Iterative reconstruction of a refractive-index profile from x-ray or neutron reflectivity measurements

Thorsten Hohage*
Institut für Numerische und Angewandte Mathematik, Lotzestrasse 16–18, 37085 Göttingen, Germany

Klaus Giewekemeyer† and Tim Salditt‡
Institut für Röntgenphysik, Friedrich-Hund-Platz 1, 37077 Göttingen, Germany

(Received 25 January 2008; revised manuscript received 22 April 2008; published 16 May 2008)

Analysis of x-ray and neutron reflectivity is usually performed by modeling the density profile of the sample and performing a least square fit to the measured (phaseless) reflectivity data. Here we address the uniqueness of the reflectivity problem as well as its numerical reconstruction. In particular, we derive conditions for uniqueness, which are applicable in the kinematic limit (Born approximation), and for the most relevant case of box model profiles with Gaussian roughness. At the same time we present an iterative method to reconstruct the profile based on regularization methods. The method is successfully implemented and tested both on simulated and real experimental data.

DOI: 10.1103/PhysRevE.77.051604 PACS number(s): 68.37.−d, 02.30.Zz, 61.05.cm, 61.05.fj

I. INTRODUCTION

X-ray and neutron reflectivity are well established techniques to study the interface profile $\rho(z)$ of thin films with molecular resolution [1,2]. Organic thin films, such as polymer layers [3], liquid wetting layers [4], or biomolecular films [5–7] on solid surfaces are well known examples, where reflectivity methods are employed. More generally, reflectivity is an important tool in the structure analysis of thin films, both in soft and hard condensed matter. In most applications, some part of $\rho(z)$ is known, i.e., the solid substrate (e.g., silicon wafer, glass slide, etc.), while most of the profile, for example, the polymer film, the lipid membrane, or the wetting layer is unknown and needs to be elucidated. In this paper we investigate from a mathematical point of view to which extent partial information on the profile can be used to uniquely determine the full profile $\rho(z)$ of the sample.

Specular reflectivity is the measured intensity $R(q)$ reflected in the specular direction (incidence angle equal to the exit angle) off a planar interface, normalized to the intensity of the incident beam. Reflectivity is measured as a function of the incidence angle α or correspondingly the momentum transfer $q=4\pi/\lambda \sin(\alpha)$ normal to the interface. $R(q)$ is uniquely determined by the laterally averaged scattering length density profile $\rho(z)$ normal to the interface. While the calculation of the $R(q)$ from $\rho(z)$ is straightforward (the forward problem), the determination of $\rho(z)$ from $R(q)$ (the inverse problem) is not. At the same time the solution of the inverse problem would make ambiguous and cumbersome data fitting obsolete, and would provide a unique scattering length density depth profile for the thin-film structure, if the conditions of uniqueness are fulfilled.

Usually, the forward problem to calculate $R(q)$ is carried out either by solving the Helmholtz equation for stratified media, i.e., for piecewise constant density profiles $\rho(z)$ parametrized by box models [8], or by solving the kinematic master equation of reflectivity [9,10], which is better suited for smooth profiles and profiles, which cannot easily be parametrized by box models or which would contain too many boxes. In the first approach of dynamic reflectivity the calculations are performed, e.g., by use of optical matrices based on full dynamic theory including multiple reflections. In the second approach of kinematic reflectivity (corresponding to the Born or weak scattering approximation), the measurable reflectivity is given by [1]

$$R(q) = R_F(q)|r(q)|^2 \tag{1a}$$

with the Fresnel reflectivity $R_F(q)$ of an ideal interface between two media with density contrast $\Delta\rho=\rho(\infty)-\rho(-\infty)$ and a structure factor $r(q)$, which is the Fourier transform of the derivative of ρ:

$$r(q) = \frac{1}{\Delta\rho}\int_{-\infty}^{\infty} \exp(-iqz)\rho'(z)dz. \tag{1b}$$

The inverse problem is therefore reduced to determine the missing phase information of the complex valued function $r(q)$, or in other words to determine the unknown phase information from the measured intensities plus supplemental information.

