Bayes-Turchin approach to x-ray absorption fine structure data analysis

H. J. Krappe and H. H. Rossner
Hahn-Meitner-Institut Berlin, Glienicker Strasse 100, D-14109 Berlin, Germany
(Received 25 July 2002; published 19 November 2002)

X-ray absorption fine structure (XAFS) data from copper, gold, and germanium are analyzed in the framework of the model-independent Bayes-Turchin approach. Compared to earlier treatments, we also obtain spring constants, besides shell radii, Debye-Waller (DW) parameters, and anharmonicity parameters. The use of spring constants instead of DW parameters reduces considerably the number of model parameters needed to achieve a satisfactory fit of the data. The *ab initio* extended XAFS code FEFF7 is used in the analysis. The various sources of uncertainty in the input data and in the FEFF7 code are carefully assessed and used in the analysis. It is shown to which degree the model parameters are determined by the data, rather than by the *a priori* assumptions. *A posteriori* errors and error correlations between model parameters are shown.

DOI: 10.1103/PhysRevB.66.184303 PACS number(s): 07.05.Tp, 61.10.Ht, 87.64.Fb

I. INTRODUCTION

Efficient codes are available to calculate the x-ray absorption cross section $\mu(k)$ approximately as a function of the x-ray wave vector k and the following model parameters: the half length of each single- or multiple-scattering path j, R_j, the Debye-Waller (DW) parameter σ_j for each scattering path, and the third cumulants $C_{3,j}$.[1] Data fitting, however, requires the solution of the inverse problem of inferring the model parameters from measured absorption coefficients. In general, there are more model parameters than data, but even if one restricts the number of model parameters to the number of available data points, the inversion problem turns out to be ill posed. A further restriction of the number of model parameters considered in the fit is therefore necessary. Intuitively guided by the so-called number of independent data points the list of model parameters is often cut down until the χ^2 fitting procedure appears to be numerically stable. Considering the linear space $\mathcal{Q}$ spanned by all model parameters, the implicit assumption is therefore made that its subspace $\mathcal{R}$, which is determined by the data, can be spanned by a finite subset of the model parameters. However, one should expect that in general the subspace $\mathcal{R}$ is spanned by orthogonal coordinates, which are oblique to the model parameters spanning the parameter space $\mathcal{Q}$. One would therefore like to have an algorithm that determines on the basis of the measured data not only the dimension of $\mathcal{R}$, but its orientation in $\mathcal{Q}$ as well.

To solve this problem we introduced in Ref. 2 the concept of an *a priori* guess of the set of model parameters and reformulated the task of fitting by looking for the shift away from the *a priori* model parameters, as required by the data. Invoking the maximum-entropy principle, the ill-posed inversion problem can then be regularized.[3] To define the procedure completely, one still has to fix the variance matrix of the *a priori* model parameters, which, in particular, determines the weight with which the *a priori* information influences the fit relative to the weight of the experimental data. We used in Ref. 2 two alternative optimization conditions, first proposed by Turchin *et al.*,[4] to obtain one overall weight parameter. It turns out that one of Turchin's conditions admits a generalization, such that more details of the *a priori* variance matrix can be determined. Below we shall use this possibility to get the relative weight for several groups of *a priori* model parameters independently.

In Ref. 2 we obtained half length parameters R_j for each scattering path independently. In order to account for geometrical constraints between these parameters, we shall take the radii of the shells surrounding the emitting atom as independent parameters and represent the half lengths of the multiple-scattering paths as functions of these radii, assuming that there are no angular distortions in the lattice. Though this is not the most general situation, we consider this simplification as appropriate for the applications we will investigate.

The DW parameters contain, in general, contributions from thermal motion of the lattice atoms and their quantum fluctuations, $\sigma_j^{\mathrm{therm}}$, and from lattice disorder, $\sigma_j^{\mathrm{disor}}$, $\sigma_j^2 = (\sigma_j^{\mathrm{therm}})^2 + (\sigma_j^{\mathrm{disor}})^2$. The thermal part can be represented more significantly in terms of a few spring constants in the neighborhood of the emitting atom,[5,6] which reduces the number of model parameters substantially without deteriorating the quality of the fit. If measurements of the same probe exist at different temperatures, one may separate the two contributions, assuming that the $\sigma_j^{\mathrm{disor}}$ are independent of the temperature. In favorable cases they can be negligible compared to $\sigma_j^{\mathrm{therm}}$. Such data can then be analyzed directly in terms of a few spring constants instead of a DW parameter for each scattering path.

In some cases third cumulants $C_{3,j}$ can be extracted from the data.[7] We have therefore always included a search for these parameters at least for the single-scattering paths. In a simple force-field model anharmonicity parameters are also determined from these third cumulants.

The paper is organized as follows. In the following section the input and output parameters of the fitting procedure and their probability distributions are defined and the models are discussed on which the analysis is based. The algorithm used for the calculation of the DW parameters is presented in some detail in Sec. III. The various sources of uncertainty affecting the fit are discussed and modeled in Sec. IV. In Sec. V the essential elements of an extended χ^2 fit and its regularization by *a priori* assumptions are summarized, general-

izing the presentation in Ref. 2. Iterative procedures to solve the resulting nonlinear systems of equations are described in Sec. VI. Section VII contains the application of the proposed method to the extended x-ray absorption fine structure (EXAFS) analysis of *K*-edge copper data, taken at three different temperatures,[8,9] L_3-edge gold data, measured at 300 K,[8] and germanium *K*-edge data, taken at 300 K.[8] A summary of the procedure and an outlook on further developments is presented in Sec. VIII.

II. DEFINITION OF THE PROBLEM

The following discussion will be based on the multiple-path expansion for x-ray absorption on a polycrystalline or amorphous sample:[10,11]

$$\chi(k)=\frac{\mu(k)-\mu_{\rm back}(k)-\mu_0}{\mu_0}$$
$$=\frac{S_0^2}{k}\sum_j N_j\frac{|f_j(k,R_j)|}{R_j^2}e^{-2k^2\sigma_j^2-2R_j/\lambda(k)}\sin\Big[2k(R_j-\delta R_j)+\phi_j(k)-\frac{4}{3}C_{3,j}k^3\Big], \quad (1)$$

with the wave number

$$k^2=\frac{2m}{\hbar^2}(h\nu-E_0)$$

and corrections to the lengths

$$\delta R_j=2\sigma_j^2\left(\frac{1}{R_j}+\frac{1}{\lambda}\right),$$

where the sum in Eq. (1) runs over all geometrically inequivalent single- and multiple-scattering paths *j*, the multiplicity of equivalent paths being counted by N_j, which is the coordination number for the single-scattering paths.

The overall amplitude S_0, the mean free path $\lambda(k)$, the scattering amplitudes $f_j(k,R_j)$, and the phases $\phi_j(k)$ for each scattering path follow from a solution of the electronic many-body problem as functions of k and the lattice geometry. In the framework of the local-density approximation together with the muffin-tin ansatz for the scattering potential, the FEFF7 code[1] provides approximate values for f, ϕ, and λ as functions of k and R_j. The recursion procedure proposed in Ref. 6 similarly allows calculation of $\sigma_j^{\rm therm}$ approximately as functions of the temperature and a few spring constants κ_s. Taking into account the reduction of the half lengths R_j of the multiple-scattering paths to those of the single-scattering paths mentioned above, we have the following as independent model parameters: the shell radii R_j, the structural disorder contributions $\sigma_j^{\rm disor}$ for all scattering paths and/or a set of spring constants κ_s (whatever applies in a specific case), and the third cumulants $C_{3,j}$ for each single-scattering path. Since S_0 cannot be obtained reliably from FEFF and since the convergence of Eq. (1) becomes poor for energies close to the edge energy E_0, which is, in addition, particularly subject to uncertainties connected with the muffin-tin approach, we will follow the usual practice and treat these two quantities also as independent model parameters.

The multiplicity numbers N_j are, in principle, also model parameters. Since we want to analyze data taken on well-ordered probes, we assume that these parameters have their ideal lattice values, and do not determine them in the fit. In the case of less well-ordered samples, this would not be appropriate and extra measures are needed to disentangle the very strong correlation between the N_j and σ_j in Eq. (1).

The absorption coefficient of the free absorbing atom $\mu_0(k)$ can, in principle, be calculated in the framework of the Hartree-Dirac approach. However, we feel that computer codes available to us at present for that purpose are not yet sufficiently fast and, at the same time, not accurate enough to treat μ_0 in Eq. (1) as a derived quantity. Since we use in some cases the EXAFS function $\chi(k)$ rather than $\mu(k)$ as input, we will follow the common practice of applying an empirical background-subtraction procedure to the measured $\mu(k)$ to obtain the EXAFS signal $\chi(k)$,[12,13] in cases where we start the analysis with the absorption coefficient. The EXAFS signal may be measured at L data points with wave numbers k_l, $l=1,\ldots,L$. The input of the fitting procedure therefore consists of the L data $\chi(k_l)$.

For each independent model parameter x_n' we introduce a corresponding *a priori* estimate $x_n^{(0)}$, e.g., for R_i an *a priori* value $R_i^{(0)}$, and the differences between the actual parameter values and their *a priori* values, $x_n'-x_n^{(0)}$, are treated as the new independent variables to be determined by the fit. It is useful to normalize these differences by quantities $\hat{x}_n$ which are of the order of the expected size of $x_n'-x_n^{(0)}$ in the final fit, so that the normalized model parameters $x_n=(x_n'-x_n^{(0)})/\hat{x}_n$ are dimensionless and will always be of a limited order of magnitude. The x_n shall be the components of the vector $\mathbf{x}$ in the N-dimensional model-parameter space $\mathcal{Q}$. The components shall be arranged according to the order: $S_0^2, E_0, R_i\,(i=1,\ldots,I), \kappa_s\,(s=1,\ldots,S), \sigma_j^{\rm disor}\,(j=1,\ldots,J), C_{3,i}\,(i=1,\ldots,I)$, where I is the number of shells, S the number of spring constants, and J the number of scattering paths considered in the fit; therefore $N=2I+J+S+2$. To simplify the notation, it is convenient to consider the dependent quantities which appear on the right-hand side of Eq. (1), $\lambda(k)$, $f_j(k)$, $\phi_j(k)$, and $\sigma_j^{\rm therm}(\kappa)$ as components of a vector $\mathbf{y}(\mathbf{k},\boldsymbol{\kappa})$. The number of these components will be designated in the following by $M=3J+1$.

If one would choose the number of data points L equal to the number of model parameters N, it may appear as if Eq. (1) yields just L algebraic equations for the same number of unknown model parameters. However, the input does not consist of L numbers, but rather L probability distributions for the $\chi(k_l)$ since the original data themselves as well as the background subtraction involve uncertainties. Therefore the output of the fit can also consist only of probabilities for the model parameters. This is even more true since the evaluation of $\mathbf{y}(k_l,\kappa_s)$ is only possible in some approximate scheme, producing additional uncertainties. Moreover, the *j*

sum in Eq. (1) has to be truncated at the term $j=J$, and the number of shells I must be limited to those of a finite cluster surrounding the absorbing atom. These truncations also contribute to the uncertainties of the fit. Equation (1) has therefore to be seen as defining a stochastic, rather than an algebraic problem.

III. LATTICE DYNAMICS

In order to represent the $\sigma_j^{\mathrm{therm}}$ in terms of a few spring constants, we use the method proposed by Poiarkova and Rehr.[6] For the sake of completeness of the presentation, and since we shall slightly extend the method, it may be useful to summarize the main steps of its derivation using the notation introduced in Refs. 5 and 6 as far as compatible with the present context.

A. Representation of DW parameters in terms of the projected density of states

We will consider a cluster of I shells surrounding the absorbing atom, each containing N_i lattice points, so that the total number of lattice points in the cluster is

$$Z=1+\sum_{i=1}^{I} N_i .$$

The lattice dynamics of the cluster is determined by the displacement vectors $\mathbf{u}'_\zeta$, $\zeta=1,\ldots,Z$ of each lattice atom. Scaling them by the mass M_ζ of the atom at site ζ, $\mathbf{u}_\zeta=\mathbf{u}'_\zeta\sqrt{M_\zeta}$, the equations of motion, in harmonic approximation, become

$$\ddot{u}_{\zeta,\varphi}=-\sum_{\zeta',\varphi'} \Phi_{\zeta,\varphi;\zeta',\varphi'} u_{\zeta',\varphi'} ,$$

where φ runs over the three Cartesian components of the vector $\mathbf{u}_\zeta$, and for central forces the dynamical matrix Φ^{sm} for the bond-streching modes is given in terms of the spring constants $\kappa_{\zeta,\zeta'}$ and the directional unit vectors $\hat{\mathbf{r}}_{\zeta,\zeta'}$ between lattice sites ζ and ζ' by[6]

$$\Phi^{\mathrm{sm}}_{\zeta,\varphi;\zeta',\varphi'}=\frac{1}{\sqrt{M_\zeta M_{\zeta'}}}\left(\delta_{\zeta,\zeta'}\sum_{\zeta''}\kappa_{\zeta'',\zeta}\hat{r}^{\varphi}_{\zeta'',\zeta}\hat{r}^{\varphi'}_{\zeta'',\zeta}-\kappa_{\zeta,\zeta'}\hat{r}^{\varphi}_{\zeta,\zeta'}\hat{r}^{\varphi'}_{\zeta,\zeta'}\right).$$

In practice, spring constants between distant points are assumed to be zero. In applications to be considered below, only forces between next and nearest next neighbors are kept, which are then labeled $\kappa_s, s=1,2$.

For lattices of the diamond type, we add bond-bending modes. Following Keating,[14] we parametrize their contribution to the potential energy by

$$V_{\mathrm{bm}}=\frac{\kappa_3}{8}\sum_{\zeta=1}^{Z}\sum_{\Delta<\Delta'=1}^{4}[(\mathbf{u}'_\zeta-\mathbf{u}'_{\zeta\Delta})\hat{\mathbf{r}}_{\zeta\Delta'}+(\mathbf{u}'_\zeta-\mathbf{u}'_{\zeta\Delta'})\hat{\mathbf{r}}_{\zeta\Delta}]^2,$$

where $\mathbf{u}'_{\zeta\Delta}, \Delta=1,\ldots,4$ are the displacement vectors of the four nearest neighbors of the lattice site ζ, $\hat{\mathbf{r}}_{\zeta\Delta}$ is the directional unit vector from the equilibrium site ζ to the site $\zeta+\Delta$, and the inner sum is over the six tetrahedron angles subtended by each pair of nearest-neighbor atoms with the atom ζ at the vertex. The contribution of these modes to the dynamical matrix is given by

$$\Phi^{\mathrm{bm}}_{\zeta,\varphi;\zeta',\varphi'}=\frac{\kappa_3}{4\sqrt{M_\zeta M_{\zeta'}}}\sum_{\Delta\neq\Delta'=1}^{4}(\delta_{\zeta\zeta'}-\delta_{\zeta'\zeta+\Delta})\hat{r}^{\varphi}_{\zeta\Delta},\hat{r}^{\varphi'}_{\zeta\Delta'} .$$

Our parameters κ_1 and κ_3 correspond to Keating's 6α and 6β, respectively.

