, except the identified structural parameters, are known exactly, and the errors in the measured intensities, $I_{hk}(\theta,\Phi,E)$, are random variables characterized by a Gaussian distribution, with zero mean value, over a finite grid of values $\{h, k, \theta_i, \Phi_i, E_i\}$. If the data set is sufficiently large (has sufficient sensitivity to the structural parameters) as to lead to parameter estimates whose errors are within the convergence region of a linear LEED theory, then the estimates of the structural parameters, $\{a_i\}$, themselves form Gaussian distributions, with means and variances which are determined by the statistics that characterize the Gaussian intensity errors. The probability distribution for an estimate, a_{est}, of a single structural parameter is thus

$$P(a_{\text{est}})=[(2\pi)^{1/2}\sigma]^{-1}\exp[-(a_{\text{est}}-\langle a\rangle)^2/2\sigma^2] , \tag{5a}$$

in which $\langle a\rangle$ is the mean of the Gaussian distribution,

$$\langle a\rangle=\int aP(a)da , \tag{5b}$$

and σ^2 is its variance

$$\sigma^2=\int(a_{\text{est}}-\langle a\rangle)^2P(a)da . \tag{5c}$$

In the case of multiple structural parameters (addressed more fully below), the parameter error distributions are correlated and are described by a multivariate Gaussian distribution, whose complete specification requires not only variance, but also covariance information. Estimation theory provides a general expression for the parameter estimation error covariance in terms of the covariance of the intensity errors. Knowledge of the parameter variances, σ_i^2, allows the assignment of confidence levels to parameter ranges around specific estimates. An estimated value, $a_{i,\text{est}}$, of the parameter represented by a_i obtained from a given analysis will lie within $\pm\sigma_i$ of the mean $\langle a_i\rangle$ of the distribution with probability 0.68 and within $\pm 2\sigma_i$ of the mean with probability 0.95. Thus, if the measurement error distribution is unbiased (has mean 0), one can expect with 95% confidence that the true value of the parameter whose estimate is $a_{i,\text{est}}$ will lie within the interval $a_{i,\text{est}}\pm 2\sigma_i$. Hence, the analysis produces a statistically significant measure of the deviation of estimated values from true values.

In the case in which intensity errors can be described as uncorrelated and of constant variance, estimates of the variances σ_i^2 as well as estimates of the parameters can be determined directly from the fitting procedure if the proper R factor is utilized to compare the measured and calculated intensities.[1] The analysis proceeds by defining the Gaussian statistics R factor, R_g, via

$$R_g(a)=\frac{1}{N-M}\sum_{\alpha}\sum_{hk}\sum_{i}[I_{hk}^{\text{th}}(a)-I_{hk}^{\text{exp}}]^2 , \tag{6a}$$

$$I_{hk}^{\text{th}}(a)=I_{hk}^{\text{th}}(\theta_\alpha,\Phi_\alpha,E_i,a) , \tag{6b}$$

$$I_{hk}^{\text{exp}}=I_{hk}^{\text{exp}}(\theta_\alpha,\Phi_\alpha,E_i) , \tag{6c}$$

in which we utilize N to indicate the total number of data points in the sums, M to indicate the number of parameters being estimated, $\alpha=\{\theta_\alpha,\Phi_\alpha\}$ to specify the grid of incident-beam directional parameters, (hk) to specify the beam indices for a given incident-beam direction, and E_i to specify the energy grid along the diffracted beams for a

given set of $\{\alpha, (hk)\}$ parameters. The estimates a are the values that minimize $R_g(a)$ and the variance is obtained from the minimum value of $R_g(a)$ and the second partial derivatives evaluated at the minimum.

If the theory predicted the absolute normalization of the intensities, then Eqs. (6) would suffice, but such is not the case.[23] Thus, Eq. (6a) is replaced by

$$R_g(a,c)=\frac{1}{F}\sum_{\alpha}\sum_{hk}\sum_{i}[cI_{hk}^{\rm th}(a)-I_{hk}^{\rm exp}]^2 , \tag{7a}$$

$$F=N-M-1 , \tag{7b}$$

in which we have introduced the required normalization constant c. This constant is determined by minimizing R_g with respect to it, giving

$$c=C_1/C_2 , \tag{7c}$$

$$C_1=\sum_{\alpha,hk,i}[I_{hk}^{\rm th}(\theta_\alpha,\Phi_\alpha,E_i,a)I_{hk}^{\rm exp}(\theta_\alpha,\ \Phi_\alpha,\ E_i)] , \tag{7d}$$

$$C_2=\sum_{\alpha,hk,i}[I_{hk}^{\rm th}(\theta_\alpha,\ \Phi_\alpha,\ E_i,\ a)]^2 . \tag{7e}$$

Equations (7) define the "Gaussian error" R factor, R_g, and thus specify our criterion for comparing calculated intensities $I_{hk}^{\rm th}(a)$ with experimental ones $I_{hk}^{\rm exp}$ when selecting structural parameters. The R_g minimizing parameter values, a_i, are the least-squares solution to the parameter estimation problem by virtue of the sum of squares form of R_g.

In the vicinity of the R-factor minimum, the R factor can be expressed as a second-order Taylor expansion about R_{g0}. Since the first derivative at the minimum is zero, the R factor expansion appears as

$$R_g(a)=R_{g0}+(a-a_{\rm est})^2\frac{1}{2}\frac{d^2R_g}{da^2}\bigg|_0 \tag{8a}$$

if the structural refinement involves only one variable parameter. If, furthermore, the errors in the various intensity data points are uncorrelated and of constant variance, and Eq. (7) is used to define R_g, then the parameter estimation variance (Ref. 1, pp. 58–60) is

$$\sigma_a^2=\frac{2R_{g0}}{F}\left[\frac{d^2R_g}{da^2}\right]_0^{-1} , \tag{8b}$$

where the subscript 0 is used to indicate evaluation of R_g or of its second derivative at the value of a which minimizes R_g. The quadratic form of the R-factor surface [Eq. (8a)] in the vicinity of the minimum follows from the assumption of a linear theory [in the case of LEED, a first order perturbation theory (e.g., the linear tensor LEED approximation[24]) relating intensity variations to the parameter variations], and the extent of the quadratic region corresponds to the convergence region of the linear theory around the R-factor minimizing structure. Equation (8b) follows from the single variable version of Menke's Eq. (3.52), which (with FR_g substituted for Menke's E) is

$$\sigma_a^2=\frac{2\sigma_e^2}{F}\left[\frac{d^2R_g}{da^2}\right]_0^{-1} , \tag{9}$$

where σ_e^2 is the variance of the zero-mean Gaussian distribution of errors in the measured intensities, the derivative is evaluated at the minimum-R structure, and F, the number of degrees of freedom in the theory/data fit, is the denominator in R_g. For constant variance data, R_g at its minimum (R_{g0}) is itself a random variable with a χ^2 distribution, F degrees of freedom, mean σ_e^2, and variance $2\sigma_e^4$. While R_{g0} assumes different values when calculated using different data sets, for a sufficiently large number of data points of moderate variance, the mean of R_{g0} is well approximated by a given realization. Hence,

$$\sigma_e^2\approx R_{g0} , \tag{10}$$

and Eq. (8b) follows. The R-factor curve thus yields not only an estimate, $a_{\rm est}$, of the structural parameter a (the minimum-R structure), but also the sensitivity information necessary to characterize the accuracy of the estimate.

As noted above, when multiple structural variables $a_1,\ldots a_N$ are to be extracted from the analysis of LEED or LEPD intensities, the estimates are correlated. The multivariable Gaussian distribution that describes the parameter estimates

$$P(\mathbf{a})=\frac{|\text{cov }\mathbf{a}|^{-1/2}}{(2\pi)^{N/2}}\exp[-\tfrac{1}{2}(\mathbf{a}-\langle\mathbf{a}\rangle)^T(\text{cov }\mathbf{a})^{-1}\times(\mathbf{a}-\langle\mathbf{a}\rangle)] , \tag{11a}$$

$$\mathbf{a}^T=(a_1,a_2,\ldots a_N) , \tag{11b}$$

is expressed in terms of the error covariance matrix, elements of which are defined as

$$[\text{cov }\mathbf{a}]_{ij}=\int da_1\ldots\int da_N[a_i-\langle a_i\rangle][a_j-\langle a_j\rangle]P(\mathbf{a}) , \tag{11c}$$

where $\mathbf{a}$ is a column vector consisting of the N structural variables. Equations (8) become

$$R_g(\mathbf{a})=R_{g0}+[\mathbf{a}-\mathbf{a}_{\rm est}]^T\left[\frac{1}{2}\frac{\partial^2R_g}{\partial\mathbf{a}^2}\right][\mathbf{a}-\mathbf{a}_{\rm est}] , \tag{12a}$$

and

$$[\text{cov }\mathbf{a}]=\frac{2R_{g0}}{F}\left[\frac{\partial^2R_g}{\partial\mathbf{a}^2}\right]_0^{-1} , \tag{12b}$$

where the bracketed derivatives of R_g are a matrix of partial second derivatives, known as the Hessian matrix,

$$\left[\frac{\partial^2R_g}{\partial\mathbf{a}^2}\right]=\begin{bmatrix}\frac{\partial^2R_g}{\partial a_1^2} & \frac{\partial^2R_g}{\partial a_1\partial a_2} & \cdots & \frac{\partial^2R_g}{\partial a_1\partial a_N}\\ \vdots & & & \\ \frac{\partial^2R_g}{\partial a_N\partial a_1} & \cdots & \cdots & \frac{\partial^2R_g}{\partial a_N^2}\end{bmatrix} . \tag{12c}$$

The analysis proceeds by finding an approximate minimum of the R_g surface and then evaluating the partial derivatives by fitting R factors for a grid of nearby structures to a quadratic form similar to that in Eq. (12a). The structure corresponding to the exact minimum of the R-factor function is determined, and the partial second derivative matrix is inverted to find [cov **a**] via Eq. (12b). The diagonal elements of the covariance matrix [defined by Eq. (11c)] are the variances, σ_{ai}^2, of the individual parameter estimates. The square root of a particular diagonal element is the standard deviation, σ_{ai}, of the estimate of the corresponding parameter. While the σ_{ai} indicate the size of the errors in the individual parameter estimates, the off-diagonal elements of the covariance matrix specify the level of correlation among errors in estimates of pairs of parameters. In the single variable parameter estimation problem, the standard deviation of the parameter estimate is determined by the scalar curvature of the R-factor function. Equation (12b) indicates that, unless the Hessian matrix is diagonal, the standard deviation of a particular parameter estimate in the multivariable estimation problem depends upon many elements of the Hessian matrix.

Appendix A presents the derivation of Eqs. (12) as an application of standard estimation theory results for covariance mapping. Appendix B describes the procedure for calculating partial derivatives of the R factor by fitting the R-factor surface to the appropriate form. These two Appendixes provide, therefore, the derivation of the equations used in the numerical analyses described in subsequent sections.