In full dynamic theory, in some cases not even knowledge of the full complex reflectivity amplitude is sufficient to recover ρ. By analogy of the Helmholtz equation to the time independent Schrödinger equation it has been known for a long time that, if bound states are present, knowledge of the phase shifts alone is not sufficient for a unique determination of the potential [11,12]. On the other hand, if no bound states exist, ρ can be reconstructed from the complex reflectivity amplitude by Marchenko inversion [13,14]. The inverse scattering problem is thus quite different for both approaches. The inversion of dynamic reflectivity has been addressed, for

*hohage@math.uni-goettingen.de; URL: www.num.math.uni-goettingen.de/hohage/
†Klaus.Giewekemeyer@phys.uni-goettingen.de; URL: www.roentgen.physik.uni-goettingen.de/
‡tsaldit@gwdg.de; URL: www.roentgen.physik.uni-goettingen.de/

example, by Majkrzak, Berk, and co-workers [15]. They have shown that the complex amplitude for neutrons specularly reflected from a film can be determined exactly through the use of buried reference layers of finite thickness or by variation of the incident or substrate medium, i.e., by combination of several measurements using contrast variation.

In certain ideal situations explicit formulas to recover the missing phase of r exist: If $\rho'(z)$ has compact support, its Fourier transform $r(q)$ has a holomorphic extension to an entire function, which by the Hadamard factorization theorem is essentially determined by its complex zeros $\{\alpha_j : j \in \mathbb{N}\}$. Since R/R_F has the analytic extension $(R/R_F)(q) = r(q)r(\overline{q})$, its set of zeros is given by $\{\alpha_j\} \cup \{\overline{\alpha_j}\}$. Therefore, for a given zero α of R/R_F it is not clear, whether α or $\bar{\alpha}$ is the corresponding zero of r. However, under certain conditions it can be guaranteed that r has no zeros in the complex upper half-plane U. Then the missing phase arg r is given explicitly by the so-called Hilbert-phase [16–18]

$$r_H(q) = \frac{2q}{\pi} \int_0^\infty \frac{\ln|r(\tilde{q})/r(q)|}{\tilde{q}^2 - q^2} d\tilde{q} \tag{2}$$

up to an affine linear function, i.e., arg $r(q) = r_H(q) + n\pi + z_0 q$ with a translation $z_0 \in \mathbb{R}$ and a sign indication $n \in \{0,1\}$. If r does have complex zeros in U, then the additional knowledge of these zeros also determines arg r uniquely up to an affine linear term, and these zeros may be included in correction terms in the phase reconstruction formula above. Variants of Eq. (2) and methods to determine zeros of r can be used for numerical phase reconstructions [2,19–23].

In this paper, we prove a uniqueness result implying identifiability of box model profiles with Gaussian interfaces if the sharpest jump is known, and its size is larger than the sum of all other jumps. Moreover, we extend the inversion techniques for kinematic reflectivity by proposing an iterative regularization method using a general nonparametric model of the profile.

The paper is organized as follows. Section II presents a brief statement of the problem in the framework of kinematic reflectivity, and the mathematical proof of uniqueness. Section III then addresses the numerical implementation of the forward problem, while Sec. IV is devoted to the ill-posed inverse problem and its solution by regularization methods. Finally, Sec. V presents numerical results on real data, before the paper closes with a short discussion.

II. UNIQUENESS

Obviously, ρ' is not uniquely determined by R since Eqs. (1a) and (1b) is also satisfied for the functions $\rho'_f(z) = \frac{\Delta\rho}{2\pi} \int_{-\infty}^{\infty} e^{i[zq+f(q)]} r(q) dq$ with an arbitrary real-valued, measurable function f. However, if additional *a priori* information on ρ' is available, it may be uniquely determined by R as the following result shows.