The eigenvalues of the dynamical matrix $\mathbf{\Phi}=\mathbf{\Phi}^{\mathrm{sm}}+\mathbf{\Phi}^{\mathrm{bm}}$ yield the frequencies $\omega_\gamma, \gamma=1,\ldots,3Z$, of the normal modes of the cluster

$$\sum_{\zeta,\zeta'=1}^{Z}\sum_{\varphi,\varphi'=1}^{3}\epsilon^{*}_{\gamma;\zeta,\varphi}\Phi_{\zeta,\varphi;\zeta',\varphi'}\epsilon_{\zeta',\varphi';\gamma}=\omega_\gamma^2\delta_{\gamma\gamma'} ,$$

where ϵ is the unitary $3Z\times 3Z$ matrix which diagonalizes the dynamical matrix Φ. In terms of its eigenvectors q_γ, the displacement vectors are

$$u_{\zeta,\varphi}=\sum_{\gamma=1}^{3Z}\epsilon_{\zeta,\varphi;\gamma}q_\gamma . \tag{2}$$

A scattering path j with n_j legs is defined as a cyclic sequence of n_j lattice points, starting and ending with the site of the absorbing atom: $\zeta_1,\ldots,\zeta_{n_j},\zeta_{n_j+1}$, where $\zeta_1\equiv\zeta_{n_j+1}$ is the absorbing atom. The thermal Debye-Waller parameter of the scattering path j is given in terms of the $\mathbf{u}'_\zeta$ by[5]

$$\sigma_j^2=\left\langle\left[\sum_{i=1}^{n_j}\sum_{\varphi=1}^{3}u'_{\zeta_i,\varphi}(\hat{r}^{\varphi}_{\zeta_i,\zeta_{i-1}}+\hat{r}^{\varphi}_{\zeta_i,\zeta_{i+1}})/2\right]^2\right\rangle,$$

where the brackets indicate a canonical ensemble average and the index ''therm'' has been suppressed. Using Eq. (2), this can be rewritten as

$$\sigma_j^2=\sum_\gamma\left|\sum_{i=1}^{n_j}\sqrt{\frac{1}{M_{\zeta_i}}}\sum_{\varphi=1}^{3}\frac{\hat{r}^{\varphi}_{\zeta_i,\zeta_{i-1}}+\hat{r}^{\varphi}_{\zeta_i,\zeta_{i+1}}}{2}\epsilon_{\zeta_i,\varphi;\gamma}\right|^2\langle q_\gamma^2\rangle,$$

where $\langle q_\gamma q_{\gamma'}\rangle=0$ for $\gamma\neq\gamma'$ has been used. A more compact notation of this expression is obtained by introducing the $3Z$-component vectors

$$|\gamma\rangle=|\epsilon_{1,1;\gamma},\epsilon_{1,2;\gamma},\ldots,\epsilon_{Z,3;\gamma}\rangle$$

and $|0\rangle$, whose ζ,φ components are given by $(1/2)\sqrt{\mu_j/M_{\zeta_i}}(\hat{r}^{\varphi}_{\zeta_i,\zeta_{i-1}}+\hat{r}^{\varphi}_{\zeta_i,\zeta_{i+1}})$ if ζ equals one of the n_j nodes ζ_i of the scattering path j and is zero else. The mass μ_j is defined by

$$\frac{1}{\mu_j}=\sum_{i=1}^{n_j}\frac{1}{M_{\zeta_i}}\sum_{\varphi=1}^{3}\left(\frac{\hat{r}^{\varphi}_{\zeta_i,\zeta_{i-1}}+\hat{r}^{\varphi}_{\zeta_i\zeta_{i+1}}}{2}\right)^2,$$

so that $|0\rangle$ is normalized to 1. The DW parameter for path j becomes, in terms of these vectors,

$$\sigma_j^2=\frac{1}{\mu_j}\sum_{\gamma=1}^{3Z}|\langle\gamma|0\rangle|^2\langle q_\gamma^2\rangle.$$

Bose-Einstein statistics yields for the canonical ensemble with temperature T

$$\omega_\gamma^2\langle q_\gamma^2\rangle=\langle n(\omega_\gamma)+1/2\rangle\hbar\omega_\gamma=\frac{\hbar\omega_\gamma}{2}\coth\frac{\hbar\omega_\gamma\beta}{2},$$

where $1/\beta=k_BT$. The DW parameter is therefore

$$\sigma_j^2=\frac{\hbar}{2\mu_j}\int_0^\infty\frac{1}{\omega}\coth\frac{\beta\hbar\omega}{2}\rho_j(\omega)d\omega, \tag{3}$$

with the projected density of states[5]

$$\rho_j(\omega)d\omega=\sum_\gamma|\langle 0|\gamma\rangle|^2\delta(\omega-\omega_\gamma)d\omega$$
$$=-\frac{1}{\pi}\mathrm{Im}\left\langle 0\left|\frac{1}{z-\Phi+i\epsilon}\right|0\right\rangle dz:=w(z)dz,$$

with $z=\omega^2$.

B. Recursion relations for the density of states

In order to approximate the Green's function $\langle 0|(z-\Phi+i\epsilon)^{-1}|0\rangle$ in an efficient way, one should take into account that the lattice dynamics in the neighborhood of the absorbing atom is predominantly determined by the forces between the atoms in the first few shells and that atoms further away from the absorbing atom have a decreasing effect. This is taken into account in the iteration procedure proposed by Haydock *et al.*[15] Starting from the vector $|0\rangle$, a set of orthonormal vectors $|\nu\rangle$ is generated by the Lanczos iteration[16]

$$a_\nu=\langle\nu|\Phi|\nu\rangle,$$
$$|v_{\nu+1}\rangle=(\Phi-a_\nu)|\nu\rangle-b_\nu|\nu-1\rangle,$$
$$b_{\nu+1}^2=\langle v_{\nu+1}|v_{\nu+1}\rangle,$$
$$|\nu+1\rangle=\frac{|v_{\nu+1}\rangle}{b_{\nu+1}},$$

$\nu=0,1,\ldots,$ with $b_0=1$ and $|-1\rangle=0$. The iteration can be continued until $b_\nu\approx 0$. In the $|\nu\rangle$ representation, the dynamical matrix Φ is easily seen to be tridiagonal with matrix elements

$$\langle\nu|\Phi|\nu'\rangle=\begin{cases}a_\nu, & \nu'=\nu\\ b_{\nu+1}, & \nu'=\nu+1\\ b_\nu, & \nu'=\nu-1\\ 0, & \text{otherwise.}\end{cases}$$

The determinant D_{-1} of the matrix $(z\delta_{\nu\nu'}-\langle\nu|\Phi|\nu'\rangle)$ is introduced together with the determinants D_ν of the matrices obtained from $(z\delta_{\nu\nu'}-\langle\nu|\Phi|\nu'\rangle)$ by removing rows and columns $0-\nu$. Since all these matrices are tridiagonal, the Laplace expansion of D_ν yields the recursion relation

$$D_\nu=(z-a_{\nu+1})D_{\nu+1}-b_{\nu+2}^2D_{\nu+2},$$

with $\nu=-1,0,\ldots,$ or

$$D_{\nu+1}/D_\nu=\frac{1}{z-a_{\nu+1}-b_{\nu+2}^2D_{\nu+2}/D_{\nu+1}}. \tag{4}$$

In terms of these determinants, the Green's function is given by

$$\left\langle 0\left|\frac{1}{z-\Phi}\right|0\right\rangle=\frac{D_0}{D_{-1}}.$$

Repeated application of the recursion relation (4) yields a representation of the Green's function as a continued fraction

$$\left\langle 0\left|\frac{1}{z-\Phi}\right|0\right\rangle=\frac{b_0^2|}{|z-a_0}-\frac{b_1^2|}{|z-a_1}-\cdots.$$

The polynomials R_{N_ν} and S_{N_ν} of the convergent R_{N_ν}/S_{N_ν}, which represents the Green's function when the continued fraction is terminated with the term N_ν, can be obtained from the recurrence relation[17]

$$Y_\nu=(z-a_{\nu-1})Y_{\nu-1}-b_{\nu-1}^2Y_{\nu-2},$$

where Y_ν may be R_ν or S_ν, and the recursion starts with

$$R_0=0,\quad R_1=1,$$
$$S_0=1,\quad S_1=z-a_0.$$

One can show that the S_ν form a set of orthogonal polynomials to the weight function $w(z)$.[18] Therefore they have simple zeros, which lie in the support of $w(z)$. The N_ν zeros of S_{N_ν} may be called z_ν. The residues of the poles of the convergent yield the weights

$$w_\nu=R_{N_\nu}(z_\nu)/S'_{N_\nu}(z_\nu),$$

where S'_ν is the derivative of S_ν with respect to z. In terms of these weights, the integral (3) becomes

$$\sigma_j^2=\frac{\hbar}{2\mu_j}\sum_{\nu=1}^{N_\nu}\frac{w_\nu}{\omega_\nu}\coth\frac{\beta\hbar\omega_\nu}{2},$$

with $\omega_\nu=\sqrt{z_\nu}$. It can also be shown[18] that the w_ν are the weights of the N_ν-point Gauss integration to the weight function $w(z)$.

In order to give an impression of the convergence of the continued fraction representation of the thermal DW parameter, we show in Fig. 1 the DW parameter for the single-scattering path to the first shell in the copper lattice at 295 K with one spring constant, $\kappa_1=27.9$ N/m. A cluster with $I=11$ shells was used in this calculation. As noticed already in Ref. 6, the $N_\nu=2$ approximation underestimates the limiting value. In the following, we shall therefore use $N_\nu=6$ as

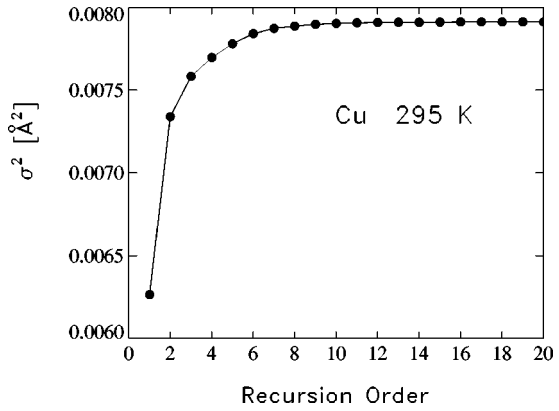

FIG. 1. Convergence of the continued fraction expansion of the thermal Debye-Waller parameter of the first single-scattering path for copper at 295 K in a dynamical model with one spring constant.

a compromise between computational speed and accuracy and accept a model uncertainty of 1.3%.

In Fig. 2 the dependence of the DW parameters in the same copper lattice for the first two scattering paths are plotted as functions of the spring constants in a model with one spring constant (full lines) and two spring constants (dash-dotted lines). In the relevant region, the DW parameter depends only weakly on the spring constants, which shows the uncertainty connected with the inverse problem of determining the spring constant from measured σ's.

C. Anharmonicity in the potential of the Einstein model

Since the third cumulants $C_{3,i}$ are determined with considerably less accuracy than the variances σ_i^2, we do not attempt to relate them to anharmonic terms in the force-field model which we used in the analysis of the variances. Instead, we employ a correlated Einstein model, generalized to contain an anharmonic term as proposed by Frenkel and Rehr.[19] They describe the lattice dynamics by a fictitious vibrational mode of two masses M, interacting via the anharmonic potential

$$V_E(x)=\frac{1}{2}\kappa_E x^2+\varpi_E x^3, \tag{5}$$

where $x=r-r_0$ is the deviation of the bond length r for single-scattering paths from its equilibrium value r_0 at zero temperature. In terms of the Einstein frequency $\omega_E^2=2\kappa_E/M$ and the Boltzmann factor $\xi(T)=\exp(-\hbar\omega_E/k_BT)$, three equations are derived in Ref. 19 relating the first three cumulants of the thermal distribution of x to the thermal expansion $a(T)=\langle r-r_0\rangle$, the Einstein frequency ω_E, and the anharmonicity ϖ_E in the potential (5),

$$a_i(T)=R_i(T)-R_i(0),$$

$$\sigma_i^2=\frac{\hbar}{M\omega_{E,i}}\frac{1+\xi_i}{1-\xi_i}, \tag{6}$$

and

$$C_{3,i}=-\frac{4\hbar^2}{M^3\omega_{E,i}^4}\frac{1+10\xi_i+\xi_i^2}{(1-\xi_i)^2}\varpi_{E,i}. \tag{7}$$

Equation (6) can be used to obtain the Einstein frequency $\omega_{E,i}$, which depends in principle on the scattering path i. But one may attempt a best fit with one Einstein temperature $\Theta_E=\hbar\omega_{E,i}/k_B$ for all single-scattering paths, allowing for a systematic error $\Delta\Theta_E$ of the Einstein model. Equation (7) yields the anharmonicity parameter ϖ_E, which we will similarly assume to be independent of the scattering path i. In this discussion we neglected a possible structural contribution to $C_{3,i}$.

In addition, the equation

$$\frac{da_i(T)}{dT}=-12\xi_i\left(\frac{\ln\xi_i}{1-\xi_i}\right)^2\frac{\varpi_{E,i}k_B}{M^2\omega_{E,i}^4} \tag{8}$$

is derived in Ref. 19, which relates the thermal expansion coefficient $a_i(T)$ to the anharmonicity parameter $\varpi_{E,i}$.

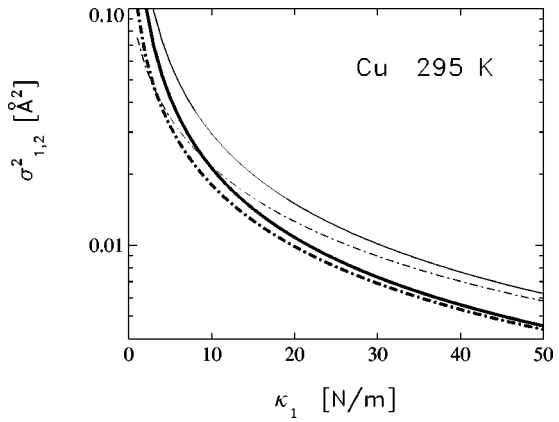

FIG. 2. Dependence of the Debye-Waller parameters on the spring constant κ_1 for the first (thick lines) and second (thin lines) single-scattering path in copper at 295 K. The full lines refer to a model with a single spring constant, the dash-dotted lines to a model with two two spring constants κ_1 and κ_2, where κ_2 was set to 3 N/m.

IV. DISCUSSION OF UNCERTAINTIES AFFECTING THE FIT

Besides uncertainties of σ^{therm} due to the approximate handling of the lattice dynamics, also the other components of the vector $\mathbf{y}$, $f_j(k_l)$, $\phi_j(k_l)$, and $\lambda(k_l)$, which are calculated in the FEFF7 code, are connected with uncertainties due to the approximate treatment of the electronic many-body problem. Of the probability distribution of $\mathbf{y}$, we do not know anyting except the most probable value $\mathbf{y}$ and an estimate for the variance of each of its components. The maximum-entropy principle then yields a Gaussian distribution $P_{\text{model}}\propto\exp[-\chi^2_{\text{model}}(\mathbf{y})/2]$, with

$$\chi^2_{\text{model}}=(\mathbf{y}-\mathbf{y}_0)^T\mathbf{D}(\mathbf{y}-\mathbf{y}_0),$$

where **D** is a diagonal matrix containing the variances of the components of **y**, $D_{jj'}=\Delta y_j^2\delta_{jj'}$. In order not to deviate too much from established practices, we use the lightface symbol χ^2 for the quadratic forms defining Gaussian distributions and employ the boldface symbol $\boldsymbol{\chi}(k)$ for the EXAFS function.

In the applications to be described below, we assume an error $\Delta f_j/f_j=7\%$, $\Delta\phi_j=0.07$ rad, and $\Delta\lambda/\lambda=10\%$, for all k_j.[20] These are rough estimates. In our applications it turns out that the uncertainties caused by the Δy_j are much smaller than those from other sources, in particular from the input data $\boldsymbol{\chi}(k_l)$. Therefore these estimates are sufficient for the present purpose.

Another source of uncertainty in the fit is the truncation of the multiple-scattering (MS) series (1). To obtain an estimate for the truncation error, let us call $\boldsymbol{\chi}_I(k)$ the value of the sum (1) when it is truncated with the Ith term, and introduce the differences $X_i(k_l)=\boldsymbol{\chi}_{I+i}(k_l)-\boldsymbol{\chi}_I(k_l)$. In Ref. 2 we defined averages

$$\overline{X(k_l)}=\frac{1}{I_{\max}}\sum_{i=1}^{I_{\max}}X_i(k_l)$$

and a correlation matrix

$$\Sigma_{ll'}=\overline{X(k_l)X(k_{l'})-X(k_l)\cdot X(k_{l'})},$$

where we use $I_{\max}=1000$.