III. APPLICATION TO Cu(100)

A. Generation of simulated experimental intensities

To provide an illustrative example of the consequences of the analysis set forth in Sec. II, we consider a case which we[25] and others,[8–10,26–28] previously have analyzed: Cu(100). For a selected structure, we generate "experimental" intensities by adding to the calculated intensities for that structure random Gaussian errors, $e = I_0 e_0$, with standard deviation $\sigma_e = I_0 \sigma_0$ at each energy and angle. e_0 is a Gaussianly distributed random variable with zero mean and variance σ_0^2, and I_0 is a normalization constant expressed in terms of the average of the calculated intensities. We describe our simulated data in terms of the standard deviation of its errors expressed as a fraction of the average intensity, $I_{\rm av}$.

The calculated intensities were generated using our current version[29] of the Duke-Laramore-Beeby angular-momentum-representation multiple-scattering programs.[30] The potential for Cu is taken to be a self-consistent nonrelativistic potential,[31] which is reduced to muffin-tin form by overlap of nearest-neighbor potentials. The resulting muffin-tin radius is $r_{\rm MT} = 1.278$ Å, and the muffin potential is $V_{\rm MT} = 11.86$ eV. Slater exchange ($\alpha = 1$) is used to describe the scattering of the low-energy electrons from the charge density inside the muffin-tin radius. Vibrations of the Cu ion cores are described using a Debye model with $\theta_D = 330$ K, $T = 298$ K, and complex phase shifts.[25,32] Six phase shifts are used to calculate the diffracted intensities from a ten-layer slab of Cu. Scattering from the top six layers of the slab is treated exactly and that from deeper layers in a layer-by-layer fashion.[33] I_0 is in units of the average diffracted intensity over all angles and energies, $I_{\rm av}$. The e_0 are generated via the random number generator in the NAG library.[34]

To emulate a practical structure analysis,[8,9,29] we calculated the diffracted intensities of the (01), (11), (02), (12), (21), (22), (03), and (13) beams in the energy range

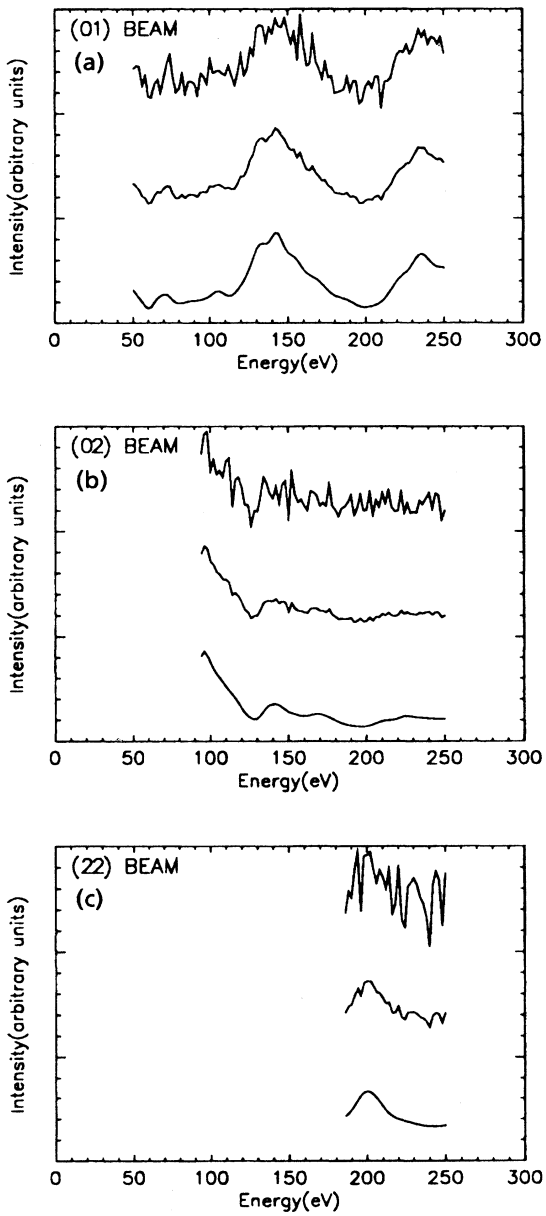

FIG. 2. Simulated experimental diffracted intensities for electrons normally incident on Cu(100) for $I_0 = 0$ (lower curve), 0.1 $I_{\rm av}$ (center curve), and 0.5 $I_{\rm av}$ (top curve). Panel (a): (01) beam; Panel (b): (02) beam; Panel (c): (22) beam.

$50 \le E \le 250$ eV at 2-eV intervals for normally incident electrons. For these electrons, symmetry greatly reduces the number of independent beams [e.g., the (01), (10), $(0\bar{1})$, and $(\bar{1}0)$ are equivalent], to those noted above over this energy range. The specular beam was removed from the data sample to emulate the analysis procedure used by Davis and Noonan.[8] In panels (a)–(c) of Fig. 2, we show simulated experimental intensities for a strong [(01)], moderate [(02)], and weak [(22)] beam, respectively, in the case that the top layer spacing is contracted by 2% (to $d_{12}=1.771\,35$ Å) relative to the bulk interlayer spacing of $d_0=1.80\,750$ Å. Within each panel are simulated intensities with error levels $\sigma_e=0$, $0.1I_{\rm av}$, and $0.5I_{\rm av}$. It is evident that the fine details in the intensity profiles are obscured by the noise of variance $\sigma_e^2=(0.1I_{\rm av})^2$ and that noise of variance $\sigma_e^2=(0.5I_{\rm av})^2$ obliterates even the shapes of the weaker beams. Thus, we see from the figures that visual comparison between calculated and measured intensities becomes of marginal utility for noise levels in the range $0.1I_{\rm av} \le \sigma_e \le 0.5I_{\rm av}$. The visible impact of the added Gaussian noise scales roughly as σ_e^2, with the results for $\sigma_e=0.1I_{\rm av}$ in Fig. 2 being not unreasonable values for typical experimental data.

B. Single-parameter analysis: Contracted top-layer spacing

To generate a single-parameter analysis, we consider the situation in which only the top layer spacing, d_{12}, for Cu(100) deviates from its bulk value. The independent structural variables for Cu(100) are illustrated in Fig. 3. Simulated experimental data were generated for $d_{12}=0.98d_B=1.771\,35$ Å and then were fit using the entire eight-beam database for $50 \le E \le 250$ eV by varying the value of d_{12} used to calculate the theoretical intensities, $I^{\rm th}$.

We performed a structure analysis using the simulated experimental data using three R factors: R_g given by Eqs. (7), the x-ray R factor,[15] R_x, and the global x-ray R factor,[11] $R_{\rm XG}$. The x-ray R factor is defined by[15]

$$R_x=\left[\frac{3}{2n}+\frac{2}{3}\right]r\ , \tag{13a}$$

$$r=\sum_{hk}(r(hk)\Delta E_{hk})\Big/\sum_{hk}\Delta E_{hk}\ , \tag{13b}$$

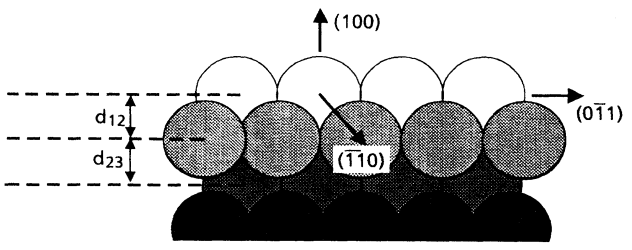

FIG. 3. Illustration of the definition of the independent surface-structure parameters for a relaxed Cu(100) surface. These are the interlayer spacings d_{ij} between the ith and jth layers. $i=1$ designates the top layer.

$$r(hk)=\sum_i[c(hk)I_{hk}^{\rm th}(E_i)-I_{hk}^{\rm exp}(E_i)]^2\Big/\sum_i[I_{hk}^{\rm exp}(E_i)]^2 \tag{13c}$$

$$c(hk)=\sum_i I_{hk}^{\rm th}(E_i)I_{hk}^{\rm exp}(E_i)\Big/\sum_i[I_{hk}^{\rm th}(E_i)]^2 \tag{13d}$$

in which n is the number of beams used in the analysis and ΔE_{hk} is the energy range over which intensities are measured for the (hk) beam. The global x-ray R factor, $R_{\rm XG}$, is defined[11] by Eqs. (13a)–(13c), but using the same normalization constant,

$$c=\sum_{hk}\Delta E_{hk}\sum_i I_{hk}^{\rm th}(E_i)I_{hk}^{\rm exp}(E_i)\Big/\sum_{hk}\Delta E_{hk}\sum_i[I_{hk}^{\rm th}(E_i)]^2\ , \tag{14}$$

for all of the beams, rather than a beam-by-beam normalization. Plots of these three R factors versus d_{12} for the case that $\sigma_e=0.1I_{\rm av}$ are shown in Fig. 4. The associated values of d_{12}(min) and, for R_g, $\sigma(d_{12})$ from Eq. (8b), are given in Tables I and II, each of which has entries derived using a different prescription for construction of the noise in the simulated measured intensities. It is evident from the tables that the extracted values of d_{12}(min) are comfortably within the calculated 95% confidence limit. All of the d_{12} estimates based on global R factors are within the 68% confidence limit. We show in Fig. 5 the comparison between the Gaussian best fit and simulated experimental data for $\sigma_e=0.1I_{\rm av}$. The calculated profiles reproduce the noise free profiles in Fig. 2 to within visual accuracy.

Several checks were made to confirm the validity of the analysis. The minimum value, R_{g0}, of R_g was compared with the variance of the data noise in order to confirm its

TABLE I. Values of d_{12} and its standard deviation predicted by quadratic minimization of various r factors. The "experimental" data were constructed by adding noise of variance σ_e^2 to intensities calculated for a Cu(100) structure with $d_{12}=1.771\,35$ Å. The four noise profiles were generated from a single set of Gaussian errors scaled to the indicated variance levels.