Theorem 1. Assume that ρ' is of the form

$$\rho'(z) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-(z-\tilde{z})^2/2\sigma^2} \{c\,\delta_{z_0}(\tilde{z}) + h(\tilde{z})\} d\tilde{z}$$

with known $\sigma \geq 0$, $c \in \mathbb{R}$, and $z_0 \geq 0$ and an unknown real-valued function $h \in L^1(\mathbb{R})$ with compact support [for $\sigma = 0$ the right-hand side is $c\,\delta_{z_0}(z) + h(z)$]. Moreover, assume that $c > \|h\|_{L^1}$. Then ρ' is uniquely determined by the values of $R(q)$ in Eqs. (1a) and (1b) for q in some interval $[0, q_{\max}]$.

Proof. Since h has compact support, the function r has a holomorphic extension to an entire function

$$r(q) = \frac{1}{\Delta\rho} e^{-(\sigma q)^2/2} [c e^{-iqz_0} + \hat{h}(q)], \quad q \in \mathbb{C}.$$

Moreover, by analyticity $R(q)$ is uniquely determined for all $q \in \mathbb{C}$ by the values on $[0, q_{\max}]$. Assume that $r(q) = 0$ for some q with $\mathrm{Im}(q) > 0$. Then $|c| \leq |ce^{-iqz_0}| = |\hat{h}(q)| \leq \|\hat{h}\|_{L^1}$ contradicting our assumption. Therefore, r has no zeros in the upper complex half-plane, and hence by the results discussed in the Introduction, ρ' is uniquely determined by R up to the sign and a translation, both of which is uniquely determined by the *a priori* knowledge of c and z_0. ■

As a particular case we obtain the following uniqueness result for box models [24]

$$\rho'(z) = \sum_{j=0}^{M-1} \frac{(\delta\rho)_j}{\sqrt{2\pi}\sigma_j} e^{-(z-z_j)^2/2\sigma_j^2} \tag{3}$$

with M interfaces with widths $\sigma_j \geq 0$ and locations $z_j \in \mathbb{R}$.

Corollary 2. If $(\delta\rho)_0, \sigma_0$, and z_0 in the box model (3) are known and the conditions

$$|(\delta\rho)_0| > \sum_{j=1}^{M-1} |(\delta\rho)_j|, \quad \sigma_0 < \min\{\sigma_1, \ldots, \sigma_{M-1}\}$$

are fulfilled, then ρ' is uniquely determined by R.

Proof. We may apply theorem 1 with $c = (\delta\rho)_0$, $\sigma = \sigma_0$, and $h(z) = \sum_{j=1}^{M-1} \frac{(\delta\rho)_j}{\sqrt{2\pi}\tilde{\sigma}_j} \exp(-\frac{(z-z_j)^2}{2\tilde{\sigma}_j^2})$, where $\tilde{\sigma}_j = \sqrt{\sigma_j^2 - \sigma_0^2}$. ■

Theorem 1 and corollary 2 imply that for (multilayer) structures, which contain a known interface of small width and which are otherwise smooth with variations that are small compared to the size of the known jump, we can expect to be able to recover the refractive index uniquely. For sharp interfaces this has been shown by Clinton [19]. Uniqueness results for other kinds of *a priori* information have been obtain by Klibanov and co-workers [16,25].

III. DISCRETIZED OPERATOR EQUATION

The measured data can be described by m pairs (q_j, y_j^δ) of a deterministic variable q_j determined by the experimentalist and a random variable y_j^δ satisfying

$$\mathbf{E}[|y_j^\delta - (R/R_F)(q_j)|^2] \leq \delta_j^2, \quad j = 1, \ldots, m, \tag{4}$$

where $\mathbf{E}$ denotes the expected value. We assume that bounds δ_j on the expected errors are provided by the experimentalist.