The original observables $\mu(k_l)$ as well as the background subtraction contribute to the uncertainty of the input $\boldsymbol{\chi}(k_l)$, of the fit. We therefore have to associate an error $\Delta\boldsymbol{\chi}_l$ with each input data point. It is convenient in latter calculations to scale $\boldsymbol{\chi}(k_l)$ by $\Delta\boldsymbol{\chi}_l$ and use the vector with components $g_l=\boldsymbol{\chi}(k_l)/\Delta\boldsymbol{\chi}_l$ as input vector. Assuming again a Gaussian distribution of the input data, the probability that $\overline{\mathbf{g}}$ is the true value when $\mathbf{g}'$ is the measured average value is given by

$$P_{\exp}(\overline{\mathbf{g}}|\mathbf{g}')=(2\pi)^{-L/2}e^{-(1/2)\chi^2_{\exp}(\overline{\mathbf{g}},\mathbf{g}')},$$

with $\chi^2_{\exp}=(\overline{\mathbf{g}}-\mathbf{g}')^T(\overline{\mathbf{g}}-\mathbf{g}')$ in matrix notation.

As in Ref. 2, the uncertainty in **g** caused by the truncation error will be modeled by the Gaussian $P_{\text{trunc}}(\mathbf{g},\mathbf{g}')\propto\exp[-\chi^2_{\text{trunc}}(\mathbf{g},\mathbf{g}')/2]$, with

$$\chi^2_{\text{trunc}}(\mathbf{g},\mathbf{g}')=[\mathbf{g}(\mathbf{x},\mathbf{y})-\mathbf{g}']^T\mathbf{B}[\mathbf{g}(\mathbf{x},\mathbf{y})-\mathbf{g}'],$$

where the variance matrix **B** is given by

$$(B^{-1})_{ll'}=\frac{\Sigma_{ll'}}{\Delta\boldsymbol{\chi}_l\Delta\boldsymbol{\chi}_{l'}}.$$

The total uncertainty due to truncation and model approximations is $P_{\text{syst}}(\mathbf{g}',\mathbf{y};\mathbf{g})=P_{\text{trunc}}P_{\text{model}}$. The conditional probability for the observable **g**, once the model parameters have the value **x**, is obtained from $P_{\exp}$ by folding with P_{syst},

$$P_{\text{cond}}(\overline{\mathbf{g}}|\mathbf{x})=\int P_{\text{syst}}(\mathbf{g}',\mathbf{y};\mathbf{g}(\mathbf{x},\mathbf{y}))P_{\exp}(\overline{\mathbf{g}}|\mathbf{g}')d^M\mathbf{y}d^L\mathbf{g}'. \tag{9}$$

The integration with respect to $\mathbf{g}'$ can be performed analytically using the formula

$$\int e^{-(1/2)\mathbf{x}^T\mathbf{A}\mathbf{x}+\mathbf{b}^T\mathbf{x}}d^L\mathbf{x}=\sqrt{\frac{(2\pi)^L}{\det(\mathbf{A})}}e^{(1/2)\mathbf{b}^T\mathbf{A}^{-1}\mathbf{b}}, \tag{10}$$

where **b** is any vector and **A** is a symmetric matrix.[21] Equation (10) is most easily proved in the eigenrepresentation of the matrix **A**. One obtains from Eq. (9)

$$P_{\text{cond}}(\overline{\mathbf{g}}|\mathbf{x})\propto\int e^{-(1/2)\chi^2_{\text{interm}}(\overline{\mathbf{g}},\mathbf{x},\mathbf{y})}d^M\mathbf{y}, \tag{11}$$

where

$$\chi^2_{\text{interm}}=-(\overline{\mathbf{g}}^T+\mathbf{g}^T\mathbf{B})(\mathbf{1}+\mathbf{B})^{-1}(\overline{\mathbf{g}}+\mathbf{B}\mathbf{g})+\mathbf{g}^T\mathbf{B}\mathbf{g}+\overline{\mathbf{g}}^T\overline{\mathbf{g}}+(\mathbf{y}-\mathbf{y}^{(0)})^T\mathbf{D}(\mathbf{y}-\mathbf{y}^{(0)}). \tag{12}$$

Finally, we need the probability $P_{\text{post}}(\mathbf{x}|\overline{\mathbf{g}})$ for the distribution of the model parameters **x**, once $\overline{\mathbf{g}}$ is given. This probability is obtained from Bayes' theorem,[3,22]

$$P_{\text{post}}(\mathbf{x}|\overline{\mathbf{g}})=\frac{P_{\text{prior}}(\mathbf{x})P_{\text{cond}}(\overline{\mathbf{g}}|\mathbf{x})}{\int P_{\text{prior}}(\mathbf{x})P_{\text{cond}}(\overline{\mathbf{g}}|\mathbf{x})d\mathbf{x}}, \tag{13}$$

in terms of the *a priori* distribution $P_{\text{prior}}(\mathbf{x})$. With our definition of **x**, we know that its *a priori* average value vanishes. Additional information on P_{prior} can be obtained on the basis of the available experimental data, as will be shown in the following section. For practical reasons, however, only a rather limited number of parameters may be calculated in this way. We will therefore assume that only the *a priori* errors α_n^{-1} of the components of **x** are given. The maximum entropy principle then yields[23]

$$P_{\text{prior}}(\mathbf{x})=\prod_{n=1}^{N}\left(\frac{2\pi}{\alpha_n}\right)^{1/2}e^{-(1/2)\alpha_n x_n^2}. \tag{14}$$

V. EXTENDED χ^2 FIT

A. Derivation of the regularized normal equations

We will first assume that the errors in **x** and **y** are sufficiently small and that the *a priori* guess of the model parameters is close to the solution of the fit, so that the function $\mathbf{g}(\mathbf{x},\mathbf{y})$ may be expanded to linear order,

$$g_l(\mathbf{x},\mathbf{y})=g_l(\mathbf{0},\mathbf{y}^{(0)}(k_l))+\sum_{n=1}^{N}G_{ln}x_n+\sum_{m=1}^{M}T_{lm}(y_m-y_m^{(0)}), \tag{15}$$

in terms of the rectangular matrices

$$G_{ln}=\left.\frac{\partial g_l}{\partial x_n}\right|_{\mathbf{x}=\mathbf{0}}+\sum_{m=1}^{M}\left.\frac{\partial g_l}{\partial y_m}\right|_{\mathbf{x}=\mathbf{0},\mathbf{y}=\mathbf{y}^{(0)}}\left.\frac{\partial y_m}{\partial x_n}\right|_{\mathbf{x}=\mathbf{0}} \tag{16}$$

and

$$T_{lm}=\left.\frac{\partial g_l}{\partial y_m}\right|_{\mathbf{x}=\mathbf{0},\mathbf{y}=\mathbf{y}^{(0)}},$$

with $\mathbf{y}^{(0)}=\mathbf{y}(\mathbf{k},\mathbf{0})$. The integral in Eq. (11) can now be evaluated analytically using Eq. (10). In terms of the $L\times L$ matrix

$$\mathbf{C}=\{\mathbf{1}-(\mathbf{1}+\mathbf{B}^{-1})^{-1}\mathbf{T}[\mathbf{D}+\mathbf{T}^{\mathbf{T}}(\mathbf{1}+\mathbf{B}^{-1})^{-1}\mathbf{T}]^{-1}\mathbf{T}^{T}\}\times(\mathbf{1}+\mathbf{B}^{-1})^{-1}, \tag{17}$$

the $N\times N$ information matrix

$$\mathbf{Q}=\mathbf{G}^{T}\mathbf{C}\mathbf{G}, \tag{18}$$

and the vector

$$\mathbf{b}^{T}=(\overline{\mathbf{g}}-\mathbf{g}_0)^{T}\mathbf{C}\mathbf{G}, \tag{19}$$

one obtains $P_{\mathrm{cond}}\propto\exp(-\frac{1}{2}\chi^2_{\mathrm{cond}})$, with

$$\chi^2_{\mathrm{cond}}=\mathbf{x}^{T}\mathbf{Q}\mathbf{x}-2\mathbf{b}^{T}\mathbf{x}+(\overline{\mathbf{g}}-\mathbf{g}_0)^{T}\mathbf{C}(\overline{\mathbf{g}}-\mathbf{g}_0)=[\mathbf{g}(\mathbf{x},\mathbf{y}^{(0)})-\overline{\mathbf{g}}]^{T}\mathbf{C}[\mathbf{g}(\mathbf{x},\mathbf{y}^{(0)})-\overline{\mathbf{g}}], \tag{20}$$

where $\mathbf{g}_0=\mathbf{g}(\mathbf{0},\mathbf{y}^{(0)})$, and $P_{\mathrm{post}}\propto\exp(-\frac{1}{2}\chi^2_{\mathrm{post}})$ with

$$\chi^2_{\mathrm{post}}=\mathbf{x}^{T}(\mathbf{Q}+\mathbf{A})\mathbf{x}-2\mathbf{b}^{T}\mathbf{x}+(\overline{\mathbf{g}}-\mathbf{g}_0)^{T}\mathbf{C}(\overline{\mathbf{g}}-\mathbf{g}_0), \tag{21}$$

where $A_{nn'}=\alpha_n\delta_{nn'}$.[2]

If the model and truncation errors $\mathbf{D}^{-1}$ and $\mathbf{B}^{-1}$ are small compared to the input errors $\Delta\boldsymbol{\chi}$, the brackets in Eq. (17) can be expanded to linear order in these matrices. One obtains the simpler expression

$$\mathbf{C}=(\mathbf{1}+\mathbf{B}^{-1}+\mathbf{T}\mathbf{D}^{-1}\mathbf{T}^{T})^{-1}. \tag{22}$$

If one neglects the off-diagonal matrix elements $C_{ll'}$ in Eq. (22), Eq. (20) can be rewritten in the particularly transparent form,

$$\chi^2_{\mathrm{cond}}=\sum_{l=1}^{L}\left[\frac{\overline{g_l}-g(k_l;\mathbf{x},\mathbf{y}^{(0)})}{\Delta g_l^{\mathrm{eff}}}\right]^2,$$

with

$$(\Delta g_l^{\mathrm{eff}})^2=1+\sum_{m=1}^{M}\left[\frac{\partial g(k_l;\mathbf{0},\mathbf{y})}{\partial y_m}\Delta y_m\right]^2+(B^{-1})_{ll},$$

i.e., in this approximation the experimental errors (which are equal to unity with our normalization of $\mathbf{g}$), the model errors, and the truncation errors $B^{-1/2}$ add quadratically to the effective error $\Delta g_l^{\mathrm{eff}}$.

The *a posteriori* expectation value of the model parameters,

$$\overline{\mathbf{x}}:=\langle\mathbf{x}\rangle_{\mathrm{post}}=\int\mathbf{x}P_{\mathrm{post}}(\mathbf{x})d^N\mathbf{x}$$

follows from solving the normal equations

$$\sum_{n'=1}^{N}(Q_{nn'}+\alpha_n\delta_{nn'})\overline{x_{n'}}=b_n. \tag{23}$$

The *a posteriori* variance matrix is

$$\langle(x_n-\overline{x_n})(x_{n'}-\overline{x_{n'}})\rangle_{\mathrm{post}}=(Q+A)^{-1}_{nn'}. \tag{24}$$

If we call q_1 as the largest and q_N as the smallest eigenvalue of $\mathbf{Q}$, the ratio q_1/q_N is an extremly large number, reflecting the ill-conditioned nature of our inversion problem. Choosing all eigenvalues α_n such that $q_1\gg\alpha_n\gg q_N$, the matrix $\mathbf{A}$ in Eq. (23) is seen to regularize the inversion problem $\mathbf{Q}\overline{\mathbf{x}}=\mathbf{b}$ in the sense of Tikhonov and Arsenin[24] since the smallest eigenvalue q_{reg} of the matrix $\mathbf{Q}+\mathbf{A}$ is given by

$$q_{\mathrm{reg}}:=\min_{\mathbf{x}}[\mathbf{x}^{T}(\mathbf{Q}+\mathbf{A})\mathbf{x}]=\min_{\mathbf{x}}[\mathbf{x}^{T}\mathbf{Q}\mathbf{x}]+\min_{\mathbf{x}}[\mathbf{x}^{T}\mathbf{A}\mathbf{x}]>\alpha_{\min}.$$

Therefore the smallest eigenvalue $\alpha_{\min}$ of $\mathbf{A}$ may be chosen such that the condition number q_1/q_{reg} of the matrix $\mathbf{Q}+\mathbf{A}$ does not become dangerously large. The introduction of the matrix $\mathbf{A}$ through the *a priori* probability P_{prior} has therefore been called stochastic regularization.[4]

B. Determination of the regularization parameters

In order to obtain estimates for the eigenvalues α_n, we slightly generalize a procedure introduced by Turchin and Nozik.[21] We start with the assumption that there is a probability distribution for the vector $\boldsymbol{\alpha}$. The conditional probability for the inhomogeneous term $\mathbf{b}$ in Eq. (23), once $\boldsymbol{\alpha}$ is given, is taken to be

$$P_{\mathrm{cond}}(\mathbf{b}|\boldsymbol{\alpha})=\int P_{\mathrm{cond}}(\overline{\mathbf{g}}|\mathbf{x})P_{\mathrm{prior}}(\mathbf{x})d^N\mathbf{x}=\mathrm{const}\sqrt{\frac{\det\mathbf{A}}{\det(\mathbf{Q}+\mathbf{A})}}\,e^{(1/2)\mathbf{b}^{T}(\mathbf{Q}+\mathbf{A})^{-1}\mathbf{b}}, \tag{25}$$

where the normalization constant is independent of $\boldsymbol{\alpha}$. We assume that the *a priori* probability $P_{\mathrm{prior}}(\boldsymbol{\alpha})$ for $\boldsymbol{\alpha}$ is constant in a sufficiently large area $\mathcal{A}$ in $\boldsymbol{\alpha}$ space,[25] defined by, say, $\alpha_{\max}:=q_1>\alpha_n>10^{-8}q_1:=\alpha_{\min}$, where q_1 is the largest eigenvalue of the information matrix Q. Bayes' theorem then yields $P_{\mathrm{post}}(\boldsymbol{\alpha}|\mathbf{b})\propto P_{\mathrm{cond}}(\mathbf{b}|\boldsymbol{\alpha})$ on $\mathcal{A}$, and 0 otherwise. If the function $P_{\mathrm{post}}(\boldsymbol{\alpha}|\mathbf{b})$ is sharply peaked as a function of $\boldsymbol{\alpha}$ for fixed $\mathbf{b}$, one may use the α_n of this maximum in Eq. (23) instead of the distribution in $\boldsymbol{\alpha}$. Differentiating Eq. (25) with respect to α_n,

$$\partial_{\alpha_n}\ln P_{\mathrm{cond}}(\mathbf{b}|\boldsymbol{\alpha})=\frac{1}{2}\partial_{\alpha_n}[\ln(\det\mathbf{A})-\ln(\det(\mathbf{Q}+\mathbf{A}))]+\frac{1}{2}\mathbf{b}^{T}\partial_{\alpha_n}(\mathbf{Q}+\mathbf{A})^{-1}\mathbf{b},$$

and using the relation

$$\partial_{\alpha_m}(Q+A)^{-1}_{nn'}=-(Q+A)^{-1}_{nm}(Q+A)^{-1}_{n'm}$$

yields the N nonlinear equations

$$A_{nn}^{-1}-(Q+A)_{nn}^{-1}-\left(\sum_{n'=1}^{N}(Q+A)_{nn'}^{-1}b_{n'}\right)^2=0, \quad (26)$$

where $n=1,\ldots,N$ for the extremum of $P_{\mathrm{cond}}(\mathbf{b}|\boldsymbol{\alpha})$. Those components α_n of the solution $\boldsymbol{\alpha}$ of Eq. (26) which are larger than $\alpha_{\max}$ or which are negative, i.e., lie outside the area $\mathcal{A}$, are replaced by $\alpha_{\max}$, those between 0 and $\alpha_{\min}$ are replaced by $\alpha_{\min}$.