σ_e	0.05 $I_{\rm av}$	0.10 $I_{\rm av}$	0.25 $I_{\rm av}$	0.50 $I_{\rm av}$
d_{12} from R_x	1.7709	1.7705	1.7699	1.7693
d_{12} from $R_{\rm XG}$	1.7713	1.7713	1.7719	1.7740
d_{12} from R_g	1.7714	1.7715	1.7718	1.7723
$\sigma_{d_{12}}$ from R_g	0.0004	0.0007	0.0018	0.0035
d_{12} from R_g (01)	1.7708	1.7702	1.7685	1.7657
d_{12} from R_g (11)	1.7714	1.7715	1.7717	1.7721
d_{12} from R_g (02)	1.7742	1.7768	1.7841	1.7935
d_{12} from R_g (12)	1.7740	1.7768	1.7858	1.8112
d_{12} from R_g (22)	1.7723	1.7757	1.7671	1.7690
d_{12} from R_g (03)	1.7601	1.7534	1.7516	1.7267
d_{12} from R_g (13)	1.7699	1.7700	1.7500	1.7720
d_{12} from R_g (21)	1.7698	1.7680	1.7613	1.7433
Mean d_{12} from R_g (hk)	1.7703	1.7703	1.7675	1.7692
$\sigma^s_{d_{12}}$ from R_g (hk)	0.0045	0.0076	0.0133	0.0264

suitability as a substitute for σ_e^2 [Eq. (10)] in the derivation of the covariance expression, Eq. (8b). R_{g0} was found to be within 10% of the variance of the data noise in all cases. It also is useful to check the scaling of the results with the variance of the data errors. From Eq. (9) [and Eq. (A10) in Appendix A], we expect $\sigma^2(d_{12})$ to scale linearly with σ_e^2 for a constant number of data points. Two sets of calculations were performed to test this hypothesis. The first set of structure searches was performed using four simulated data sets, each of which was generated using intensity errors derived from a single random Gaussian intensity error profile scaled to the desired level. The calculated parameter variances, $\sigma^2(d_{12})$, are shown in Panel (a) of Fig. 6 opposite the scaling levels, I_0^2 (I_0 in units of average intensity). The second set of searches was performed using simulated data sets each of which was generated from separate random Gaussian error profiles with variances σ_0^2. The cal-

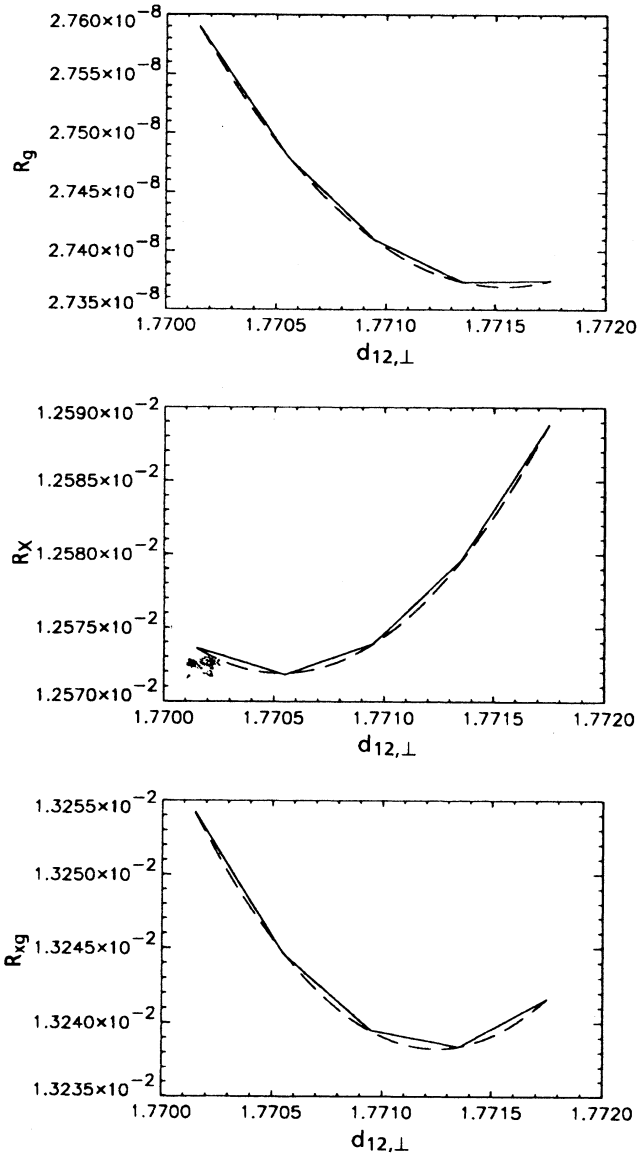

FIG. 4. Plots of R factors versus the independent structural parameter d_{12} for the experimental data simulated using $d_{12}=1.771\,35$ Å $\sigma_0=0.1I_{av}$. Panel (a): $R_g, d_{12}(\min)=1.7715$; Panel (b): $R_x, d_{12}(\min)=1.7705$ Å; Panel (c): $R_{XG}, d_{12}(\min)=1.7713$ Å.

TABLE II. Values of d_{12} and its standard deviation predicted by quadratic minimization of various r factors. The "experimental" data were constructed by adding noise of variance σ_e^2 to intensities calculated for a Cu(100) structure with $d_{12}=1.771\,35$ Å. Each of the four noise profiles was generated using a different set of random Gaussian errors.

σ_e	0.025 I_{av}	0.05 I_{av}	0.17 I_{av}	0.25 I_{av}
d_{12} from R_x	1.7711	1.7709	1.7702	1.7699
d_{12} from R_{XG}	1.7713	1.7713	1.7714	1.7719
d_{12} from R_g	1.7714	1.7714	1.7717	1.7718
$\sigma_{d_{12}}$ from R_g	0.0002	0.0004	0.0012	0.0018
d_{12} from R_g (01)	1.7711	1.7708	1.7694	1.7685
d_{12} from R_g (11)	1.7714	1.7714	1.7716	1.7717
d_{12} from R_g (02)	1.7728	1.7742	1.7801	1.7841
d_{12} from R_g (12)	1.7727	1.7740	1.7807	1.7858
d_{12} from R_g (22)	1.7717	1.7723	1.7557	1.7671
d_{12} from R_g (03)	1.7647	1.7600	1.7498	1.7516
d_{12} from R_g (13)	1.7698	1.7700	1.7697	1.7500
d_{12} from R_g (21)	1.7706	1.7698	1.7653	1.7613
Mean d_{12} from R_g (hk)	1.7706	1.7703	1.7678	1.7675
$\sigma^s_{d_{12}}$ from R_g (hk)	0.0026	0.0045	0.0108	0.0133

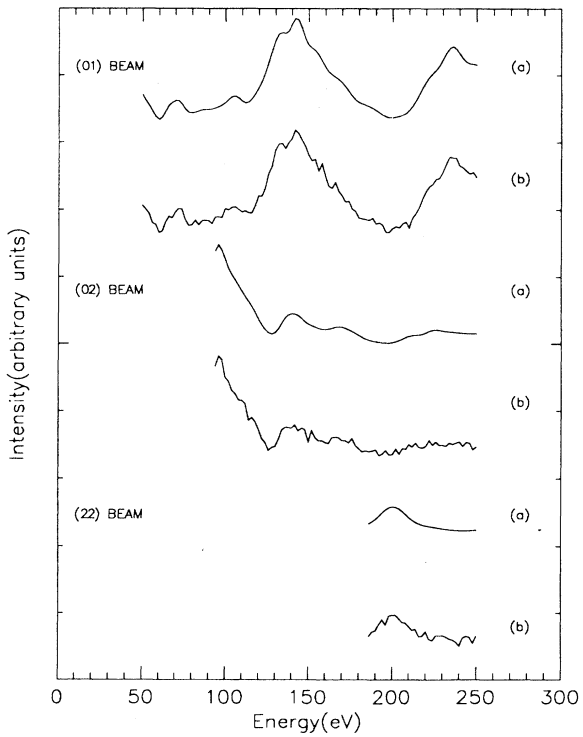

FIG. 5. Comparison of the Gaussian best-fit and simulated experimental intensity profiles for $\sigma_0=0.1I_{av}$.

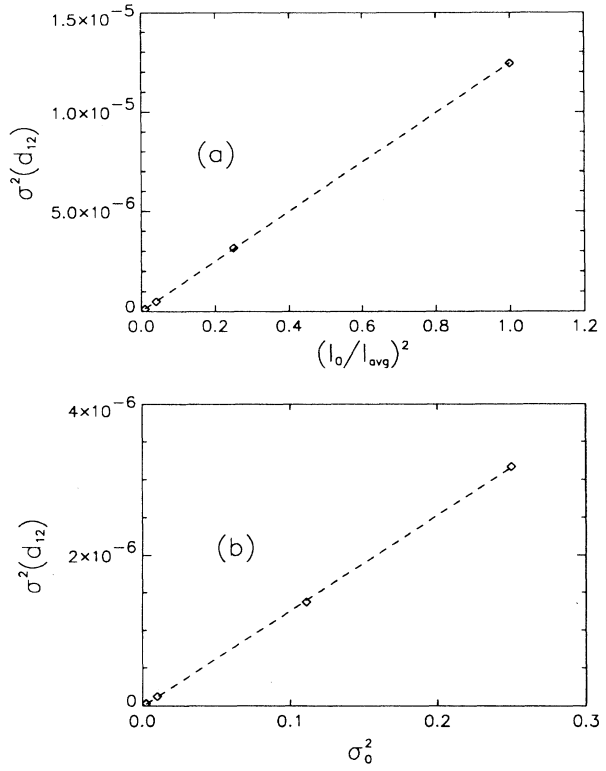

FIG. 6. Scaling of $\sigma^2(d_{12})$ with σ_e^2. Panel (a): Errors used to construct simulated data for each noise level were generated by scaling a single random noise profile of variance $(0.5)^2$; Panel (b): Each simulated data set was generated from a different random noise profile of variance, σ_0^2; the random error was scaled by $I_0=0.5I_{\rm av}$.

culated parameter variances are shown in Panel (b) of Fig. 6. Both sets of calculations demonstrate the linear scaling of the parameter variance with the data error variance. Furthermore, from Table I, we see that with increasing data variance the uncertainty, $\sigma_{d_{12}}$, in the parameter estimate (i.e., of d_{12} calculated by minimizing R_g), grows in proportion with σ_e. Moreover, it is of interest that at moderate data error levels both of the global R factors, R_g and $R_{\rm XG}$, give more accurate values of d_{12} than does the beam-by-beam normalized R factor, R_x. It is evident from Fig. 2, Fig. 5, and Tables I and II that accurate structural parameters can be extracted from simulated measured intensities that appear extremely noisy if the noise is uncorrelated and has a Gaussian distribution.

A common[6,8,35,36] method of estimating uncertainties in structural parameters for relaxed metal surfaces is to determine parameter values from each beam independently and to interpret the variance of the beam-specific parameter estimates as a measure of the uncertainty in the selected value. We refer to these calculations as estimates of the sampling uncertainties, because they measure the extent to which presumably independent samples of the data (i.e., the individual beams) yield the same values for the structural parameters. Performing this analysis for data sets with a broad range of noise levels gives the results presented in the lower rows of Tables I and II. The sampling uncertainties $\sigma^s(d_{12})$ for these cases are the estimated standard deviations of the corresponding Gaussian distributions given in the last row of each table. It is clear from the tables that this estimation procedure generates less accurate estimates and significantly larger variances than use of the global estimation procedures based on R_g or $R_{\rm XG}$. The sampling estimates appear to be sensitive to large deviations obtained from a few beams. For our calculations, much of the variance results from the (03), (13), and (12) beams. It should be of little surprise that the $I(V)$ profiles of a few individual beams may have low sensitivity to particular parameters,[37] and that the R-factor surface for a single beam might display multiple minima and assume its lowest value at a structure, which is well removed from the global minimum of the multibeam R factor. It is more plausible to view large differences between single-beam structures as a measure of the spacing between local minima than as a measure of uncertainty in the parameters determined by a proper optimization. Issues related to the low sensitivity of single-beam $I(V)$ data to particular parameters can be expected to be of even greater concern in multiparameter structure problems. The large sampling errors and erroneous standard deviations displayed in the last two rows of Tables I and II can be presumed to be modest, relative to what is likely to be observed when the sampling method is applied to more complex problems.