The geometric mean of the components of $\boldsymbol{\alpha}$ may be called α^*. We then introduce scaling factors $\tilde{\alpha}_n=\sqrt{\alpha_n/\alpha^*}$ and rescale the model parameters

$$\tilde{x}_n=\tilde{\alpha}_n x_n. \quad (27)$$

Equation (16) shows that, using $\tilde{\mathbf{x}}$ instead of $\mathbf{x}$, the vector $\mathbf{b}$ and the matrices $\mathbf{G}$, $\mathbf{Q}$, and $\mathbf{A}$ scale like $\tilde{b}_n=b_n/\tilde{\alpha}_n$, $\tilde{G}_{ln}=G_{ln}/\tilde{\alpha}_n$, $\tilde{Q}_{nn'}=Q_{nn'}/(\tilde{\alpha}_n\tilde{\alpha}_{n'})$, and $\tilde{A}_{nn'}=\alpha^*\delta_{nn'}$. It is easily seen that the rescaled quantities are independent of the somewhat arbitrary choice of the original scaling parameters $\hat{x}_n$. Since $\tilde{\mathbf{A}}$ commutes with $\tilde{\mathbf{Q}}$, the equations

$$\sum_{n'=1}^{N}(\tilde{Q}_{nn'}+\alpha^*\delta_{nn'})\tilde{x}_{n'}=\tilde{b}_n$$

decouple in the eigenrepresentation of the matrix $\tilde{\mathbf{Q}}$. If the components of the vectors $\tilde{\mathbf{x}}$ and $\tilde{\mathbf{b}}$ in this representation are called ξ_n and β_n, respectively, and the eigenvalues of $\tilde{\mathbf{Q}}$ are called $\tilde{q}_n$, the normal equations (23) become

$$(\tilde{q}_n+\alpha^*)\xi_n=\beta_n, \quad n=1,\ldots,N. \quad (28)$$

In terms of these quantities, Eq. (26) becomes

$$\frac{1}{\alpha^*}-\frac{1}{\tilde{q}_n+\alpha^*}-\frac{\beta_n^2}{(\tilde{q}_n+\alpha^*)^2}=0,$$

valid for any n because of our special scaling, except for those n where the solution of Eq. (26) has been replaced by the boundary values $\alpha_{\min}$ or $\alpha_{\max}$. This equation can be solved for α^*. Using Eqs. (28) and (24) formally for $\alpha=0$, one obtains

$$\alpha^*=\frac{1}{(\xi_n(\boldsymbol{\alpha}=0))^2-(\Delta\xi_n(\boldsymbol{\alpha}=0))^2},$$

where $(\Delta\xi_n(\boldsymbol{\alpha}=0))^2=1/\tilde{q}_n$.

This equation shows that α^* becomes very large when $\xi_n(\boldsymbol{\alpha}=0)$ approaches $\Delta\xi_n(\boldsymbol{\alpha}=0)$, and would even become negative for $\xi_n(\boldsymbol{\alpha}=0)<\Delta\xi_n(\boldsymbol{\alpha}=0)$, had we not restricted $\boldsymbol{\alpha}$ to the area $\mathcal{A}$. In other words, if the *a priori* model-parameter vector $\mathbf{x}^{(0)}$ approaches the area of one standard deviation around the (hypothetical) purely experimental mean value $\overline{\mathbf{x}'}$, the regularization parameter becomes equal to q_1, the largest eigenvalue of the matrix Q. On the other hand, if the *a priori* parameters $\mathbf{x}^{(0)}$ are far away from $\overline{\mathbf{x}'}$, α^* becomes very small.

C. Interpretation of *a posteriori* model parameters and errors

As in Ref. 2, we call $\mathcal{R}$ the space spanned by the eigenvectors of those eigenvalues for which $\tilde{q}_n>\alpha^*$. Its orthogonal complement in the total model space $\mathcal{Q}$ shall be $\mathcal{P}$. The structure of Eq. (28) shows that in $\mathcal{P}$ space the solution of the normal equations is predominantly determined by the *a priori* assumptions, whereas in $\mathcal{R}$ space the result is—as in ordinary χ^2 fitting—determined by the data. For the interpretation of the result of the fit, it is therefore convenient to define an approximate projection s_n of the model parameter x_n into the space $\mathcal{R}$ by

$$s_n^2:=\sum_{n'=1}^{N}Q_{nn'}(Q+A)_{nn'}^{-1}=\sum_{n'=1}^{N}\tilde{Q}_{nn'}(\tilde{Q}+\alpha^*I)_{n'n}^{-1}$$
$$=\sum_{n'=1}^{N}U_{nn'}^2\frac{\tilde{q}_{n'}}{\tilde{q}_{n'}+\alpha^*},$$

where $U_{nn'}$ is the unitary matrix that transforms $\tilde{\mathbf{Q}}$ to its diagonal form. If the ordered sequence of eigenvalues $\tilde{q}_n$ decreases very rapidly, one has

$$s_n^2\approx\sum_{n'=1}^{i_{\mathcal{R}}}U_{nn'}^2,$$

where $i_{\mathcal{R}}$ is the dimension of $\mathcal{R}$ space, for which one finds

$$i_{\mathcal{R}}\approx\mathrm{tr}\,\mathbf{Q}(\mathbf{Q}+\mathbf{A})^{-1}.$$

The quantities α_n, $i_{\mathcal{R}}$, and s_n^2 depend not only on the data and their errors and on the type of model on which the analysis is based (in our case, FEFF7 for the electronic and the force-field model of Sec. III for the vibrational properties), but also very essentially on the choice of the *a priori* vector $\mathbf{x}^{(0)}$. According to the logic of the Bayesian analysis the information that goes into the determination of all *a posteriori* quantities is first taken from the *a priori* assumptions. The experimental information is only brought in to the extent that the data require a modification of these assumptions, i.e., something ''new'' has been learned from the experiment. If, for instance, a set of data is analyzed with an *a priori* that is taken to be the result of a previous analysis of the same data, then nothing new can be learned and therefore α^* would now be equal to q_1 and $i_{\mathcal{R}}=0$.

The dimension $i_{\mathcal{R}}$ is expected to be smaller than the ''number of independent data points'' $N_d=(2/\pi)(k_{\max}-k_{\min})\Delta R+2$,[26] where $\Delta R=R_I-R_1$, because N_d represents the maximal amount of information that can be accommodated on the interval $k_{\max}-k_{\min}$ if it were optimally arranged. The dimension $i_{\mathcal{R}}$, however, represents the amount of information obtained by the experiment beyond the *a priori* information, taking all error sources into account. It is not possible to define the information content of a given set of experimental data completely independent of the choice of *a priori* model parameters. Quantities which characterize rather closely what one naively means by the quality of a set of experimental data (without referring to *a priori* assumptions) are obtained by choosing $\mathbf{x}^{(0)}$ far away from the final

fit and considering the projections $\hat{s}_n^2$ into the $\mathcal{R}$ space, resulting from this choice of the *a priori* data. In the numerical examples to be discussed later, it turned out that by shifting the *a priori* value of each single model parameter sufficiently far away from its final fit value, its projection $\hat{s}_n^2$ can always be brought close to unity.

For many model parameters, the s_n^2 turn out to be significantly different from 1, but not 0. The space $\mathcal{R}$, where the data determine the fit, is therefore *not* spanned by just $i_{\mathcal{R}}$ of the model parameters x_n. It is rather spanned by the first $i_{\mathcal{R}}$ of the ξ_n. This has important consequences for the proper interpretation of the *a posteriori* errors and error correlations (24). Only for those model parameters for which $s_n^2 \approx 1$ the *a posteriori* error has the usual meaning, accounting for experimental errors, uncertainties in the model, and various truncation errors. For the other x_n, only a fraction is determined in this way. To associate an error with the total parameter x_n, not just its projection into $\mathcal{R}$ space, it is unavoidable to rely on *a priori* information and its estimated uncertainty. Therefore, also the parameters that are poorly determined by the data have a finite *a posteriori* error.

It is also seen that it may be misleading to restrict the model-parameter space from the very beginning to a sufficiently small dimension, so that the normal equations (23) do not require any regularization. Some of the parameters kept in this approach may in reality be only partly determined by the data. Nevertheless, they appear to follow from the data with reasonable ''experimental'' errors since they tend to mock up some of the model parameters that are left out and become distorted in this way.

Sometimes the quality of a fit is discussed in terms of an expression $\tilde{R}$ similar to χ^2_{cond}, Eq. (20), but with the matrix $\mathbf{C}$ proportional to the unit matrix.[27] It must be stressed that our solution of the normal equations (23) minimizes χ^2_{post} and *not* $\tilde{R}$. Apart from an attraction of our solution towards the *a priori* solution because of the matrix $\mathbf{A}$ in Eq. (21), our matrix $\mathbf{C}$ weighs input in k space more strongly in the direction of those of its eigenvectors that belong to the larger eigenvalues, whereas in $\tilde{R}$ all data points have been given the same weight.

Since the two formal model parameters S_0^2 and E_0 are in general not of interest in themselves, we integrate the probability $P_{\mathrm{cond}}(\mathbf{b}|\mathbf{x})$ over these two variables. Using Eq. (10), one obtains the Gaussian distribution $P^{\mathrm{red}}_{\mathrm{cond}}(\mathbf{b}^{\mathrm{red}}|\mathbf{x}^{\mathrm{red}}) \propto \exp(\chi^2_{\mathrm{red}}/2)$, with

$$\chi^2_{\mathrm{red}} = \sum_{n,n'=3}^{N} x_n Q^{\mathrm{red}}_{nn'} x_{n'} - 2\sum_{n=3}^{N} b_n^{\mathrm{red}} x_n + \sum_{l,l'=1}^{L} [\overline{g_l} - g_l(0)] C_{ll'} [\overline{g_{l'}} - g_{l'}(0)],$$

where for $n,n' = 3, \ldots, N$,

$$Q^{\mathrm{red}}_{nn'} = Q_{nn'} - \sum_{\nu,\nu'=1}^{2} Q_{n\nu} P_{\nu\nu'} Q_{\nu' n'}$$

and

$$b_n^{\mathrm{red}} = b_n - \sum_{\nu,\nu'=1}^{2} Q_{n\nu} P_{\nu\nu'} b_{\nu'},$$

and $P_{\nu\nu'}$ is the inverse of the 2×2 matrix $Q_{\nu\nu'}$ with $\nu,\nu' = 1,2$. The mean values $\overline{x_n}$ follow from the reduced normal equations

$$\sum_{n'=3}^{N} (Q^{\mathrm{red}}_{nn'} + \alpha_n \delta_{nn'}) \overline{x_{n'}} = b_n^{\mathrm{red}}, \quad n = 3, \ldots, N, \tag{29}$$

and the reduced variance matrix is $(Q^{\mathrm{red}}_{nn'} + \alpha_n \delta_{nn'})^{-1}$, with $n,n' = 3 \ldots, N$.

D. Alternative strategy for the determination of the overall regularization parameter α^*

In the space of the scaled model parameters $\tilde{\mathbf{x}}$, the strength parameter α^* is the most probable regularization in the Bayesian sense. For certain purposes it is useful to determine instead the largest regularization parameter $\hat{\alpha}^*$ compatible with the data. The condition for this parameter,

$$\langle \chi^2_{\mathrm{cond}}(\tilde{\mathbf{x}}) \rangle_{\mathrm{post}} = L, \tag{30}$$

was first proposed by Turchin.[25] In the ill-posed case we are considering here, this equation generalizes a known result for well-posed problems:[28] If all errors are properly estimated, the minimum of χ^2 with respect to the model parameters is equal to the number of degrees of freedom. Condition (30) is equivalent to the nonlinear equation

$$\chi^2_{\mathrm{cond}}(\overline{\mathbf{x}}(\hat{\alpha}^*)) = L - \mathrm{tr}\tilde{\mathbf{Q}}(\tilde{\mathbf{Q}} + \hat{\alpha}^* \mathbf{I})^{-1}.$$

Since the largest possible weight of the regularization is larger than the most probable weight, one finds $\hat{\alpha}^* \geqslant \alpha^*$. Solutions $\overline{\mathbf{x}}$ corresponding to $\hat{\alpha}^*$ are useful to decide between two competing models on the basis of a given measurement.[28]

VI. ITERATIVE SOLUTIONS

The expansion of $\mathbf{g}(\mathbf{x},\mathbf{y})$ to linear order used in Eq. (15) is only justified if the *a priori* guess $\mathbf{x}^{(0)}$ is sufficiently close to the solution $\overline{\mathbf{x}}$ of the normal equations (23). In general, this will not be true. In such a case we expand $\mathbf{g}(\mathbf{x},\mathbf{y}^{(0)})$ around $\mathbf{x}^{(\nu)}$ supposed to be sufficiently close to $\overline{\mathbf{x}}$ to allow a linear expansion. From Eqs. (14) and (20), one obtains

$$\chi^2_{\mathrm{post}}(\mathbf{x},\boldsymbol{\alpha}) = \mathbf{x}^T(\mathbf{Q}^{(\nu)} + \mathbf{A})\mathbf{x} - 2\mathbf{b}^{(\nu)T}\mathbf{x} - 2\mathbf{x}^{(\nu)T}\mathbf{Q}^{(\nu)}\mathbf{x} + \mathbf{x}^{(\nu)T}\mathbf{Q}^{(\nu)}\mathbf{x}^{(\nu)} + 2\mathbf{b}^{(\nu)T}\mathbf{x}^{(\nu)} + \chi^2_{\mathrm{cond}}(\mathbf{x}^{(\nu)}).$$

The information matrix $\mathbf{Q}^{(\nu)}$ and the inhomogeneous term $\mathbf{b}^{(\nu)}$ are given by Eqs. (18) and (19), where

$$G^{(\nu)}_{ln} = \left.\frac{\partial g_l}{\partial x_n}\right|_{\mathbf{x}=\mathbf{x}^{(\nu)}} + \sum_{m=1}^{M} \left.\frac{\partial g_l}{\partial y_m}\right|_{\mathbf{x}=\mathbf{x}(\nu),\mathbf{y}=\mathbf{y}(k_l,\mathbf{x}^{(\nu)})} \left.\frac{\partial y_m}{\partial x_n}\right|_{\mathbf{x}=\mathbf{x}^{(\nu)}} \tag{31}$$

and

$$T_{lm}^{(\nu)}=\left.\frac{\partial g_l}{\partial y_m}\right|_{\mathbf{x}=\mathbf{x}^{(\nu)},\mathbf{y}=\mathbf{y}(k_l,\mathbf{x}^{(\nu)})}$$

are to be used. Minimizing $\chi^2_{\rm post}$ with respect to $\mathbf{x}$ yields the equations

$$\sum_{n'=1}^{N}[Q_{nn'}^{(\nu)}(x-x^{(\nu)})_{n'}+A_{nn'}^{(\nu)}x_{n'}]=b_n^{(\nu)}, \tag{32}$$

$n=1,\ldots,N$, which are used to obtain an improved solution $\mathbf{x}^{(\nu+1)}=\mathbf{x}$, thus constructing a sequence of approximants $\mathbf{x}^{(\nu)}$, starting with $\mathbf{x}^{(0)}=0$.

In the first few iteration steps, the matrix $\mathbf{A}^{(\nu)}$ is just used to regularize the matrix inversion needed to solve Eq. (32) in each iteration step ν. Here one can choose a regularization matrix proportional to the unit matrix, $A_{nn'}^{(\nu)}=\alpha^{(\nu)}\delta_{nn'}$. This scheme is the iteratively regularized Gauss-Newton procedure,[29] a stabilized version of the Levenberg-Marquardt algorithm.[30] A strategy is needed to choose the sequence $\alpha^{(\nu)}$ in the iteration. The starting value $\alpha^{(0)}$ is chosen such that

$$\mathrm{tr}[\mathbf{Q^{(0)}}/(\mathbf{Q^{(0)}}+\mathbf{A^{(0)}})]\approx 1.$$

We then decrease α in each step by a factor of 10.

As in Ref. 2, we simplify the calculation of $\partial y_m/\partial x_n|_{\mathbf{x}=\mathbf{x}^{(\nu)}}$ in Eq. (31): For the components of $\mathbf{x}$, which correspond to S_0, E_0, and to the half lengths R_i, the derivative is taken to be independent of ν, for the Debye-Waller parameters σ_j^2 and the third cumulants $C_{3,j}$ it is taken equal to zero, and only for the spring constants κ_s it is recalculated in each iteration step.

Comparison of Eq. (23) with Eq. (32) shows that the analog of Eq. (26) in the nonlinear case is obtained by replacing $\mathbf{b}$ in Eq. (26) by $\mathbf{b}^{(\nu)}+\mathbf{Q}^{(\nu)}\mathbf{x}^{(\nu)}$. With the abbreviation

$$\Delta_n(\boldsymbol{\alpha}^{(\nu)})=\sum_{n'=1}^{N}(Q^{(\nu)}+A^{(\nu)})_{nn'}^{-1}\left(b_{n'}^{(\nu)}+\sum_{n''=1}^{N}Q_{n'n''}^{(\nu)}x_{n''}^{(\nu)}\right), \tag{33}$$

one obtains

$$(A^{(\nu)})_{nn}^{-1}-(Q^{(\nu)}+A^{(\nu)})_{nn}^{-1}-\Delta_n^2(\boldsymbol{\alpha}^{(\nu)})=0. \tag{34}$$

It turns out that the trace of this equation is negative in the first few iteration steps. After this quantity has changed its sign, we start to solve Eqs. (32) and (34) simultaneously. To reduce the necessary computational effort somewhat, we assumed that there are only up to seven different coefficients α_n. Two, α_1 and α_2, for the parameters S_0 and E_0, respectively, a third one, for the half radii R_i, a fourth and fifth one for the Debye-Waller parameters σ_j of the single- and the multiple-scattering paths, respectively, or for the two spring constants κ_s, and α_6 and α_7 for the first and for the rest of the third cumulants $C_{3,i}$. If the model parameters from $n=n_{s-1}+1$ to $n=n_s$ correspond to α_s, we obtain the gradient $h_s(\boldsymbol{\alpha}^{(\rm red)})$ of $\ln P_{\rm post}(\mathbf{b}|\boldsymbol{\alpha}^{(\rm red)})$ as function of $\boldsymbol{\alpha}^{(\rm red)}=(\alpha_1,\ldots,\alpha_7)$ by summing over Eqs. (34),

$$h_s(\boldsymbol{\alpha}^{(\rm red)}):=\frac{1}{2}\sum_{n=n_{s-1}+1}^{n_s}\{(A^{(\nu)})_{nn}^{-1}-(Q^{(\nu)}+A^{(\nu)})_{nn}^{-1}-\Delta_n^2(\boldsymbol{\alpha}^{(\rm red)})\}=0,\quad s=1,\ldots,7; \tag{35}$$

and the Hessian is

$$H_{ss'}=\frac{1}{2}\sum_{n=n_{s-1}+1}^{n_s}\sum_{n'=n_{s'-1}+1}^{n_{s'}}\{-(\alpha_n^{(\nu)})^{-2}\delta_{nn'}+(Q^{(\nu)}+A^{(\nu)})_{nn'}^{-1}[(Q^{(\nu)}+A^{(\nu)})_{nn'}^{-1}+2\Delta_n(\boldsymbol{\alpha}^{(\rm red)})\Delta_{n'}(\boldsymbol{\alpha}^{(\rm red)})]\}, \tag{36}$$

where $s,s'=1,\ldots,7$ and the upper index (ν) on $\boldsymbol{\alpha}^{(\rm red)}$ has been dropped.

To solve Eqs. (35) for $\boldsymbol{\alpha}^{(\rm red)}$, the quasi-Newton algorithm is used.[31] It yields the sequence of approximations $\boldsymbol{\alpha}^{(\rm red)(\mu)}$,

$$\boldsymbol{\alpha}^{(\rm red)(\mu+1)}=\boldsymbol{\alpha}^{(\rm red)(\mu)}+p_\alpha^{(\mu)}\Delta\boldsymbol{\alpha}^{(\mu)},$$

to the solution, where $p_\alpha^{(\mu)}$ is an underrelaxation factor and the Newton step $\Delta\boldsymbol{\alpha}^{(\mu)}$ follows from solving the linear equations

$$\mathbf{H}(\boldsymbol{\alpha}^{(\rm red)(\mu)})\Delta\boldsymbol{\alpha}^{(\mu)}=-\mathbf{h}(\boldsymbol{\alpha}^{(\rm red)(\mu)}).$$

The iteration is started with the last $\alpha=\alpha^*$ obtained in the preceding iteration cycle

$$\alpha_s^{(\rm red)(0)}=\alpha^*,\quad s=1,\ldots,7.$$

To determine the under-relaxation factors $p_\alpha^{(\mu)}$, we use a strategy proposed in Ref. 16. One first chooses $p_\alpha^{(\mu)}=\min(1,2p_\alpha^{(\mu-1)})$ and determines the auxiliary quantity $\overline{\Delta}\boldsymbol{\alpha}^{(\mu)}$ by solving the equations

$$\mathbf{H}(\boldsymbol{\alpha}^{(\rm red)(\mu)})\overline{\Delta}\boldsymbol{\alpha}^{(\mu)}=-\mathbf{h}(\boldsymbol{\alpha}^{(\rm red)(\mu)}+p_\alpha^{(\mu)}\Delta\boldsymbol{\alpha}^{(\mu)}).$$

If the affine invariant condition

$$\sum_{s=1}^{7}(\overline{\Delta}\alpha_s^{(\mu)})^2\leqslant\left(1-\frac{p_\alpha^{(\mu)}}{2}\right)^2\sum_{s=1}^{7}(\Delta\alpha_s^{(\mu)})^2 \tag{37}$$

is fulfilled, $p_\alpha^{(\mu)}$ is kept. Otherwise it is reduced by a factor of 2 and the test (37) is repeated. Alternatively, one may determine the minimum of

$$F(\boldsymbol{\alpha}^{(\rm red)})=\sum_{s=1}^{7}h_s^2$$

by a standard gradient procedure.[31] Once a solution $\boldsymbol{\alpha}^{(\rm red)}$ of Eqs. (35) is found, those $\alpha_s^{(\rm red)}$ that are larger than $\alpha_{\rm max}$ or negative were set equal to $\alpha_{\rm max}$, and Eqs. (32) are solved.

When Eqs. (35) are solved for the first time, the components of the solution vector $\alpha_s^{(\rm red)}$ are generally very different from α^* of the previous iteration. To stabilize the μ iteration, we again employ here an iteratively regularized Gauss-Newton procedure. For that purpose we introduce an *a priori* probability

$$P_{\text{prior}}(\boldsymbol{\alpha}) \propto \exp\left[-\epsilon^{(\alpha)} \sum_{s=1}^{7} \sum_{n=n_{s-1}+1}^{n_s} (\alpha_n - \alpha^*)^2 \right],$$

peaked at the α^* of the preceding ν iteration, instead of the rectangular distribution between $\alpha_{\min}$ and $\alpha_{\max}$ used so far to obtain $P_{\text{post}}(\boldsymbol{\alpha}|\mathbf{b})$ from $P_{\text{cond}}(\mathbf{b}|\boldsymbol{\alpha})$ in Bayes' theorem. Equation (35) is then to be replaced by

$$h_s(\boldsymbol{\alpha}^{(\text{red})}) - \epsilon^{(\alpha)} \sum_{n=n_{s-1}+1}^{n_s} (\alpha_n - \alpha^*) = 0,$$

and the Hessian becomes

$$H_{ss'}^{(\text{reg})} = H_{ss'} - \epsilon^{(\alpha)} \sum_{n=n_{s-1}+1}^{n_s} \sum_{n'=n_{s'-1}+1}^{n_{s'}} \delta_{nn'}.$$

In the first iteration, we choose $\epsilon^{(\alpha)}$ equal to the absolute value of the largest negative eigenvalue of $H_{ss'}$ and reduce this value in each step by a factor of 3. If one of the eigenvalues of $H_{ss'}^{\text{reg}}$ changes its sign, $\epsilon^{(\alpha)}$ is increased by a factor of 2. The optimization of $\boldsymbol{\alpha}^{(\text{red})}$ is done at most for three cycles of the ν iteration.

It is sometimes necessary to use an under-relaxation strategy also for the cycle defined by Eq. (32). One therefore calculates a provisional

$$x_n^{(\nu+1)} = \overline{x_n^{(\nu)}} + p_x^{(\nu')} \Delta x_n, \tag{38}$$

with

$$\Delta x_n = \sum_{n'=1}^{N} (Q^{(\nu)} + A^{(\nu)})_{nn'}^{-1} [b_{n'}^{(\nu)} - A_{n'n}^{(\nu)} x_{n'}^{(\nu)}]$$

and $p_x^{(\nu')} = \min(1, 2p_x^{(\nu'-1)})$, starting with $\nu' = 1, p_x^{(0)} = 1$. One also calculates

$$\overline{\Delta x}_n = \sum_{n'=1}^{N} (Q^{(\nu)} + A^{(\nu)})_{nn'}^{-1} [b_{n'}(\mathbf{x}^{(\nu+1)}) - A_{n'n}^{(\nu)} x_{n'}^{(\nu+1)}].$$

If the monotony condition

$$\sum_{n=1}^{N} (\overline{\Delta x}_n)^2 \leqslant \left(1 - \frac{p_x^{(\nu')}}{2}\right)^2 \sum_{n=1}^{N} (\Delta x_n)^2 \tag{39}$$

is not fulfilled, one sets $p_x^{(\nu'+1)} = p_x^{(\nu')}/2$ and the test (39) is repeated. Otherwise, the new approximant $\mathbf{x}^{(\nu+1)}$ from Eq. (38) is kept and rescaled: $\overline{x_n^{(\nu+1)}} = x_n^{(\nu+1)} \tilde{\alpha}_n$. With

$$\hat{\alpha}_n^{(\nu)} = \prod_{\nu'=1}^{\nu} \tilde{\alpha}_n^{(\nu')},$$

one obtains $Q_{nn'}^{(\nu+1)} = Q_{nn'}(\mathbf{x}^{(\nu+1)})/(\hat{\alpha}_n^{(\nu+1)} \hat{\alpha}_{n'}^{(\nu+1)})$ and $b_n^{(\nu+1)} = b_n(\mathbf{x}^{(\nu+1)})/\hat{\alpha}_n^{(\nu+1)}$ and repeats the iteration cycle with Eqs. (35) and (33). The iteration is terminated when the norm $\Sigma_n \Delta x_n^2$ becomes sufficiently small.

When the *a priori* estimate for a model parameter is very close to its final fit value, the probability $P_{\text{post}}(\boldsymbol{\alpha}|\mathbf{b})$ is no longer sharply peaked in $\boldsymbol{\alpha}$ space. This expresses the fact that the final fit is very insensitive to the relative weights with which the experimental data and the *a priori* data enter into the procedure. Even though the position of the maximum of P_{post} is then not a good approximation for the average value of $\boldsymbol{\alpha}$, we use it nevertheless because of the considerable numerical effort necessary to perform seven-dimensional averaging integrals over $\boldsymbol{\alpha}$ space. This procedure does not affect the $\overline{\mathbf{x}}$, but the projections s_n^2 depend strongly on $\boldsymbol{\alpha}$. Their values are therefore rather meaningless in our calculations for those model parameters whose fitted value is very close to the corresponding *a priori* value.

When $P_{\text{post}}(\boldsymbol{\alpha})$ is represented by a broad mountain with a small maximum whose position is rather sensitive to small changes of $\mathbf{x}$, we found poor convergence of our iteration procedure where we determine a new $\boldsymbol{\alpha}$ satisfying $P_{\text{post}} = \max$ in each step of the iteration with respect to $\mathbf{x}$, Eq. (32), instead of solving Eqs. (32) and (35) simultaneously. We therefore kept $\boldsymbol{\alpha}$ fixed when $\chi^2_{\text{post}}(\mathbf{x}, \boldsymbol{\alpha})$ appeared to increase, rather than decrease, in successive iterations ν.

To determine the quantities $\hat{s}_n$, Eqs. (35) and (33) are solved with the last $\mathbf{Q}$, $\mathbf{b}$, and $\mathbf{x}$, but with one *a priori* value $x_n^{(0)}$ at a time shifted by a sufficiently large multiple of its standard deviation away from its fit value to ensure that $\hat{s}_n^2$ gets close to 1.

VII. ANALYSIS OF GERMANIUM, GOLD, AND COPPER DATA

The data from Refs. 8 and 9 consisted mostly of the EXAFS function $\boldsymbol{\chi}(k_l)$. In two cases absorption coefficients $\mu(k_l)$ were given. We analyzed them by standard procedures to obtain $\boldsymbol{\chi}(k_l)$. We then treat the EXAFS function in all cases as the input of our fitting procedure. Conventional methods to obtain $\boldsymbol{\chi}$ from the raw data make it very difficult to determine "experimental" errors, let alone cross correlations for these data. We therefore associated, admittedly somewhat *ad hoc*, the following distribution function with the uncertainty of $\boldsymbol{\chi}(k)$:

$$\Delta \boldsymbol{\chi}_l = A_1 \Theta(k_l - k_{\text{cut}}) + A_2 \exp(k_l / k_{\text{exp}}). \tag{40}$$

There are several reasons why the data for very small k should not be included in the fit: Standard procedures to extract the EXAFS function $\boldsymbol{\chi}(k)$ from the measured absorption coefficient $\mu(k)$ [without firm knowledge of $\mu_0(k)$] result in particularly large uncertainties for small k. In addition, our expression for the truncation error $\boldsymbol{\Sigma}$ tends to underestimate the actual truncation error for small k, since convergence of the MS series (1) may not be reached with the one thousand terms, which we included in the construction of $\Sigma_{ll'}$ and because of the restriction of the scattering to a given finite cluster size. The reason is the increase of the mean free path $\lambda(k)$ with decreasing k, allowing increasingly longer scattering paths to contribute with a non-negligible amplitude to the sum (1). One also has to keep in mind the roughly exponential increase of the number of possible MS paths with the number of shells included in the cluster, compensating partly their smaller amplitude. We therefore chose $A_1 = 10$, which is sufficiently large to completely cut off the

TABLE I. Input and output parameters for the fit of germanium data taken at 300 K.

Data points	$L=393$	$k_1=0.2,\ldots,k_L=19.8$ Å^{-1}				
	$L_{\text{eff}}=331$	$k_{\text{cut}}=3.3$ Å^{-1}, $k_{\text{exp}}=28.56$ Å^{-1}, $A_2=0.002$				
Cluster size for scattering	$I=7$	$R_I=7.349$ Å				
Cluster size for vibrations	$I=21$	$A=417$, $R_I=12.89$ Å				
Number of paths	$J=20$	5% Amplitude threshold				
Number of legs $(n_j)_{\max}=4$						
Model parameters	$N=18$	S_0^2	E_0	R_i	κ_1,κ_3	$C_{3,i}$
or	$N=36$	S_0^2	E_0	R_i	σ_j^2	$C_{3,i}$
their *a priori* values	$\mathbf{x}^{(0)}$	1.0	11104 eV	a(fcc)	100 N/m	0.0
			K edge	=5.6574 Å	14 N/m	
or					$\theta_D=360$ K	

data for $k<k_{\text{cut}}$. We choose k_{cut} between 3 and 3.5 Å^{-1} since it turned out that smaller k would lead to serious deviations between the experimental $\chi(k)$ and the one calculated with our choice of the *a priori* model parameters. In particular, retaining smaller k_l in the fit would lead to unacceptably large deviations between fitted scattering half path radii and their values obtained from the well-known lattice constants. The number of input data points is therefore reduced to the L_{eff} points with $k_l \geqslant k_{\text{cut}}$. A lower bound on the parameter A_2 follows from the requirement that Eq. (30) should still have a solution. The parameter k_{exp} is finally obtained from taking $\Delta\chi_L/\chi(k_L)=0.1\%$.