C. Two parameter analysis: Double-layer relaxation

Cu(100) is believed to exhibit at least a double-layer surface relaxation. The two analyses thereof[8–10] yield results for these relaxations which differ, however, by more than the errors quoted by one of the two groups of authors. If we write

$$d_{ij}=d_B+\Delta d_{ij} \tag{15}$$

to describe the actual layer spacings, we can specify the two sets of results in terms of Δd_{ij}. Davis and Noonan[8,9] give

$$\Delta d_{12}=-0.0199\ \text{Å}\pm 0.0072\ \text{Å}\ , \tag{16a}$$

$$\Delta d_{23}=+0.0307\ \text{Å}\pm 0.0108\ \text{Å}\ , \tag{16b}$$

whereas Mayer *et al.*[10] give

$$\Delta d_{12}=-0.0380\ \text{Å}\pm 0.0307\ \text{Å}\ , \tag{17a}$$

$$\Delta d_{23}=+0.0081\ \text{Å}\pm 0.0307\ \text{Å}\ . \tag{17b}$$

The generous error estimates of Mayer *et al.* render the two analyses marginally compatible. We thought it of interest, however, to repeat the structure analysis using a procedure equivalent to that of Noonan and Davis to determine if there exists a level of random noise at which the two analyses became compatible. If such were the

case, then uncertainties in the experimental LEED intensities could account for the deviations between the two groups' results. If not, these deviations must result from systematic errors associated, e.g., with the construction of the model of the diffraction process. Thus, for each of the structures given by Eqs. (16) and (17), we generated simulated measured intensities as before for a range of values of σ_e, $0.05I_{\rm av} \le \sigma_e \le 0.5I_{\rm av}$.

A two part search was used to locate the R factor minimizing structure for each simulated data set. Approximate structures were located using a simplex algorithm. These structures are shown in Table III. Subsequently, the minimum of each R_g surface was located by fitting a set of R_g values corresponding to a grid of d_{12}, d_{23} values to a quadratic form and analytically determining the structure which corresponded to the R_g minimum. (See Appendix B.) These structures are shown in Table IV. Also included in the table are the square roots of elements of the covariance matrices. Square roots (σ) of the variances (σ^2) in the estimates of d_{12} and d_{23} are labeled as σ_{12} and σ_{23}; the square root of the off-diagonal matrix element, which describes the correlation of uncertainties in estimates of d_{12} and d_{23}, is labeled as $\sigma_{12,23}$.

The parameter estimation errors (i.e., the deviations between the estimated and input values) again are small, even for noise levels which appear to be quite large. These errors scale with the standard deviation of the measurement noise as do the elements of the covariance matrices. The errors in estimating d_{23} are approximately 5 times as great as those in d_{12}. Moreover, the d_{23} uncertainty variances are approximately 2.5 times the size of the d_{12} variances. The quadratically extrapolated best fit structures are slightly outside the 95% confidence intervals around the input structure, although the estimated d_{12} values are themselves well within the 95% intervals around the input value of d_{12}. Cross terms in the covariance matrices identify correlation between the uncertainties in the first and second interlayer spacings. It is these correlations that explain the larger size of the *errors* in d_{12} that occur when d_{23} is allowed to vary. The d_{12} *variances* can also be seen to be larger when d_{23} is identified as an unknown structural parameter than when d_{23} is assumed to be known.

The estimation errors and covariances calculated here indicate that the presence of uncorrelated Gaussian measurement noise does not suffice to explain the differences between the structures determined by Davis and Noonan and by Mayer *et al.* Correlated measurement errors and/or deficiencies in the model are likely explanations for the discrepancy between the two structures. The small data set at off-normal incidence of Mayer *et al.* presumably interacted with model deficiencies differently than did the large data set at normal incidence of Davis and Noonan. In addition, the structural estimate of Davis and Noonan is an average of structural estimates calculated from single beam data sets. As single-beam data sets have difficulty selecting the best among local R-factor minima, equivalent beam estimates often have large correlated errors which displace the average structure from the global R-factor minimum.

TABLE III. Interlayer spacings predicted by a two-dimensional downhill simplex, with the range of the final simplex < 0.002 Å in each dimension.

Structure	σ_e	0	0.05 $I_{\rm av}$	0.25 $I_{\rm av}$	0.5 $I_{\rm av}$
Davis and Noonan	d_{12} (Å)	1.7876	1.7879	1.7899	1.7923
	d_{23} (Å)	1.8382	1.8294	1.8228	1.8145
Mayer *et al.*	d_{12} (Å)	1.7695	1.7683	1.7697	1.7690
	d_{23} (Å)	1.8156	1.8241	1.8236	1.8236

TABLE IV. Values of d_{12} and d_{23}, which minimize R_g and their standard deviations obtained from the covariance matrix.

Structure	σ_e	0	0.05 $I_{\rm av}$	0.25 $I_{\rm av}$	0.5 $I_{\rm av}$
Davis & Noonan	d_{12} (Å)	1.7876	1.7880	1.7899	1.7923
	d_{23} (Å)	1.8382	1.8262	1.8265	1.8145
	σ_{12} (Å)		0.0004	0.0020	0.0041
	σ_{23} (Å)		0.0010	0.0049	0.0100
	$\sigma_{12,23}$ (Å)		0.0004	0.0021	0.0045
Mayer *et al.*	d_{12} (Å)	1.7695	1.7699	1.7718	1.7738
	d_{23} (Å)	1.8156	1.8139	1.8048	1.7944
	σ_{12} (Å)		0.0004	0.0020	0.0041
	σ_{23} (Å)		0.0010	0.0048	0.0094
	$\sigma_{12,23}$ (Å)		0.0004	0.0021	0.0043

IV. DISCUSSION

Quantitative error analysis is a complex topic.[1,38,39] Many variants exist in various specialties associated with the structure determination of solids, especially X-ray diffraction[22,38,39] and LEED.[4] Our purpose in this section is to provide a compact overview of the essential ingredients of these methods and to compare them to the one proposed above.

To achieve this goal, the section is organized into four parts. In Sec. IV A, we sketch the derivation of the present method from the dual assumptions of linearity of outputs (intensity variations) with inputs (structural parameter variations) and of uncorrelated Gaussian noise in the measurement of the intensities as the sole source of uncertainty in the analysis. This subsection provides the concepts and nomenclature needed to compare our method with others. In Sec. IV B, we note that the present methodology is remarkably similar, but not identical, to several others, including a few commonly used to analyze LEED intensity data. In Sec. IV C, we expand our scope to include those analyses used in x-ray crystallography and in certain heuristic analyses of LEED intensities. Then, having developed the requisite concepts and mathematics in Secs. IV A–IV C, in Sec. IV D we give a concise comparison of our methodology with others commonly used in LEED. Presentation of the essential requisite details is relegated to the appendixes.

A. Derivation of the methodology

In this subsection, we indicate how the twin assumptions of linearity and uncorrelated Gaussian measure-

ment noise lead to the methodology specified in Sec. II. The purpose of this discussion is to permit the comparison of this methodology with others proposed for the same purpose.

Symbolically, a linear LEED theory (e.g., the linear tensor LEED approximation[24]) relates intensity differences, $I^{\rm th}(a)-I^{\rm th}(a_0)$, to the structural model, a (specifically, the variation, $m=a-a_0$, of the structure from a reference model, a_0), via a linear mapping, G. The inverse mapping, M, can be defined in such a way as to yield the least-squares solution, $a=a_0+M[I^{\rm exp}-cI^{\rm th}(a_0)]$, $M=(G^TG)^{-1}G^T$ when applied to the data residual. If the theory is perfect, then a may be determined by minimizing the least-squares R factor (Appendix A),

$$R_g=\frac{1}{F}[I^{\rm exp}-cI^{\rm th}(a)]^T[I^{\rm exp}-cI^{\rm th}(a)]\ , \tag{18}$$

rather than by direct inversion. The normalization factor, F (equal to the number of data points less the number of parameters being estimated including a scale factor), is selected so as to make R_g a measure of the mean-square deviation of the theory/data intensity difference. The parameter error covariance is related to the covariance of the data residuals, d, via[1]

$$[\text{cov } a]=M[\text{cov } d]M^T\ . \tag{19}$$

In the case of uncorrelated measurement errors of constant variance, [cov d] is a diagonal matrix of equal matrix elements σ_e^2 and Eq. (19) becomes

$$[\text{cov } a]=\sigma_e^2[G^TG]^{-1}\ . \tag{20}$$

Within the region of parameter space where the data residual d is linearly related to parameter variations ($d=Gm$, $m=a-a_0$), $R_g(a)$ can be represented as a second-order Taylor series expansion about its minimum $R_g(a_0)$,

$$R_g(a)=R_g(a_0)+[a-a_0]^T\left[\frac{1}{2}\frac{\partial^2 R_g}{\partial \mathbf{a}^2}\right]_0[a-a_0]\ , \tag{21}$$

and the jk elements of the Hessian matrix, $\partial^2 R_g/\partial \mathbf{a}^2$, may be expressed in the simple form

$$\frac{1}{2}\frac{\partial^2 R_g}{\partial a_j \partial a_k}=\frac{1}{F}\sum_i \frac{\partial d_i}{\partial a_j}\frac{\partial d_i}{\partial a_k}=\frac{1}{F}\sum_i \frac{\partial I_i^{\rm th}}{\partial a_j}\frac{\partial I_i^{\rm th}}{\partial a_k}\ . \tag{22}$$

(Note that the expression for the Hessian loses its simplicity if the data-model relationship is nonlinear.) From Eq. (22), it follows that (Appendix A)

$$\left[\frac{F}{2}\frac{\partial^2 R_g}{\partial \mathbf{a}^2}\right]_0=G^TG \tag{23}$$

and hence from Eq. (20)

$$[\text{cov } a]=\frac{2\sigma_e^2}{F}\left[\frac{\partial^2 R_g}{\partial \mathbf{a}^2}\right]_0^{-1}\ . \tag{24}$$

As described in Sec. II, if the theory is perfect, R_{g0} is an estimate of σ_e^2 and the expression is rewritten as

$$[\text{cov } a]=\frac{2R_{g0}}{F}\left[\frac{\partial^2 R_g}{\partial \mathbf{a}^2}\right]_0^{-1}\ , \tag{25}$$

which is Eq. (12b). Use of a least-squares R-factor leads, therefore, to an expression for the parameter estimation error covariance in which the data variance, σ_e^2, and the inverse mapping matrix, M, [both of which appear explicitly in the general expression, Eq. (19)] are replaced by the R factor and its second partial derivatives at the R-factor minimum. Thus, *in the presence of a perfect theory, the effect of Gaussian measurement noise on parameter estimation can be determined without explicit knowledge of the data error level and without calculating the data/model mapping matrix, $G=[\partial I/\partial a]_0$, let alone its inverse, M, provided that the structure is determined using the least-squares Gaussian R factor, R_g.*

The error analysis, like the structure analysis itself, uses the intensities directly, without any preprocessing other than perhaps averaging of equivalent $I(E)$ profiles.[8,9] This procedure contrasts with common practice in bulk crystallography where significant data reduction (peak integration, background subtraction, and scaling to compensate for geometry and polarization effects) is performed prior to initiating structure searches and where parameter error covariances are calculated from the covariances of the reduced data.[38] Thus, while the bulk structure estimation problem (from X-ray data) is free of the modeling imperfections that complicate surface-structure determination by LEED, the analysis of parameter uncertainties in surface structures determined by LEED data is facilitated by the absence of data preprocessing, which typically complicates an error analysis.

B. Closely related analysis methods

Variations of the analysis described here admit the use of alternative, though closely related, R factors. In the case of uncorrelated measurement errors of known variance (not necessarily constant), statistically significant parameter estimation covariances can be calculated for structures determined using an R factor, which is a weighted sum of squares with weights chosen to be the inverse of the data variance, $w_i=1/\sigma_i^2$, in which case the residuals are effectively normalized by the standard deviation, σ_i. The data/model relationship is rewritten as $d'=G'm$, where $d'=Ad$, $G'=AG$, and A is a diagonal normalization matrix with elements $1/\sigma_i$. (Note that d is now being used to represent the data, not the residual.) The least-squares R factor becomes $R_g'=(1/F)[d'-G'm]^T[d'-G'm]$, the least-squares inverse, M, becomes $M'=(G'^TG')^{-1}G'^T$, and the covariance becomes[38]

$$[\text{cov } a]=M'M'^T=[G'^TG']^{-1}=\left[\frac{F}{2}\frac{\partial^2 R_g'}{\partial \mathbf{a}^2}\right]_0^{-1} \tag{26a}$$

or

$$[\text{cov } a]=\left[\frac{1}{2}\frac{\partial^2 R}{\partial \mathbf{a}^2}\right]_0^{-1}\ , \tag{26b}$$

for the χ^2 R factor

$$R = FR'_g = [d'-G'm]^T[d'-G'm] = \sum_i \frac{1}{\sigma_i^2}[d-Gm]_i^2 \ . \tag{26c}$$

The matrix being inverted in Eq. (26b) is referred to as the curvature matrix. The particularly simple form of Eq. (26b) follows from the fact that the covariance matrix of the normalized data is an identity matrix.

If only the relative measurement error variances are known and the R factor, R''_g, is a weighted sum of squares with weights inversely proportional to the data variance $w_i = k/\sigma_i^2$, k arbitrary, the residuals are effectively normalized by $\sqrt{k}/\sigma_i$. In this case, the variance of the normalized data is k and, if, as we are assuming, the theory is perfect, k is estimated by R''_g (assuming that R''_g includes a $1/F$ factor). The parameter estimation error covariance assumes the familiar form,

$$[\text{cov } a] = \frac{2R''_{g0}}{F}\left[\frac{\partial^2 R''_g}{\partial \mathbf{a}^2}\right]_0^{-1} \ . \tag{27}$$

Any R factor that differs from R''_g by a constant factor (for instance, the R factor identical to R''_g, but with a normalization different from $1/F$) not only leads to the same surface structure but also can be, used in Eq. (27) to calculate the error covariance. Similarly, any R factor that differs from R_g by a constant factor not only leads to the same surface structure as that obtained by minimizing R_g, but also can be used in Eq. (25) to calculate the error covariance. One such alternative is the R factor,

$$R = \Sigma[I^{\text{exp}} - cI^{\text{th}}(a)]^2 / \Sigma (I^{\text{exp}})^2 \ , \tag{28}$$

with c defined as in Eq. (7c)–(7e). This R factor (which treats all data points equivalently) is a variation of the R factor known as R_2;[16] the latter [see Eq. (44a)] weighs individual beam R factors, $R_{2,hk}$, according to the beam's energy range.

A closely related R factor is the alternatively normalized

$$R'_2 = C_g R_g = \frac{N}{F}\frac{\sum_{\alpha,hk,i}[I^{\text{exp}} - cI^{\text{th}}(a)]^2}{\sum_{\alpha,hk,i}(I^{\text{exp}})^2} \ , \tag{29a}$$

$$C_g = \frac{N}{\sum_{\alpha,hk,i}(I^{\text{exp}})^2} \ , \tag{29b}$$

where R_g, $F = N - M - 1$, and c are as in Eq. (7a)–(7e). Since R'_2 differs from R_g by a multiplicative constant, C_g, its minimization leads to the same structure as does minimization by R_g, and it can be used in Eq. (27) to calculate the error covariance. While R_g is an estimate of the variance, σ_e^2, of the residual, $I^{\text{exp}} - cI^{\text{th}}(a)$, R'_2 is an estimate of the data variance in units of mean-squared intensity,

$$R'_2 \approx \frac{\sigma_e^2}{\langle I^2 \rangle} \ . \tag{30}$$

This R factor is labeled R'_2 because of its similarity to the previously introduced R_2, which uses constant (beam independent) weights and a single scale factor.

C. Characteristics of alternative error analysis methods

Least-squares R factors and covariance analyses are used by x-ray crystallographers to determine bulk crystal structures and parameter estimation error levels.[22,38,39] X-ray R factors are typically weighted sums of squares either of the form defined by Eq. (13), which are not useful for covariance analysis, or with weights, W, corresponding to the inverse of the data variance, i.e.,

$$R = [d-Gm]^T W [d-Gm] \ , \tag{31a}$$

in which case the covariances can be calculated[39] as

$$[\text{cov } a] = [G^T W G]^{-1} \ . \tag{31b}$$

If the data are assumed to be uncorrelated, the covariance matrix and weight matrices become diagonal with diagonal elements σ_i^2, and $1/\sigma_i^2$, respectively, so the analysis becomes the same as that described by Eqs. (26b), (26c). The structure search requires estimated standard deviations, σ_i, for each of the reduced data points as inputs. Various schemes[38,39] are used to determine the σ_i's involving the assessment of the contributing errors, comparisons of symmetrically equivalent beams, or examination of residuals in preliminary calculations. A fundamental difference between LEED analyses of the surface structure and bulk structure analyses, which use x ray and neutron data, is that the latter make use of scaled integrated peak intensities rather than the intensity measurements themselves. This is true for both monoenergetic experiments, which require intensity integrations over solid angles (or one dimensional integrations with geometrical scaling) and diffuse diffraction experiments, in which intensities are integrated over energy in order to determine the area of each Bragg peak. Because of the weakness of the scattering processes involved in bulk structure experiments, the diffraction process functions as a physical realization of a Fourier transform of the crystal structure with each integrated intensity corresponds to the real part of a Fourier coefficient. The integrated intensities corrected for background, emission, and polarization effects are thus the fundamental observables.

Within the field of surface crystallography, covariance analysis is uncommon. The most widely used methods use R-factor surface curvature to calculate parameter variations from R-factor variations.[4,17,40,41] In such methods, a deterministic mapping of specific variations is reinterpreted as a mapping appropriate for statistical quantities, i.e., variances and standard deviations. Specifically, the variation with a given parameter, a, of the R factor in the vicinity of its minimum, $R_0 = R(a_{\text{est}})$, is described quadratically as

$$R(a)=R_0+\left[\frac{1}{2}\frac{\partial^2 R}{\partial a^2}\right]_0(\Delta a)^2 . \tag{32}$$

The square of the parameter change, $\Delta a = a - a_{\text{est}}$, can thus be written in terms of the corresponding change in the R factor, $\Delta R = R(a) - R_0$,

$$(\Delta a)^2=\left[\frac{1}{2}\frac{\partial^2 R}{\partial a^2}\right]_0^{-1}\Delta R . \tag{33}$$

While this relation is typically established by eye using an R-factor plot (see, e.g., the discussion of Fig. 2 in Ref. 17), as illustrated in Fig. 7, Eq. (33) provides the analytical description. The specific variations Δa (or, more frequently, $2\Delta a$) and ΔR are then interpreted as being analogous to standard deviations and the above relation is invoked as mapping the standard deviation of the R factor (σ_R, sometimes mistakenly referred to as the variance) to the variance, σ_a^2, of the parameter error distribution,

$$\sigma_a^2=\left[\frac{1}{2}\frac{\partial^2 R}{\partial a^2}\right]_0^{-1}\sigma_R . \tag{34}$$

Unfortunately, this seemingly natural relationship is not correct. If the R-factor minimum can be construed as having a χ^2 distribution with F degrees of freedom (as would be the case if the R factor is a sum of squares of N zero-mean random variables with Gaussian distributions of constant variance, subject to M constraints, such that $F=N-M$), then the ratio of the standard deviation of R_0 to its mean, $\langle R_0\rangle$, is

$$\frac{\sigma_R}{\langle R_0\rangle}=\frac{\sqrt{\text{var}R}}{\langle R_0\rangle}=\left[\frac{2}{F}\right]^{1/2} . \tag{35}$$

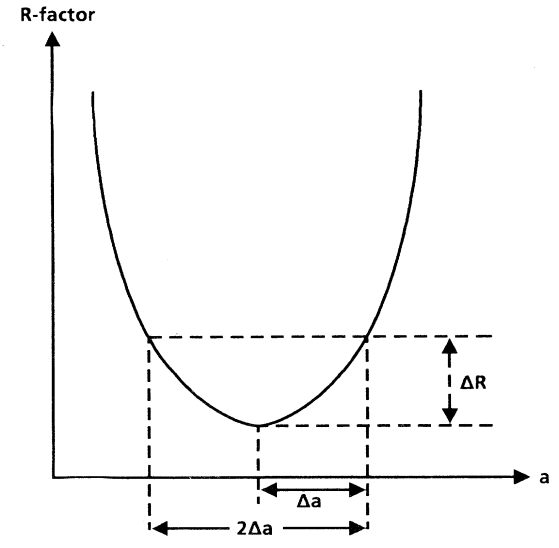

FIG. 7. R factor plot, illustrating the generation of an uncertainty Δa in a structural parameter in terms of a range, ΔR, of R factor values from its minimum.