The result of the fit is represented in r space for each scattering path j by a function $d_j(r)$ whose first three cumulants are given by R_j, σ_j^2, and $C_{3,j}$, and the area under which is proportional to the coordination number N_j. Maximum entropy yields the expression[23]

$$d_j(r)=\frac{N_j}{2\pi}\int_{-\infty}^{\infty}\exp\left[-i(r-R_j)k-\frac{\sigma_j^2}{2}k^2-i\frac{C_{3,j}}{6}k^3\right]dk$$
$$=\frac{N_j}{\pi}\int_0^{\infty}e^{-(\sigma_j^2/2)k^2}\cos\left[(r-R_j)k+\frac{C_{3,j}}{6}k^3\right]dk. \quad (41)$$

The integration is done numerically. If $C_{3,j}$ is zero, the integral is of course trivial and yields a Gaussian of width σ_j, peaked at $r=R_j$, and with area N_j. The total fit in r space is given by

$$d(r)=\sum_{j=1}^{J} d_j(r). \quad (42)$$

To find the Einstein frequency $\bar{\omega}_E$ that fits the Debye-Waller parameters σ_i of Eq. (6) optimally, we minimize the expression

$$\chi_\sigma^2(\omega_E)=\sum_{n,n}{}' (\tilde{x}_n+\tilde{x}_n^{(0)}-\tilde{x}_E)(\tilde{Q}+\tilde{A})_{nn'}(\tilde{x}_{n'}+\tilde{x}_{n'}^{(0)}-\tilde{x}_E)$$

with respect to ω_E, where the prime on the sum means that the summations are restricted to those values of n and n' that correspond to the Debye-Waller parameters σ_i of single-scattering paths and

$$\tilde{x}_n^{(0)}=\frac{(\sigma_n^2)^{(0)}}{\hat{x}_n}\tilde{\alpha}_4,$$

$$\tilde{x}_E=\frac{\omega_0}{\omega_E}\ \frac{1+\xi(\omega_E)}{1-\xi(\omega_E)},$$

where $\omega_0=\hbar\tilde{\alpha}_4/(M\hat{x}_n)$ and with the Boltzmann factor $\xi=\exp(-\hbar\omega_E/k_BT)$. From $d\chi_\sigma^2/d\omega_E=0$, one finds the nonlinear equation

$$\frac{1}{\omega_E^2(1-\xi)^2}\left[1-\xi^2+2\xi\frac{\hbar\omega_E}{k_BT}\right]\sum_{nn'}{}' (\tilde{Q}+\tilde{A})_{nn'}\left[\tilde{x}_{n'}+\tilde{x}_{n'}^{(0)}-\frac{\omega_0}{\omega_E}\frac{1+\xi}{1-\xi}\right]=0 \quad (43)$$

for the optimal Einstein frequency $\bar{\omega}_E$. Expanding χ_σ^2 to quadratic order in the vicinity of $\bar{\omega}_E$, one obtains

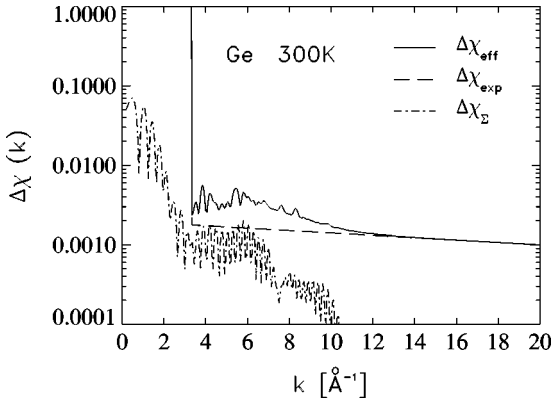

FIG. 3. Errors affecting the input of the fit of Ge data. Full line represents diagonal part of the total error $\Delta\chi_{\text{eff}}$; dashed line represents assumed experimental error $\Delta\chi_I$, Eq. (40); dot-dashed line represents diagonal part of the truncation error $\Sigma_{ll}^{1/2}$.

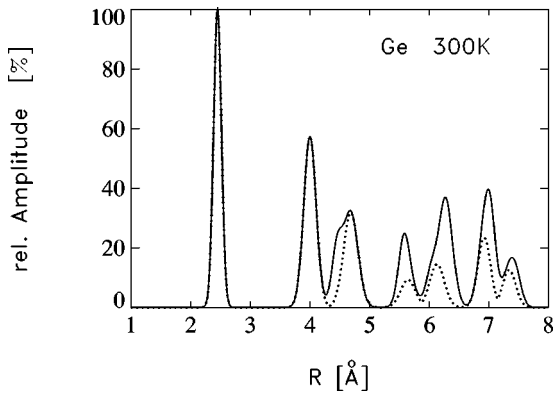

FIG. 4. Fit of the Ge data with freely variable Debye-Waller parameters result in *r* space as sum of Gaussians. Dotted line represents contribution of the single-scattering paths; for the first two peaks the two lines coincide.

$$\chi^2_\sigma(\omega_E)=\chi^2_\sigma(\overline{\omega}_E)+\frac{(\omega_E-\overline{\omega}_E)^2}{\Delta\omega_E^2},$$

with the model uncertainty of the Einstein model,

$$(\Delta\omega_E)^{-2}=\left(\frac{d\tilde{x}_E}{d\omega_E}\right)^2{\sum_{nn'}}'\,(\tilde{Q}+\tilde{A})_{nn'}, \tag{44}$$

where

$$\frac{d\tilde{x}_E}{d\omega_E}=-\frac{\omega_0}{\omega_E^2}\left(1-\xi^2+2\xi\frac{\hbar\omega_E}{k_BT}\right)(1-\xi)^{-2}.$$

A. Analysis of germanium data

The input for the analysis of the germanium data of Ref. 8 is collected in Table I, together with the parameters defining the dimensions of the fit problem. In this case we started from $\mu(k_l)$, from which we obtained $\boldsymbol{\chi}(k_l)$ with the AUTOBK routine, version 2.1.[32] The error function (40), the square root of the diagonal matrix elements $\sqrt{\Sigma_{ll}}$, and the effective error $\Delta\boldsymbol{\chi}_l^{\rm eff}=\Delta g_l^{\rm eff}\Delta\boldsymbol{\chi}_l$ are shown in Fig. 3. A MACMASTER correction of 4.1×10^{-4} Å^2 was used.

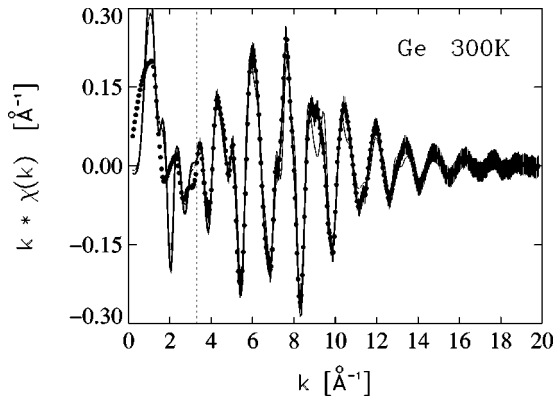

FIG. 5. Fit of the Ge data in *k* space is shown by a thick full line. *A priori* assumptions yield the thin line. Points with error bars are the input data. The vertical dotted line is at $k_{\rm cut}$.

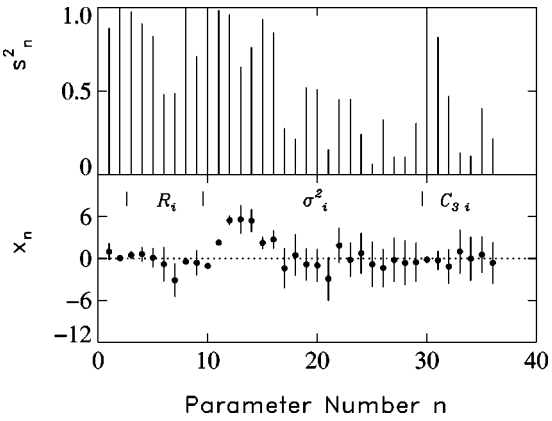

FIG. 6. In the lower frame, the deviations of the fitted model parameters from their *a priori* values with *a posteriori* error bars are shown in dimensionless units for the σ fit of the Ge data. Scaled projections into $\mathcal{R}$ space, s_n^2, are in the upper frame. Ordering of the model parameters from left to right is $S_0, E_0, R_i, \sigma_j^2, C_{3,i}$.

Our *a priori* assumptions for the model parameters are also shown in Table I. The value for E_0 was here and in the following cases obtained from the code ATOMS, version 2.46,[32] and the values for the half path lengths $R_j^{(0)}$ were calculated from the lattice constant *a*, assuming an ideal lattice. Values for the lattice constants were taken here and in the following cases from Ref. 33. *A priori* values for the DW parameters $\sigma_j^{2(0)}$ were derived from the correlated Debye model with the Debye temperature $\theta_D=360$ K (from Ref.

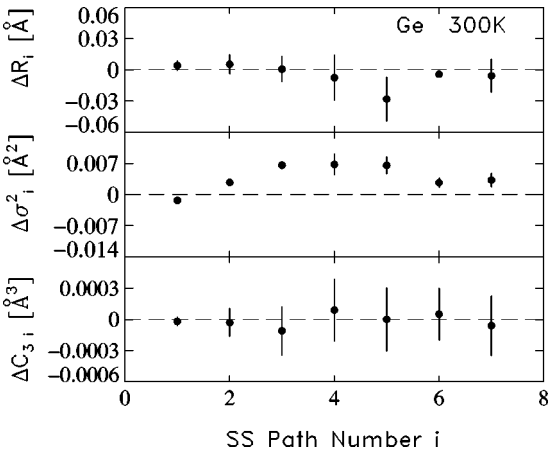

FIG. 7. Deviation of the fitted model parameters from their *a priori* values in absolute units for the single-scattering paths after the parameters S_0 and E_0 have been integrated out.

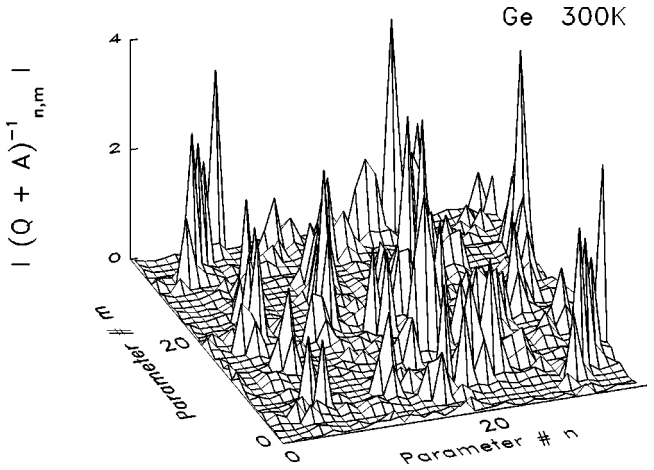

FIG. 8. Modulus of the nondiagonal matrix elements of the correlation matrix $(\mathbf{Q}+\mathbf{A})^{-1}$ for the σ fit of Ge data.

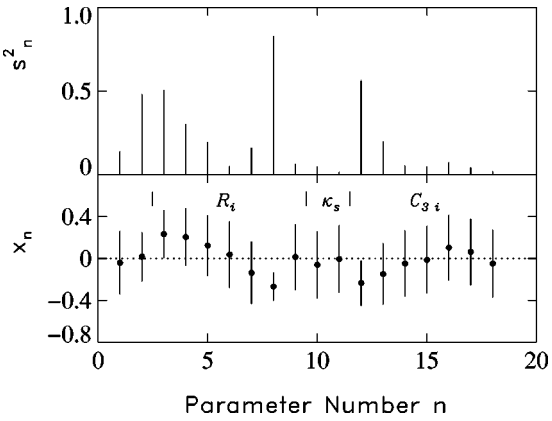

FIG. 9. In the lower frame, the deviations of the fitted model parameters from their *a priori* values with *a posteriori* error bars are shown in dimensionless units for the κ fit of the Ge data. Scaled projections into $\mathcal{R}$ space, s_n^2, are in the upper frame. Ordering of the model parameters from left to right is $S_0, E_0, R_i, \kappa_s, C_{3,i}$.

34). The *a priori* spring constants were obtained from plots like the one of Fig. 2, using the first two σ_j^2 of the preceding fit to determine $\kappa_1^{(0)}$ and $\kappa_3^{(0)}$.

The result of the fit with the free DW parameters is shown in *r* space in Fig. 4 in terms of a sum of Gaussians. The dashed line gives the contribution of the single-scattering paths only. The fit in *k* space is presented in Fig. 5 together with the *a priori* EXAFS function $\boldsymbol{\chi}_{\text{prior}}$. Also shown are the input data with their errors $\Delta\boldsymbol{\chi}_l^{\text{eff}}$. We emphasize again that our fit minimizes χ^2_{post} and not the *R* function of Ref. 8. We do not only associate different weights with different data points k_l, but also take cross correlations into account, which cannot be represented in this figure.

More detailed information on this fit is presented in Figs. 6–8. The deviation of the fitted model parameters from their *a priori* values is shown in Fig. 6 in terms of the dimensionless quantities $\overline{x_n}$ together with their *a posteriori* errors. It is seen that the $\overline{x_n}$, which correspond to S_0^2, E_0, R_i, and $C_{3,i}$ remain zero within their error bars. Only the σ_j^2 have been shifted away from their *a priori* values. Therefore something ''new'' has been learned from the data only for the DW parameters. The s_n^2 are therefore close to 1 for all single-scattering paths. But even the very small shifts of the radii compared to their *a priori* values lead to fairly large s_n^2 values in most cases. That they are not all equal to 1 does not mean that the measurement would not be very sensitive to the half path radii. It only means that the assumed *a priori* values for the radii were perfectly compatible with the data. If the *a priori* values for the R_i were artificially moved away from their actual fit values by 8 standard deviations, the resulting $\hat{s}_n^2$ become larger than 0.9, which shows that the radii are indeed well determined by the data. For $\sigma_1^2, \ldots, \sigma_7^2$ already a shift of 12 standard deviations is needed to bring the corresponding $\hat{s}_n^2$ above 0.9.

The dimension $i_{\mathcal{R}}$ of $\mathcal{R}$ space is shown in Table II. Also given for comparison is the number of independent data points, N_d (calculated with k_{cut} instead of $k_{\min}$). It is seen to be much larger than the dimension of $\mathcal{R}$ space. The condition number $z_{\text{cond}}=q_1/q_N$ of the information matrix $Q_{nn'}$ in Table II shows the ill-conditioned character of the 36-dimensional fit problem. In Fig. 7 the shift of the model parameters away from their *a priori* values is presented in absolute units for the single-scattering paths after the parameters S_0^2 and E_0 have been integrated out, Eq. (29). The modulus of the nondiagonal matrix elements of the variance matrix $(Q+A)^{-1}_{nn'}$ is plotted in Fig. 8. It shows strong correlations between the R_i and the multiple-scattering σ_j^2 as well as the $C_{3,i}$. The optimal fit of the σ_j^2 to a correlated Einstein model yields the Einstein temperature shown in Table II, together with its uncertainty $\Delta\theta_E$.