Assuming that the mean, $\langle R_0\rangle$, can be approximated by a given realization, R_0, Eq. (34) becomes

$$\sigma_a^2=\left[\frac{1}{2}\frac{\partial^2 R}{\partial a^2}\right]_0^{-1}\left[\frac{2}{F}\right]^{1/2}R_0 . \tag{36}$$

For comparison, the correct expression for parameter estimation error variance/covariance in the presence of intensity errors [Eq. (8b) for the variance of a single variable parameter] is rewritten here in terms of the χ^2 R factor, $R = FR_g = [d - Gm]^T[d - Gm]$, which is an unscaled sum of squares of Gaussianly distributed errors with a constant variance, as

$$\sigma_a^2=\left[\frac{1}{2}\frac{d^2 R}{da^2}\right]_0^{-1}\frac{R_0}{F} . \tag{37}$$

The $1/F$ factor in the properly derived expression is seen to be replaced by $(2/F)^{1/2}$ in the commonly used Eq. (36). Thus, reinterpretation of the $\Delta R, (\Delta a)^2$ relation in Eq. (33) as a mapping of statistical parameters characterizing distributions leads to an incorrect relation even in the simplest case of a single variable parameter.

Another error analysis reported in the literature[37] consists of calculating standard deviations for estimation error in a given parameter by inverting and scaling the corresponding diagonal element of the curvature matrix. We refer to this as a one-dimensional variance calculation. When the system being modeled has only one unknown parameter (and the curvature matrix consists of a single element), this is correct. If there is more than one unknown parameter, however, the curvature matrix inversion required for proper calculation of the parameter error variances [diagonal elements of the covariance matrix, see Eq. (12b)] is poorly approximated by inversion of the diagonal elements alone, unless the off-diagonal elements of the curvature matrix are all negligible. As the latter condition is satisfied infrequently in the case of two unknown parameters and almost never when the number of parameters is three or more, approximation of the full variance calculation with one-dimensional inversions leads to a misrepresentation of the errors.

An alternative to inverting the full covariance matrix directly is to diagonalize it, i.e., identify the linear combinations of parameter errors that are uncorrelated and transform the curvature matrix into this basis of independent parameter errors. The variances of these linear combinations of parameter errors can then be calculated one at a time from their own curvatures. Determination of the variances of the original parameter estimates requires back transformation to the original basis, hence, for a given parameter, involves combining the variances of all the independent basis elements that include that parameter. Such a procedure is equivalent to the method being advocated here in terms of both the results and the level of effort required.

This alternative procedure is sometimes described in geometric rather than analytic terms.[37] Diagonalization of the curvature matrix is equivalent to identifying the principle axes of an ellipsoid in parameter space corresponding to a surface of constant R_g, where R_g is within

a few percent of R_{g0}. The curvatures along the principle axes are the same as the diagonal elements of the diagonalized curvature matrix, and thus can be inverted individually. Again, determination of the variances of the original parameter estimates requires back transformation of the curvatures to the original basis.

D. Comparison with other LEED/LEPD error analyses

Having indicated the general principles of parameter estimation error variance analysis, we can now relate our methodology to some of the analysis methods currently in use. One of the most widely used procedures for estimating errors is that introduced by Pendry.[17] Parameter error variances are calculated from an R-factor standard deviation by means of a mapping, which is based on the relationship [Eq. (33)] between changes in the R factor and changes in individual structural parameters, as embodied in the R-factor curvature. Under the assumption that peaks in the $I(E)$ curves are well separated, the Pendry R factor is hypothesized to exhibit χ^2 statistics with N degrees of freedom corresponding to N Lorentzian intensity peaks (see Appendix C), such that the standard deviation of its minimum value, R_P, can be expressed in terms of its mean, $\langle R_P \rangle$. Thus, if the latter is approximated by a given realization ($\langle R_P \rangle \approx R_P$), the standard deviation of R_P is trivially calculated from R_P itself as

$$\sigma_{R_P} = \left[\frac{\sigma_{R_P}}{\langle R_P \rangle} \right] \langle R_P \rangle \approx \left[\frac{2}{N} \right]^{1/2} R_P \ . \tag{38}$$

For each parameter estimate, an error bar is assigned whose magnitude (illustrated geometrically in Fig. 2 of Ref. 17) is the full width at $R = R_P \pm \sigma_{R_P}$ of the R-factor curve as a function of the parameter of interest, i.e., $2\Delta a$ where Δa is as defined above ($\Delta a = a - a_{\text{est}}$) and $(\Delta a)^2$ is as specified by Eq. (33). As indicated following Eq. (37), the deterministic relation described by Eq. (33) is not appropriate for mapping statistical moments and yields results which differ from those calculated with proper statistical relations. Furthermore, its frequent application to multiparameter structural estimation problems suffers from neglect of correlation among parameter estimation errors. As discussed above, parameter variance calculations based on inversion of individual diagonal elements, $[\partial^2 R / \partial a^2]_{ii}$, of the full Hessian matrix [Eq. (12c)] yield variances which are typically poor approximations to the true variances.[42] Only if R-factor curvature with respect to a given parameter, a_i, is calculated from an R-factor curve which is evaluated by optimizing all other parameters a_j at each value of a_i, does individual element inversion yield a statistically meaningful estimate of the parameter variances.

In contrast to the nonstatistically based methods, which neglect the effects of correlations, the error analysis method of Adams and co-workers,[35,37,43–46] bears great structural similarity to the methods developed here. Parameter estimation error variances are calculated using the relation

$$\sigma_{ai}^2 = \frac{\varepsilon_{ii} R}{F} \ , \tag{39}$$

where ε_{ii} is a diagonal element of the inverse of the Hessian matrix,

$$H = \left[\frac{\partial^2 R}{\partial a^2} \right] \ , \tag{40}$$

and F is the number of degrees of freedom, equal to the number of data points less the number of variable parameters. Equation (39) is a variation of the relation found in the reference[47] cited by Adams[37], which is

$$\sigma_{ai}^2 = \varepsilon_{ii} \ , \tag{41a}$$

where ε_{ii} is a diagonal element of the inverse of the curvature matrix,

$$\varepsilon_{ii} = \left[\frac{1}{2} \frac{\partial^2 R_{\chi^2}}{\partial a^2} \right]_{ii} \ , \tag{41b}$$

and R_{χ^2} is the χ^2 R factor defined in Eq. (26c). The expression for the parameter estimation error variance given by Eqs. (41) is identical to that of the diagonal elements of the full covariance matrix described above in Eq. (26b) and is the relation appropriate for a sum of squares R factor with each squared residual weighted by its variance. As shown in Sec. IV B, Eq. (39), with an appropriate choice of R factor, is also a statistically valid expression if ε and F are properly defined. Unfortunately Adams' definition (see above) of the error matrix, ε, differs by a factor of 2 from that of the original reference,[47] and this error persists in the analyses that follow. More troublesome, however, is the practice of employing Eq. (39) with R factors for which it is not designed. While Adams *et al.* in their original presentation[37] recognized the statistical importance of using χ^2-type R factors with a constant scale factor, c, in the calculation of parameter estimation errors, later works embodied other R factors with implications that need to be made explicit.

One important R factor commonly used[35,37,42–46] together with the above analysis is $\hat{R}_2$ defined by

$$\hat{R}_2 = \frac{1}{N} \sum_{hk} N_{hk} R_{2,hk} \ , \tag{42a}$$

where N_{hk} is the number of data points in the hk beam, N is the total number of data points,

$$N = \sum_{hk} N_{hk} \ , \tag{42b}$$

and $R_{2,hk}$ is the individual beam R factor,

$$R_{2,hk} = \sum_i w_{hk} (I_{hk,i}^{\text{exp}} - c_{hk} I_{hk,i}^{\text{th}})^2 \ , \tag{42c}$$

with weights (constant for a given beam)

$$w_{hk} = 1 \Big/ \sum_i (I_{hk,i}^{\text{exp}})^2 \ , \tag{42d}$$

and scale factors

$$c_{hk} = \sum_i I_{hk,i}^{\text{exp}} \Big/ \sum_i I_{hk,i}^{\text{th}} \ . \tag{42e}$$

[Reference 35 gives an expression for the single-beam R factor [Eq. (42c)], in which the w_{hk} are squared. This is in contrast to the usual definition and, as it would lead to an R factor with units corresponding to those of (intensity)$^{-2}$, presumably is a notational error.] $\hat{R}_2$ is a version of the R_2 R factor introduced by Van Hove, Tong, and Elconin[16] and used, for example, by Davis and Noonan.[8] It can be rewritten as

$$\hat{R}_2 = \sum_{hk,i} w_{hk} (I^{\rm exp}_{hk,i} - c_{hk} I^{\rm th}_{hk,i})^2 , \tag{43a}$$

where the sum is now over all intensities in all beams,

$$w_{hk} = \frac{N_{hk}}{N(\Sigma_i I^{\rm exp}_{hk,i})^2} = \frac{1}{N\langle I^2_{hk}\rangle} , \tag{43b}$$

and $\langle I^2_{hk}\rangle$ is the mean-squared intensity of a given beam.

Unfortunately, use of $\hat{R}_2$ does not automatically lead to a consistent statistically significant error analysis. As described in Sec. IV B, simple expressions such as those in Eqs. (25)–(27), (39), and (40) exist for parameter error variance when either (1) the residual errors are uncorrelated and have constant variance and the R factor is an unweighted sum of squared residuals or (2) the matrix [see Eq. (31) or Ref. 39] used to weight residuals in the R-factor sum is the inverse of the covariance, which describes residual errors. The beam specific weights in Eq. (43) preclude a simple form on the basis of the first set of conditions. For Eq. (39) (corrected for the missing factor of 2) to be valid, the R factor must weight each residual, if not with the inverse of its variance, then with a weight proportional to the inverse of its variance. Hence, an implicit assumption of using $\hat{R}_2$ in Eq. (39) is that σ^2_{hk} is proportional to $\langle I^2_{hk}\rangle$ such that

$$w_{hk} \propto \frac{1}{\sigma^2_{hk}} , \tag{44}$$

in order for Eq. (39) to be valid. While such an assumption may be plausible, it is debatable, and needs to be made explicit. The statistical implications of including beam-specific scale factors, c_{hk}, also is not clear. Thus, while the introduction of curvature matrix inversion to the LEED error analysis methodology has greatly improved the treatment of parameter estimation error correlation, the specific analyses reported in the literature incorporate elements of covariance analysis methodology in a less than rigorous fashion.

A procedure analogous to that used for R_2 by Adams *et al.* was applied to R_1, defined

$$R_1 = \sum_{hk,i} |c I^{\rm th}_{hk}(E_i) - I^{\rm exp}_{hk}(E_i)| \tag{45}$$

by Kleinle and co-workers[20,21] to define an R factor, R_{DE}, suitable for use in fast automated structural searches. R_{DE} was proposed because of its insensitivity to the spacing between energy values (E_i) along an intensity profile for a given beam. Like R_2, however, R_{DE} is a beam-by-beam normalized R factor, whose weights are not purported to be inverses of the data variances. While we agree with the general philosophy of data selection articulated by Klein *et al.*, we selected R_g instead of R_{DE} because of its connection with Gaussian error statistics which provides the benefit of a well-defied statistical interpretation of the error analysis defined in Sec. II.

V. SYNOPSIS

Our purposes in this paper have been to adapt a well known[1,38] statistically significant error analysis procedure to the case of structure determination via analysis of LEED and LEPD intensity data; to demonstrate that this adaptation produces sensible results compatible with its promises in a specific case, i.e., Cu(100); and to compare it with the myriad of alternatives proposed in the literature to illuminate its similarities and differences with them. Specifically, we demonstrate that surface-structure parameter estimation based on dynamical analyses of LEED intensity data can be expressed as a well-defined procedure that yields parameter estimates, whose errors have magnitudes and correlations which are statistically significant in the limit that measurement errors are the sole source of uncertainty in surface-structure parameter estimates. The quality of the parameter estimates is described in terms of the distribution of the parameter estimation errors and is determined by the data error distribution. The error analysis is facilitated by making the additional assumptions that the *measurement* error distributions are Gaussian and that the errors fall within a range in which a linear LEED theory is adequate. These assumptions are useful, because linear systems map zero-mean Gaussian distributions to zero-mean Gaussian distributions, and hence lead to the result that *parameter* estimation errors will themselves be Gaussianly distributed. An additional convenience of using Gaussian error models follows from the fact that Gaussian distributions are fully characterized by their means and covariances; in the zero-mean case, a valid expression for the parameter error covariance is sufficient to fully characterize the parameter estimation errors. We illustrated the analysis methodology by defining a Cu(100) surface structure and interaction model, calculating the intensities that are determined by that model, adding Gaussian measurement errors to the calculated intensities, and then attempting to recover the surface structure from the simulated data, while characterizing the error levels and correlations. Finally, we compared our analysis to others used for structure analysis via LEED, LEPD, and x-ray crystallography to illuminate the similarities and differences between the various procedures. We believe that the procedure described in this paper affords a simple, meaningful assessment of the uncertainties in the structural parameters extracted from LEED and LEPD intensity data induced by errors in the measured intensities. The procedure is not perfect: it does not describe the consequences of systematic errors due to model imperfections. Nevertheless, it does provide a procedure for LEED and LEPD structure determinations based on statistical error analysis methods; it is simple; and it is at the current state-of-the-art in relating the mathematics of statistical error analysis to the craft of structure analysis via LEED and LEPD.

ACKNOWLEDGMENTS

The authors are indebted to S. Jeyadev for numerous conversations and references concerning statistical analysis of the experimental measurements and to M. S. Warthman for her splendid preparation of this manuscript.

APPENDIX A: PARAMETER ESTIMATION ERROR COVARIANCE

The linear covariance theory for parameter estimation errors provides an expression for the covariance matrix that describes parameter estimation errors that result from least-squares estimation using data, d, which is linearly related to the parameters of interest, m, and is corrupted by random Gaussian errors, e, such that

$$d = Gm + e \ . \tag{A1}$$

In a region around an approximate structure, a_0, the calculation of LEED intensities, I, can be linearized

$$d \leftarrow [I^{\text{exp}} - I^{\text{th}}(a_0)] \ ,$$

$$m \leftarrow [a - a_0] \ ,$$

$$G \leftarrow \left[\frac{\partial I^{\text{th}}}{\partial \mathbf{a}} \right]_{a_0} \ ,$$

to yield a data/model parameter relationship, $d = Gm + e$,

$$[I^{\text{exp}} - I^{\text{th}}(a_0)] = \left[\frac{\partial I^{\text{th}}}{\partial a} \right]_{a_0} [a - a_0] + e \ , \tag{A2}$$

whose least-squares solution, $m = Md$, $M = (G^T G)^{-1} G^T$, has an error covariance,

$$[\text{cov } m] = M [\text{cov } d] M^T \ , \tag{A3}$$

where [cov d] is the covariance of the Gaussian errors e. Since the covariance of the whole value parameter estimates, [cov a], is equivalent to the covariance of the parameter estimation error [cov m],

$$[\text{cov } a] = M [\text{cov } d] M^T \ . \tag{A4}$$

For the case in which the data errors, e, are uncorrelated and of constant variance, σ_e^2, that is

$$[\text{cov } d] = \sigma_e^2 I \ , \tag{A5}$$

where I is an identity matrix, the covariance expression simplifies to

$$[\text{cov } a] = \sigma_e^2 (G^T G)^{-1} \ . \tag{A6}$$

F times the Gaussian R factor, where F is the number of data points less the number of parameters being estimated,

$$FR_g = [I^{\text{exp}} - I^{\text{th}}(a)]^T [I^{\text{exp}} - I^{\text{th}}(a)] \ , \tag{A7}$$

can be expressed in terms of the residual

$$[I^{\text{exp}} - I^{\text{th}}(a)] = [I^{\text{exp}} - I^{\text{th}}(a_0)] - [I^{\text{th}}(a) - I^{\text{th}}(a_0)]$$
$$= d - Gm \ , \tag{A8}$$

where $m = a - a_0$. Evaluating the second partial derivatives of FR_g with respect to the parameters a at the minimum R_{g0} yields the matrix equation

$$\left[\frac{F}{2} \frac{\partial^2 R_g}{\partial \mathbf{a}^2} \right]_0 = G^T G \ . \tag{A9}$$

The covariance expression can, therefore, be written as

$$[\text{cov } a] = \sigma_e^2 \left[\frac{F}{2} \frac{\partial^2 R_g}{\partial \mathbf{a}^2} \right]_0^{-1} \ , \tag{A10}$$

and, by substituting R_{g0} for σ_e^2 as justified in Sec. II, rewritten as

$$[\text{cov } a] = \frac{2R_{g0}}{F} \left[\frac{\partial^2 R_g}{\partial \mathbf{a}^2} \right]_0^{-1} \ . \tag{A11}$$

In the absence of a theory of absolute intensities, a scale factor c is included in the R factor

$$R_g = \frac{1}{F} [I^{\text{exp}} - cI^{\text{th}}(a)]^T [I^{\text{exp}} - cI^{\text{th}}(a)] \ , \tag{A12}$$

where F is now reduced by 1 since an additional parameter is being estimated, and the linear relationship between perturbations in the data and perturbations in the model parameters becomes

$$[I^{\text{exp}} - c_0 I^{\text{th}}(a_0)] = c_0 \left[\frac{dI^{\text{th}}}{da} \right] [a - a_0]$$
$$+ I^{\text{th}}(a_0)(c - c_0) + e \ , \tag{A13}$$

where c_0 is the R-factor minimizing scale factor defined in Eq. (7) and c is the true ratio of noise-free experimental intensities to intensities calculated by a perfect theory for the correct structure.

The augmented vector of unknowns is

$$m' = \begin{bmatrix} a - a_0 \\ c - c_0 \end{bmatrix} \ ,$$

and the augmented G matrix is

$$G' = \left[c \frac{dI^{\text{th}}}{da} \middle| I^{\text{th}} \right] \ .$$

Least-squares solution leads to the covariance expression,

$$[\text{cov } a] = \sigma_e^2 (G'^T G')^{-1} \ , \tag{A14}$$

analogous to Eq. (6). One half of the matrix of second partial derivatives of FR_g at the minimum R_{g0} becomes

$$\left[\frac{F}{2} \frac{\partial^2 R_g}{\partial \mathbf{a}'^2} \right]_0 = G'^T G' \ , \tag{A15}$$

where a' represents an augmented vector that includes c, and the covariance expression, therefore, retains the form given by Eq. (A10), with a replaced by a',

$$[\text{cov } a'] = \frac{2R_{g0}}{F}\left[\frac{\partial^2 R_g}{\partial \mathbf{a}'^2}\right]_0^{-1} . \quad \text{(A16)}$$

Equation (A16) is the desired result: i.e., the explicit derivation of Eqs. (12) in the text from the combined assumptions of a linear response in I to small variations in $\mathbf{a}$ in the vicinity of $\mathbf{a}_{\text{est}}$ and of uncorrelated Gaussian noise in I.

APPENDIX B: FITTING THE R-FACTOR SURFACE TO A QUADRATIC FORM

The curvature or Hessian matrix,

$$H = \left[\frac{\partial^2 R_g}{\partial \mathbf{a}^2}\right] , \quad \text{(B1)}$$

with elements $[H]_{ij} = \partial^2 R_g / \partial a_i \partial a_j$ used to calculate the error covariance of the parameter estimates is determined by fitting a set of R-factor values, $R_g(a_k)$, for structures, $\mathbf{a}_k$, to a quadratic form

$$R_g(\mathbf{a}_k) = R_g(\mathbf{a}) + [\mathbf{a}_k - \mathbf{a}]^T \nabla R_g(\mathbf{a}) + \tfrac{1}{2}[\mathbf{a}_k - \mathbf{a}]^T [H][\mathbf{a}_k - \mathbf{a}] , \quad \text{(B2)}$$

where $\mathbf{a}$ specifies the parameter values of an approximate structure presumed close to the R-factor minimizing structure, $\mathbf{a}_{\text{est}}$, and the gradient, $\nabla R_g(\mathbf{a})$, is the vector of derivatives $\partial R_g / \partial a_i$ at $\mathbf{a}$. Since for each $\mathbf{a}_k$ the parameter perturbations, $\mathbf{a}_k - \mathbf{a}$, (relative to the reference structure, $\mathbf{a}$) and the R-factor values, $R_g(\mathbf{a}_k)$, [as well as the reference value, $R_g(\mathbf{a})$] can be simply calculated, the unknowns in the set of equations (B2) are the elements of the gradient vector, $\nabla R_g(\mathbf{a})$, and of the partial derivative matrix, H. The set of Eq. (B2) can be recast as a set of linear equations (one for each structure, $\mathbf{a}_k$) in matrix form in which a vector of R-factor variations are related to a vector of unknown partial derivatives through a matrix of parameter perturbations. Solution of the linear equations yields values for both the gradient vector and curvature matrix.