The result of the second fit, involving the spring constants κ_1 and κ_3 instead of the DW parameters, is in *r* space as well as in *k* space almost indistinguishable from Figs. 4 and 5, respectively. The details of the differences between the two fits can be seen in Fig. 9. The smaller number of available fit

TABLE II. Dimension of the model-parameter space $i_{\mathcal{R}}$; number of independent data points N_d; condition number z_{cond} of $\mathbf{Q}$; fitted force-field parameters κ_1, κ_3 for Ge and κ_1, κ_2 for Au; Einstein temperature θ_E; and anharmonicity parameter ϖ_E of Eq. (5).

		$i_{\mathcal{R}}$	N_d	z_{cond}	κ_1 [N/m]	$\kappa_{2,3}$ [N/m]	θ_E [K]	ϖ_E[eV Å^{-3}]
Ge	σ	20	55	6×10^6			272.2 ± 2.7	0.43 ± 1.15
	κ	3	55	1×10^4	99.9 ± 0.4	14.0 ± 0.3		
Au	σ	18	41	8×10^6			127.8 ± 1.9	-1.56 ± 0.28
	κ	9	41	2×10^4	31.0 ± 0.7	-11.5 ± 2.0		

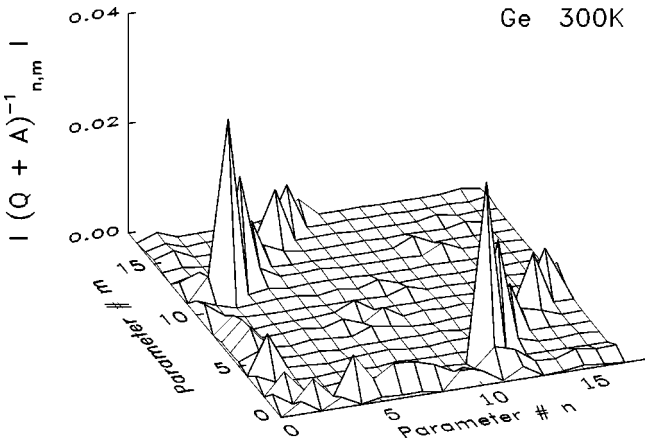

FIG. 10. Same as Fig. 8, but for the κ fit of the Ge data.

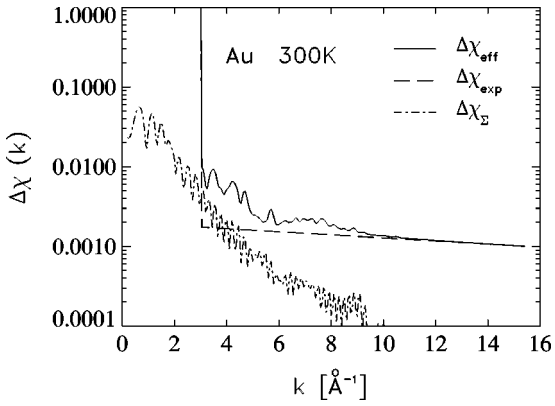

FIG. 11. Same as Fig. 3, for the Au data.

parameters has the consequence that R_6 is now shifted away from its *a priori* value. Consequently, also the corresponding s_8^2 is now close to 1 since the data appear to have shifted it. If one insists that at least the first half path length R_1 should not deviate from its *a priori* value, calculated with the well-known lattice constant for germanium, one has to conclude that the model used for the fit was not quite appropriate. It may, for instance, not be allowed to disregard $\sigma_j^{\rm disor}$ as we did implicitly when we identified σ_j with $\sigma_j^{\rm therm}$ in the second fit.

We find $\kappa_1=99.9$ N/m and $\kappa_3=14.0$ N/m in disagreement with Keating's result of $\kappa_1=228$ N/m, $\kappa_3=77$ N/m.[14] However, Keating fitted to the elastic parameters, i.e., to the acoustic modes, whereas EXAFS is predominantly sensitive to the optical-phonon spectrum. This may explain some of the discrepancy. In Ref. 6, where the same EXAFS data were fitted with a force-field model involving three bond-stretching constants $\kappa_1=120$ N/m, $\kappa_2=4$ N/m, $\kappa_3^{\rm stretching}=-1.1$ N/m, besides a bond-bending mode, reasonable agreement with our κ_1 is found. In summary, we conclude that the data of Ref. 8 may require some contribution of $\sigma_j^{\rm disor}$ to the DW parameters. The $C_{3,i}$ vanish in both analyses within their error bars. Anharmonicities are therefore not clearly seen in these data.

The modulus of the nondiagonal matrix elements of the error-correlation matrix are shown in Fig. 10. There are now strong correlations only between R_i and $C_{3,i}$ for all i.

B. Analysis of gold data

The input data and *a priori* values of the model parameters for the fit of the L_3-edge gold data from Ref. 8 are summarized in Table III. This time we started the analysis with the EXAFS function $\boldsymbol{\chi}(k_l)$ with which we associated the error functions shown in Fig. 11. A MACMASTER correction of 4×10^{-4} Å^2 was used.

Again, the input was analyzed once with all Debye-Waller parameters as freely varying model parameters and in a second run with two bond-streching spring constants κ_1 and κ_2. Though for a metal such a two-body force-field model is not obvious, we follow a practice often used in the literature,[35] considering it as an effective dynamical model. The *a priori* values of the model parameters were determined as in the previous case and are collected in Table III.

The result of the fits in r space and k space is shown in Figs. 12 and 13, respectively. The difference between the two fits is not visible in either of these plots. A more detailed picture of the fits is presented in Fig. 14, which shows the shift of the (dimensionless) model parameters with respect to their *a priori* values in the σ fit and in Fig. 15, in which the shifts are given in absolute units for the κ fit. In both cases,

TABLE III. Input and output parameters for the fit of gold data taken at 300 K.

Data points	$L=305$	$k_1=0.2$ Å^{-1}, . . . , $k_L=15.4$ Å^{-1}				
	$L_{\rm eff}=249$	$k_{\rm cut}=3.0$ Å^{-1}, $k_{\rm exp}=22.22$ Å^{-1}, $A_2=0.002$				
Cluster size for scattering	$I=7$	$R_I=7.63$ Å				
Cluster size for vibrations	$I=18$	$A=500$, $R_I=12.57$ Å				
Number of paths	$J=25$	5% Amplitude threshold				
Number of legs $(n_j)_{\rm max}=4$						
Model parameters	$N=18$	S_0^2	E_0	R_i	κ_1,κ_2	$C_{3,i}$
or	$N=41$	S_0^2	E_0	R_i	σ_j^2	$C_{3,i}$
their *a priori* values	$\mathbf{x}^{(0)}$	1.0	11918 eV	a(fcc)	30.7 N/m,	0.0
			L_3 edge	=4.078 Å	-12 N/m	
or					$\theta_D=180$ K	

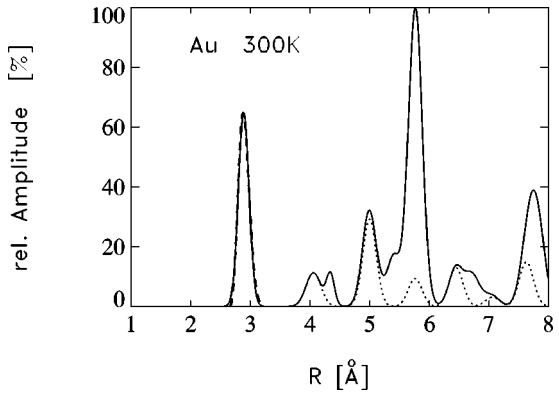

FIG. 12. Fit of the Au data with freely variable Debye-Waller parameters result in r space as sum of Gaussians. Dotted line represents the contribution of the single-scattering paths, for the first two peaks the dotted line and the full line coincide. Dashed line represents the inclusion of an anharmonicity according to Eq. (41) for the first peak.

R_1 is shifted upwards by slightly more than 1 standard deviation. This is surprising since the lattice constant should yield a reliable nearest-neighbor distance. However, R_1 is also contained in the half path lengths of several multiple-scattering paths. The truncation error $\Sigma_{ll'}$ connected with these paths may not be properly accounted for. Also the Debye-Waller parameter σ_1^2 of the first single-scattering path is significantly shifted away from the correlated Debye estimate. The anharmonicity parameter of this path is different from 0 in both fits.

The values of these shifts in absolute units are, however, rather small. In both fits, R_1 is shifted only by some 10^{-2} Å. The anharmonicity is barely visible in the first peak of Fig. 12, indicated by the long-dashed line, though the curve was obtained from the full expression (41). From $C_{3,1}$, an anharmonicity term ϖ_E in the force field was calculated with Eq. (13). It is shown in the last column of Table II. For comparison, one can roughly determine ϖ_E within the Einstein model from the empirical, macroscopic, thermal expansion coefficient $\alpha = d \ln r/dT$ using Eq. (8). To provide the necessary microscopic length scale, we divide Eq. (8) by the distance of the nearest neighbor R_1. Using Table 4f-1 of Ref. 36, we obtain $\varpi_E = -1.32$ eVÅ^{-3}. We also consider the ratio between the sizes of the anharmonic and the harmonic term of the Einstein potential (5) at the average amplitude σ_E, Eq. (6), for the given temperature: $R_E = 4|\varpi_E|\sigma_E/(\omega_E^2 M) = 0.1$. This quantity is more significant for the anharmonicity than the tiny effect of the latter on the distribution function of Fig. 12.

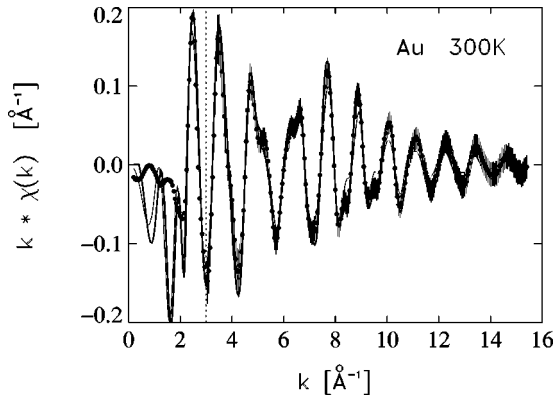

FIG. 13. Same as Fig. 5, but for the Au data.

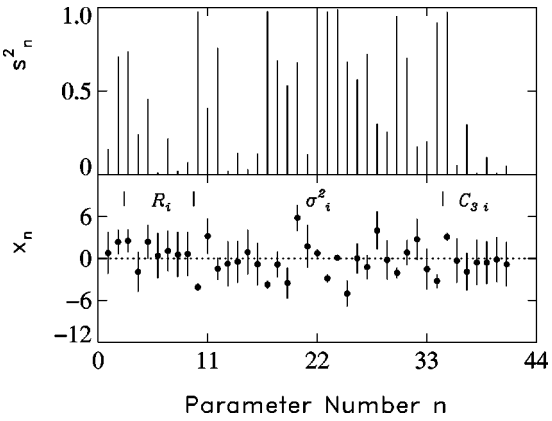

FIG. 14. Same as Fig. 6, but for the Au data.

Again the spring constants $\kappa_1 = (31.0 \pm 0.7)$ N/m and $\kappa_2 = (-11.5 \pm 2.0)$ N/m come out considerably smaller than in a fit to the elastic constants. From the fitted Debye-Waller parameters, an effective Einstein temperature θ_E was calculated. It is shown together with its uncertainty $\Delta\theta_E$ in Table II.

The modulus of the error correlation matrix elements $|(Q+A)^{-1}_{nn'}|, n \neq n'$, is shown in Figs. 16 and 17 for the two fits. In the σ fit, there are correlations between all model parameters; in particular, between R_i and $C_{3,i}$, and among the σ's. For the κ fit, there are strong correlations between R_i and $C_{3,i}$ for all i and between the two κ_s's.

C. Analysis of copper data

We use data taken on copper at three different temperatures: EXAFS functions $\chi(k_l)$ at 10 K and 300 K from

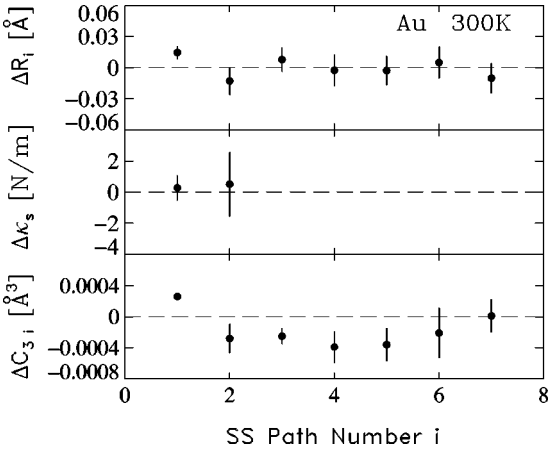

FIG. 15. For the κ fit of the Au data; deviation of the fitted model parameters from their *a priori* values in absolute units for the single-scattering paths after the parameters S_0 and E_0 have been integrated out.

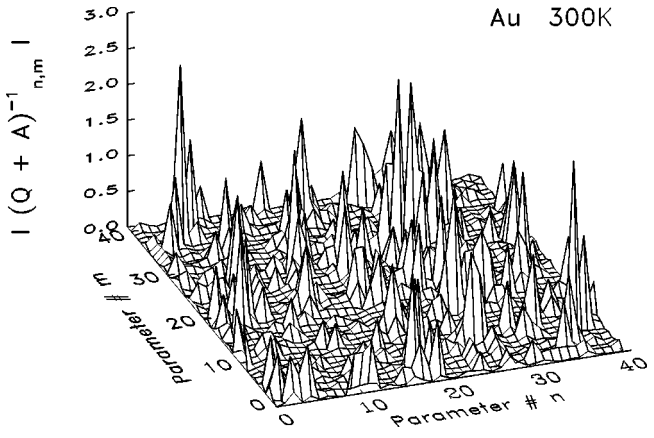

FIG. 16. Modulus of the nondiagonal matrix elements of the correlation matrix $(\mathbf{Q}+\mathbf{A})^{-1}$ for the σ fit of Au data.

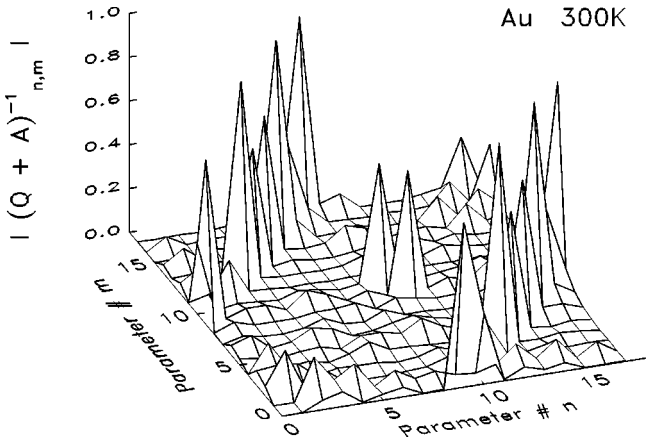

FIG. 17. Same as Fig. 16, but for the κ fit of the Au data.

Newville,[8] and for an intermediate temperature of 80 K we extracted the EXAFS function from Fig. 1 of Ref. 9. In the latter case, larger experimental uncertainties had to be associated with the input data than in other cases in order to obtain reasonable fits at all. In particular, $k_{\mathrm{cut}}=4.4$ Å^{-1} was required. Detailed information on the input parameters, the internal size parameters, and the choice of *a priori* values for the model parameters for the three sets of data is collected in Table IV. The MACMASTER correction used in all three cases was 5.2×10^{-4} Å^2. Again, in each case two analyses are performed, one with all Debye-Waller parameters varied freely and a second in a model with two bond-stretching constants κ_1 and κ_2. A Debye temperature of 315 K was used[34] to generate *a priori* values $(\sigma_j^2)^{(0)}$.