First the set of structures, $\mathbf{a}_k$, is specified. For an M parameter problem, an M-dimensional grid of structures in the vicinity of $\mathbf{a}$ is specified by identifying a step size, Δ_{a_i}, and a range of parameter values, $a_i \pm k_i \Delta a_i$, for $k_i = 0, 1, \ldots, k_{i,\max}$ for each variable parameter a_i. $R_g(\mathbf{a}_k)$ is calculated for each of the $K = \prod_i (2k_{i,\max}+1)$ structures, $\mathbf{a}_k$, on the specified grid. A vector of length K with elements

$$Y_k = R_g(\mathbf{a}_k) - R_g(\mathbf{a}) , \quad \text{(B3)}$$

is constructed and a vector $\mathbf{C}$ of length $L = M(M+3)/2$ is defined with elements corresponding to the unknowns, i.e., the M elements of the gradient vector, $\nabla R_g(\mathbf{a})$, and the $M(M+1)/2$ distinct elements of the symmetric partial derivative matrix, H. A $K \times L$ matrix A of first- and second-order parameter perturbations is formed with elements

$$[A]_{kl} = \begin{cases} (a_{k,l} - a_l) & \text{if } 1 \le l \le M \\ (a_{k,p} - a_p)(a_{k,q} - a_q) & \text{if } M+1 \le l \le M + M(M+1)/2 , \end{cases} \quad \text{(B4)}$$

such that each row corresponds to a specific structure, $\mathbf{a}_k$. The set of K equations in L unknowns displayed in Eq. (B2) is rewritten as

$$\mathbf{Y} = \mathbf{AC} . \quad \text{(B5)}$$

Since K is always greater than L by choice, Eq. (B5) is an overdetermined system of equations; its least-squares solution, $\mathbf{C}$ (determined using the NAG library routine F04AMF), specifies the values of the elements of the gradient vector, $\nabla R_g(\mathbf{a})$, and the curvature matrix, H. The number of points, $K = \Pi_i (2k_{i,\max}+1)$, in the grid (as determined by the extent, $2k_{i,\max}$, of the grid in each dimension, $k = 1, \ldots, M$) must be chosen to be sufficiently large relative to the number, L, of unknown derivatives being estimated to force to a low level the errors in the partial derivative estimates caused by roughness in the R-factor surface, due to noise in the intensities. For data with low noise, the R-factor surface can usually be fit using a grid of structures of length 3 in each dimension (i.e., $k_{i,\max} = 1$ for each parameter, a_i); in this case, the number, K, of points in the grid (i.e., the number of structures for which the R factor is evaluated) is 3^M. The ratio which characterizes the extent of overdetermination

$$\frac{K}{L} = \frac{3^M}{M(M+3)/2} , \quad \text{(B6)}$$

is thus 9/5 for two variable parameters (one structural parameter and one scaling parameter) and larger for analyses that involve more variable parameters.

Once the curvature matrix and gradient have been determined, the quality of the quadratic fit is evaluated by calculating the root-mean-square deviation, $\sigma_R = \sqrt{\sigma_R^2}$, between the calculated R factors, $R_g(\mathbf{a}_k)$, and the fit R-factor function, where

$$\sigma_R^2 = \sum_k \left[R_g(\mathbf{a}_k) - \left[R_g(\mathbf{a}) + [\mathbf{a}_k - \mathbf{a}]^T \nabla R_g(\mathbf{a}) + \tfrac{1}{2}[\mathbf{a}_k - \mathbf{a}]^T [H][\mathbf{a}_k - \mathbf{a}] \right] \right]^2 . \quad \text{(B7)}$$

If the fit is adequate ($\sigma_R \ll 0.1\max[R_g(\mathbf{a}_k)-R_g(\mathbf{a})]$), the curvature matrix can be inverted and used to calculate the parameter estimation covariance as prescribed by Eq. (12b). Furthermore, the R-factor function can be used to determine the correction, $\mathbf{a}_{\text{est}}-a$, to the approximate structure, $\mathbf{a}$, by requiring the gradient

$$\nabla R_g(\mathbf{a}')=\nabla R_g(\mathbf{a})+[H][\mathbf{a}'-\mathbf{a}] \tag{B8}$$

at the R-factor minimizing structure, $\mathbf{a}_{\text{est}}$, to equal 0. Then, the correction is determined analytically

$$\mathbf{a}_{\text{est}}-\mathbf{a}=-[H]^{-1}\nabla R_g(\mathbf{a}) \tag{B9}$$

and can be added to the approximate structure, $\mathbf{a}$, to yield an improved estimate, $\mathbf{a}_{\text{est}}$, of the unknown structural parameters. The intensities and the R factor are calculated for the refined structure. If the refined structure is outside the range of the original grid, the calculated R factor is compared with the value predicted by the R-factor function in order to confirm the validity of the quadratic form for the R factor in the region which includes not only the original grid points, but also the refined structure, $\mathbf{a}_{\text{est}}$.

As the primary function of the fitting procedure is to determine the curvature matrix in order to calculate the parameter estimation error covariance, it is preferable to initiate the procedure with a reference structure, $\mathbf{a}$, determined by a well-converged parameter search so as to ensure that a grid of modest extent around the reference structure includes the R-factor minimizing structure. "Modest extent" in this context means small relative to the parameter perturbations whose signature in the intensity domain can be adequately described by the first-order perturbation theory. It is within such a range that the curvature matrix can be expected to be constant and the R-factor surface properly modeled by the quadratic expression presented in Eq. (B2).

[1] W. Menke, *Geophysical Data Analysis: Discrete Inverse Theory,* International Geophysics Series (Academic, San Diego, 1989), Vol. 45, p. 1.

[2] M. A. van Hove, W. H. Weinberg, and C.-M. Chan, *Low-Energy Electron Diffraction,* Springer Series in Surface Sciences Vol. 6 (Springer-Verlag, Berlin, 1986), p. 1.

[3] K. F. Canter, C. B. Duke, and A. P. Mills, Jr., in *Chemistry and Physics of Solid Surfaces VII,* edited by R. Vanselow and R. Howe, Springer Series in Surface Sciences Vol. 22 (Springer-Verlag, Berlin, 1990), p. 183.

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

[5] X. M. Chen, K. F. Canter, C. B. Duke, A. Paton, D. L. Lessor, and W. K. Ford, Phys. Rev. B **48**, 2400 (1993).

[6] F. Jona, D. Sondericker, and P. M. Marcus, J. Phys. C **13**, L155 (1980).

[7] C. B. Duke, Appl. Surf. Sci. **11/12**, 1 (1982).

[8] H. L. Davis and J. R. Noonan, J. Vac. Sci. Technol. **20**, 842 (1982).

[9] H. L. Davis and J. R. Noonan, Surf. Sci. **126**, 245 (1983).

[10] R. Mayer, C.-S. Zhang, K. G. Lynn, W. E. Frieze, F. Jona, and P. Marcus, Phys. Rev. B **35**, 3102 (1987).

[11] T. N. Horsky, G. R. Brandes, K. F. Canter, C. B. Duke, A. Paton, A. Kahn, S. F. Horng, K. Stevens, K. Stiles, and A. P. Mills, Jr., Phys. Rev. B **46**, 7011 (1992).

[12] C. B. Duke, Adv. Chem. Phys. **27**, 1 (1974).

[13] K. Heinz, Appl. Phys. A **41**, 3 (1986).

[14] C. W. Tucker, Jr. and C. B. Duke, Surf. Sci. **29**, 237 (1972).

[15] E. Zanazzi and F. Jona, Surf. Sci. **62**, 61 (1977).

[16] M. A. van Hove, S. Y. Tong, and M. H. Elconin, Surf. Sci. **64**, 85 (1977).

[17] J. B. Pendry, J. Phys. C **13**, 937 (1980).

[18] C. B. Duke, S.L.Richardson, A. Paton, and A. Kahn, Surf. Sci. **127**, L135 (1983).

[19] P. G. Cowell, M. Prutton, and S. P. Tear, Surf. Sci. **177**, L915 (1986).

[20] G. Kleinle, W. Moritz, D. L. Adans, and G. Ertl, Surf. Sci. **219**, L637 (1989).

[21] G. Kleinle, W. Moritz, and G. Ertl, Surf. Sci. **238**, 119 (1990).

[22] B. Schwarzenbach *et al.*, Acta. Crystallogr. Sec. A **45**, 63 (1989).

[23] C. B. Duke, N. O. Lipari, and U. Landman, Phys. Rev. B **6**, 2454 (1973).

[24] P. J. Rous, Prog. Surf. Sci. **39**, 3 (1992).

[25] A. H. Weiss, I. J. Rosenberg, K. F. Canter, C. B. Duke, and A. Paton, Phys. Rev. B **27**, 867 (1983).

[26] G. Capart, Surf. Sci. **26**, 429 (1971).

[27] J. B. Pendry, J. Phys. C **4**, 2514 (1971).

[28] G. E. Laramore, Phys. Rev. B **9**, 1204 (1974).

[29] W. K. Ford, T. Guo, D. L. Lessor, and C. B. Duke, Phys. Rev. B **42**, 8952 (1990).

[30] G.E. Laramore and C. B. Duke, Phys. Rev. B **5**, 267 (1972).

[31] F. Herman and S. Skillman, *Atomic Structure Calculations* (Prentice-Hall, Englewood Cliffs, NJ, 1963).

[32] C. B. Duke, N. O. Lipari, and U. Landman, Phys. Rev. B **8**, 2454 (1973).

[33] R. J. Meyer, C. B. Duke, A. Paton, A. Kahn, E. So, J. L. Yeh, and P. Mark, Phys. Rev. B **19**, 5194 (1979).

[34] *NAG Fortran Library Introductory Guide* (Numerical Algorithms Group, Oxford, 1991), Mark 15, Chap. G05.

[35] D. L. Adams, V. Jensen, X. F. Sun, and J. H. Vollesen, Phys. Rev. B **38**, 7913 (1988).

[36] P. R. Watson, F. R. Sheperd, D. C. Frost, and K. A. R. Mitchell, Surf. Sci. **72**, 562 (1978).

[37] D. L. Adams, H. B. Nielsen, and M. A. Van Hove, Phys. Rev. B **20**, 4789 (1979).

[38] E. Prince, *Mathematical Techniques in Crystallography and Materials Science* (Springer-Verlag, New York, 1982), pp. 77–118.

[39] J. S. Rollett, in *Methods and Applications in Crystallographic Computing,* edited by S. R. Hall and T. Ashida (Oxford University Press, Oxford, 1984), pp. 161–173.

[40] R. D. Diehl, M. Lindroos, A. Kearsley, C. J. Barnes, and D. A. King, J. Phys. C **18**, 4069 (1985).

[41] P. G. Cowell and V. E. de Carvalho, Surf. Sci. **187**, 175 (1987).

[42] A. A. Lazarides, C. B. Duke, A. Paton, and A. Kahn, Phys. Rev. B **52**, 14 895 (1995).

[43]D. L. Adams and H. B. Nielsen, Surf. Sci. **107**, 305 (1981).
[44]H.B. Nielsen and D. L. Adams, J. Phys. C **15**, 615 (1982).
[45]D. L. Adams, H. B. Nielsen, and J. N. Andersen, Surf. Sci. **128**, 294 (1983).
[46]J. N. Andersen, H. B. Nielsen, L. Petersen, and D. L. Adams, J. Phys. C **17**, 173 (1984).
[47]P. R. Bevington, *Data Reduction and Error Analysis in the Physical Sciences* (McGraw-Hill, New York, 1969).