The result of the fit in k space is shown only for the σ fit of the 300-K data in Fig. 18. In Fig. 19, the fit in r space is shown for the κ fits of the 10-K and 300-K data. The temperature broadening of the peaks is clearly seen. A more detailed analysis of the fits shows that for 10 K, all model parameters remain at their *a priori* values in the σ fit and in the κ fit, as expected. Only two of the σ_i^2 and $C_{3,1}$ are slightly shifted, indicating a very small structural disorder. The analysis of the 80-K data shows the same picture, except that the anharmonicity parameter $C_{3,1}$ is now larger and in

TABLE IV. Input and output parameters for the fit of copper data taken at 10, 80, and 300 K.

Data points 10 K	$L=397$	$k_1=0.1$ Å^{-1}, . . . , $k_L=19.9$ Å^{-1}				
	$L_{\mathrm{eff}}=328$	$k_{\mathrm{cut}}=3.5$ Å^{-1}, $k_{\mathrm{exp}}=6.643$ Å^{-1}, $A_2=0.02$				
80 K	$L=288$	$k_1=1.9$ Å^{-1}, . . . , $k_L=16.25$ Å^{-1}				
	$L_{\mathrm{eff}}=238$	$k_{\mathrm{cut}}=4.4$ Å^{-1}, $k_{\mathrm{exp}}=8.566$ Å^{-1}, $A_2=0.03$				
300 K	$L=330$	$k_1=1.0$ Å^{-1}, . . . , $k_L=17.45$ Å^{-1}				
	$L_{\mathrm{eff}}=279$	$k_{\mathrm{cut}}=3.5$ Å^{-1}, $k_{\mathrm{exp}}=12.59$ Å^{-1}, $A_2=0.004$				
Cluster size for scattering	$I=9$	$R_I=7.67$ Å				
Cluster size for vibrations	$I=18$	$A=500$, $R_I=11.11$ Å				
Number of paths for 10 K	$J=54$					
80 K	$J=53$	5% Amplitude threshold				
300 K	$J=53$					
Number of legs $(n_j)_{\max}=6$						
Model parameters	$N=22$	S_0^2	E_0	R_i	κ_1,κ_2	$C_{3,i}$
or for 10 K	$N=74$					
80 K	$N=73$	S_0^2	E_0	R_i	σ_j^2	$C_{3,i}$
300 K	$N=73$					
a priori values for 10 K	$\mathbf{x}^{(0)}$	1.0	8979 eV	a(fcc)	28.5 N/m,	0.0
			(K edge)	$=3.6032$ Å	-7.5 N/m	
80 K		1.0	8979 eV	3.6032 Å	20.0 N/m,	0.0
					1.0 N/m	
300 K		1.0	8979 eV	3.6150 Å	27.5 N/m,	0.0
					$-3,7$ N/m	
or					$\theta_D=315$ K	

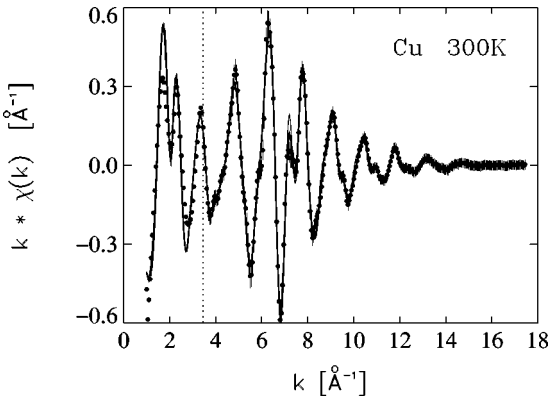

FIG. 18. Same as Fig. 5, but for the Cu data at 300 K.

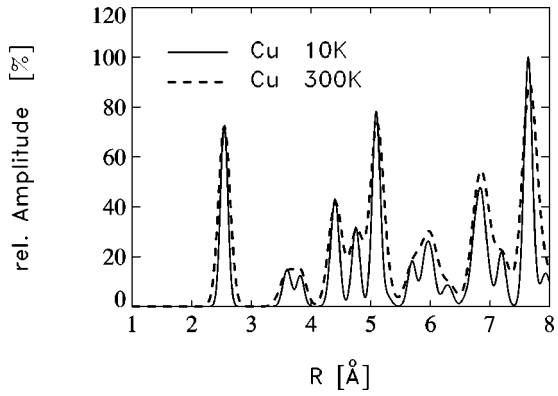

FIG. 19. Result of the fit of Cu data in *r* space; 300-K data represented by dashed line and 10-K data by full line.

the κ fit also R_1 is significanly shifted, which may either be due to structural disorder or an underestimate of the assumed experimental error. For the 300-K data all radii are unchanged, there are some deviations of the Debye-Waller parameters from their predictions in the correlated Debye model, and in both fits $C_{3,1}$ is shifted upwards. To appreciate the apparently irregular temperature dependence of the spring constants, the rather different origin of the data at 80 K compared to the two other sets has to be taken into account. Furthermore, there is a rather large error correlation between κ_1 and κ_2. However, our results show that considerable caution is necessary when force-field parameters are extracted from a single set of data without corroborating evidence from other independent measurements.

The dimension of $\mathcal{R}$ space, for comparison the number of independent data points N_d, the two spring constants, the Einstein temperature and its uncertainty, and the anharmonicity parameter ϖ_E are given in Table V for the three temperatures. The spring constant $\kappa_1 = 27.4 \pm 0.6$ N/m for 300 K may be compared with the value 27.9 N/m, obtained in Ref. 37 with a one-spring-constant model for 295-K copper data. As one might have expected, the effective Einstein temperature is not independent of the temperature of the measurement. The result for the Einstein temperature of the 300-K data is in agreement with $\theta_E = 218$ K, found in Ref. 37 for copper at 295-K. The anharmonicity parameter is obviously most affected by noise. At low temperature, $C_{3,1}$ should not be visible. In fact, our value at 10 K is only slightly more than 1 standard deviation away from zero. For 80 K, our result appears to be spurious. However, our value $C_{3,1} = (1.4 \pm 0.1) \times 10^{-4}$ Å^3 at 300 K compares well with 1.36×10^{-4} Å^3, reported in Table II of Ref. 37. The ratio R_E is 0.046 and the thermal expansion coefficient for copper from Ref. 36 yields $\varpi_E = -1.19$ eV Å^{-3}, again using R_1 as the length scale.

VIII. CONCLUDING REMARKS

A careful analysis of the problem of fitting EXAFS data to structural model parameters shows that one is dealing with a stochastic, usually ill-posed, inverse problem. Its solution requires the use of Bayes' theorem, which needs the introduction of *a priori* information into the fitting procedure. The fine tuning of the relative weights α_n with which the experimental data and the *a priori* data enter into the fit is achieved by slightly generalizing a procedure first proposed by Turchin and Nozik. Since the relation between the observables $\boldsymbol{\chi}(k_l)$ and the model parameters x_n is nonlinear, an algorithm is described for its reduction to solving a sequence of linear, well-posed equations. This nonlinear problem is coupled to an additional set of nonlinear equations to determine the α_n. The convergence of our global iteration scheme is discussed and its convergence is tested. Special attention was given to the limiting case where the *a priori* data are very close to the final fit. For the iteration with respect to the model parameters x_n, we typically found convergence after 10–20 iteration steps in our applications.

These applications involve data from monoatomic systems, where the radii should be well described by their lattice

TABLE V. Dimension of the model parameter space $i_{\mathcal{R}}$, number of independent data points N_d, condition number z_{cond} of **Q**, fitted force-field parameters κ_1, κ_2, Einstein temperature θ_E, and anharmonicity parameter ϖ_E of Eq. (5) for copper at three different temperatures.

		$i_{\mathcal{R}}$	N_d	z_{cond}	κ_1 [N/m]	κ_2 [N/m]	θ_E [K]	ϖ_E [eV Å^{-3}]
10 K	σ	17	58	2×10^{12}			184.7 ± 5.2	1.9 ± 1.2
	κ	8	58	2×10^{6}	28.3 ± 1.4	-8.3 ± 0.5		
80 K	σ	16	43	2×10^{13}			187.9 ± 6.2	4.0 ± 1.5
	κ	10	43	2×10^{4}	19.3 ± 2.5	0.9 ± 3.7		
300 K	σ	13	50	2×10^{13}			215.5 ± 1.5	-0.6 ± 0.2
	κ	3	50	2×10^{7}	27.4 ± 0.6	-4.0 ± 0.3		

constants. The emphasis of the analysis was therefore on two questions: (i) How well can the Debye-Waller parameters σ_j^2 be described either by the correlated Debye model or a force-field model with two harmonic coupling parameters; and (ii) how reliable is the determination of anharmonicity parameters in these fits.

The main limitation of the present analysis comes from the poor convergence of the multiple-scattering series (1) in the x-ray absorption near-edge structure region, which required the elimination of all data with $k<k_{\mathrm{cut}}$ from the fit. Use of alternative representations of the EXAFS function, now available,[38,39] should lead to a more complete utilization of the experimental data. A second severe drawback of any error analysis is the considerable uncertainty in associating an ''experimental'' error with the input EXAFS function $\boldsymbol{\chi}(k_l)$. An important contribution to the uncertainty comes from the fact that the zero-order contribution $\mu_0(k)$ to the multiple scattering signal cannot be obtained from FEFF with sufficient accuracy. Klementiev[40] proposed to determine the difference between the true and the FEFF expression for μ_0 from the data, introducing some smoothness requirement as *a priori* assumption. This increases the number of model parameters and it remains to be seen what the cross correlations are between these additional parameters and those structure parameters in which one is primarily interested.

For the application of our Bayes-Turchin approach, we chose data of high quality, taken on systems with a simple, rather undisturbed structure. This is useful to check our method of data analysis, but it is not necessarily typical for investigations involving EXAFS measurements. Often the data are ''dirtier,'' and one would therefore not expect to derive more model parameters from the measurement than, say, a nearest-neighbor distance R_1 and the average coordination number N_1 of the first shell. Of course, our method can be applied also to such data, only the size and relative weight of the uncertainties will be different from the examples considered above. The dimension of the model-parameter space can be chosen much smaller, which reduces the numerical effort considerably and brings it closer to the numerical needs of typical nonlinear least-squares-fitting procedures.

ACKNOWLEDGMENTS

We thank J. Rehr for valuable discussions and his continuing encouragement of this investigation, M. Newville for allowing us to use partly unpublished data, and R. Lipperheide for a careful reading of the manuscript.

[1] S.I. Zabinsky, J.J. Rehr, A. Ankudinov, R. Albers, and M.J. Eller, Phys. Rev. B **52**, 2995 (1995); A. L. Ankudinov, Ph.D. thesis, University of Washington, 1996.

[2] H.J. Krappe and H.H. Rossner, Phys. Rev. B **61**, 6596 (2000).

[3] J. Skilling, in *Maximum Entropy in Action*, edited by B. Buck and V. A. Macaulay (Oxford University, Oxford, 1991), Chap. 2.

[4] V.F. Turchin, V.P. Kozlov, and M.S. Malkevich, Usp. Fiz. Nauk **102**, 345 (1970) [Sov. Phys. Usp. **13**, 681 (1971)]; V. F. Turchin, in *Advanced Methods in the Evaluation of Nuclear Scattering Data*, Lecture Notes in Physics Vol. 236 (Springer, Berlin, 1985), p. 33.

[5] A.V. Poiarkova and J.J. Rehr, Phys. Rev. B **59**, 948 (1999).

[6] A.V. Poiarkova and J.J. Rehr, J. Synchrotron Radiat. **8**, 313 (2001).

[7] T. Lederer, D. Arvanitis, G. Comelli, L. Tröger, and K. Baberschke, Phys. Rev. B **48**, 15 390 (1993).

[8] M. Newville, Ph.D. thesis, University of Washington, 1994; (private communication).

[9] E.A. Stern, B.A. Bunker, and S.M. Heald, Phys. Rev. B **21**, 5521 (1980).

[10] E. A. Stern, in *X-Ray Absorption*, edited by D. C. Koningsberger and R. Prins (Wiley, New York, 1988), p. 3.

[11] G. Bunker, Nucl. Instrum. Methods Phys. Res. **207**, 437 (1983).

[12] J. A. Victoreen, J. Appl. Phys. **19**, 855 (1948); *International Tables for X-Ray Cristallography III*, edited by K. Lonsdale *et al.*, (Kynoch, Birmingham, England, 1962), Sec. 3.2.

[13] M. Newville, P. Livins, Y. Yacoby, J.J. Rehr, and E.A. Stern, Phys. Rev. B **47**, 14 126 (1993).

[14] P.N. Keating, Phys. Rev. **145**, 637 (1966).

[15] R. Haydock, V. Heine, and M.J. Kelly, J. Phys. C **8**, 2591 (1975).

[16] P. Deuflhard and A. Hohmann, *Numerical Analysis* (de Gruyter, Berlin, 1995), Secs. 4.2 and 8.5.

[17] O. Perron, *Die Lehre von den Kettenbrüchen* (Teubner, Stuttgart, 1957), Vol. 2, Sec. 1.

[18] G. Szegö, *Orthogonal Polynomials* (American Mathematical Society, New York, 1959), Chap. 3.

[19] A.I. Frenkel and J.J. Rehr, Phys. Rev. B **48**, 585 (1993).

[20] J. J. Rehr (private communication).

[21] V.F. Turchin and V.Z. Nozik, Akad. Nauk. SSSR, ser. FIZ. atm. i okeana **5**, 29 (1969) [Bull. Acad. Sci. USSR, Atmos. Oceanic Phys. **5**, 14 (1969)].

[22] H. Richter, *Wahrscheinlichkeitstheorie* (Springer, Berlin, 1966), Chap. 3.

[23] A. Tagliani, J. Math. Phys. **34**, 326 (1993).

[24] A. N. Tikhonov and V. Y. Arsenin, *Solutions of Ill-Posed Problems* (Winston, Washington, DC, 1977).

[25] V.F. Turchin, Zh. vychisl. matem. i matem. fiz. **7**, 1270 (1967) [USSR Comput. Math. Math. Phys. **7**, 79 (1967)].

[26] E.A. Stern, Phys. Rev. B **48**, 9825 (1993).

[27] N. Newville, Physics Department of the University of Washington, FEFFIT manual, 1995 (unpublished).

[28] H.J. Krappe and H.H. Rossner, Z. Phys. A **314**, 149 (1983).

[29] P. Deuflhard, H. Engel, and O. Scherzer, Inverse Probl. **14**, 1081 (1998).

[30] D.W. Marquardt, J. Soc. Ind. Appl. Math. **11**, 431 (1963).

[31] J. E. Dennis, Jr. and R. B. Schnabel, *Numerical Methods for*

Unconstrained Optimization and Nonlinear Equations (Prentice-Hall, Englewood Cliffs, NJ, 1983), Chap. 6.
[32] E. A. Stern, M. Newville, B. Ravel, Y. Yacoby, D. Haskel, Physica B **208&209**, 117 (1995).
[33] R. W. G. Wyckoff, *Crystal Structures* (Wiley, New York, 1963), Vol. 1.
[34] N. W. Ashcroft and N. D. Mermin, *Solid State Physics* (Holt, Rinehart and Winston, New York, 1976).
[35] E. Sevillano, H. Meuth, and J.J. Rehr, Phys. Rev. B **20**, 4908 (1979).
[36] *American Institute of Physics Handbook*, 3rd ed., edited by D. E. Gray (McGraw-Hill, New York, 1972).
[37] Nguyen Van Hung and J.J. Rehr, Phys. Rev. B **56**, 43 (1997).
[38] J.J. Rehr and R.C. Albers, Rev. Mod. Phys. **72**, 621 (2000).
[39] F.J. Garcia de Abajo, M.A. Van Hove, and C.S. Fadley, Phys. Rev. B **63**, 075404 (2001).
[40] K.V. Klementiev, J.Phys. D, J. Phys. D **34**, 209 (2001).