Moreover, we assume that ρ' is real-valued such that R is symmetric. Hence, we can complement our measured data

vectors by the pairs $(-q_j, y_j^\delta)$. Furthermore, the point (0,1) is included in the data vector. Although no measurement at $q_j=0$ can be performed, the forward Eqs. (1a) and (1b) imply this point. We further assume that an interval $[0,a]$ containing the support of ρ' is known. The unknowns of the discretized inverse problem will be the values of ρ' at n equidistant points, $\phi_l=\rho'(la/n)$. Approximating the integral in Eq. (1b) by the trapezoidal rule we obtain

$$r(q) \approx \frac{a}{n(\Delta\rho)} \sum_{l=0}^{N-1} \exp\left(\frac{ilqa}{n}\right)\phi_l, \quad |\tilde{q}_j| \le q_{\max}.$$

To make this approximation accurate we have to choose n such that $q_{\max}a/n \ll 1$ since in this case the integrand is slowly varying in l and the trapezoidal rule is very accurate for smooth periodic functions. We point out that this typically requires to choose the number n of unknowns much larger than the number m of data. This poses no problem as the regularization method discussed below even works for an infinite-dimensional space of unknowns and finitely many data.

Defining the auxiliary grid $\tilde{q}_j=j\Delta_q$ with $\Delta_q=\frac{2\pi n}{aN}$ we obtain

$$\left(\frac{R}{R_F}\right)(\tilde{q}_j) \approx \left| \frac{a}{n(\Delta\rho)} \sum_{l=0}^{N-1} \exp(i2\pi lj/N)\phi_l \right|^2$$

for $j=-\frac{N}{2},\ldots,\frac{N}{2}-1$, and the right-hand side can be computed efficiently by the fast Fourier transform ($\mathcal{F}$) in $O[N\ln(N)]$ operations. (For notational convenience we assume N to be even.) To make Δ_q sufficiently small, we may have to choose $N \ge n$ setting $\phi_l=0$ for $l \ge n$. Since the q_j's are not necessarily equidistant, we use an interpolation matrix $J \in \mathbb{R}^{2m+1\times N}$ such that

$$J\left[\left(\frac{R}{R_F}\right)(\tilde{q}_j)\right]_{j=-N/2,\ldots,N/2-1} \approx \left[\left(\frac{R}{R_F}\right)(q_j)\right]_{j=1,\ldots,2m+1}.$$

More specifically, we approximate R/R_F by the value of the interpolation polynomial of degree 5 using the six points of the grid $\{\tilde{q}_j\}$ closest to q_j. Note that J is a sparse matrix.

Hence, the phase retrieval problem can be written as a discretized operator equation

$$F(\phi) = y^\delta \tag{5}$$

with the data vector $y^\delta=[y_j^\delta]$ and the mapping $F:\mathbb{R}^n \to \mathbb{R}^{2m+1}$ given by $F(\phi)=Jf_{\mathrm{sq}}(\mathcal{F}_N L\phi)$. Here $L:\mathbb{R}^n \to \mathbb{R}^N$ is the mapping, which elongates a vector $[\phi_j]_{j=0,\ldots,n}$ by adding $N-n-1$ zeros at the end, the mapping $f_{\mathrm{sq}}:\mathbb{C}^N \to \mathbb{R}^N$ is defined by $f_{\mathrm{sq}}[v_j]=[|v_j|^2]$, and $(\mathcal{F}_N)_{j,l}=a/[n(\Delta\rho)]\exp(i2\pi lj/N)$ for $j=-N/2,\ldots,N/2-1$ and $l=0,\ldots,(N-1)$.

IV. REGULARIZATION

Since the solution to Eq. (5) is not unique and particularly in the presence of noise many different profiles ρ can explain the data, it is important to incorporate further *a priori* information in the inversion process, e.g., solving the minimization problem

$$\hat{\phi} = \mathrm{argmin}_\phi(\|F(\phi)-y^\delta\|_Y^2 + \alpha\|\phi-\phi_0\|_X^2) \tag{6}$$

for some regularization parameter $\alpha>0$ and an initial guess ϕ_0 featuring, e.g., known interfaces. This method is known as nonlinear Tikhonov regularization [26,27] and can be derived from the perspective of Bayesian statistics by minimizing the posterior distribution [28].

For the performance of the method it is crucial to choose appropriate inner products on $X=\mathbb{R}^n$ and $Y=\mathbb{R}^{2m+1}$. Assuming the errors δ_j to be $N(0,\delta_j)$ distributed and considering the log-likelihood function leads to the choice $\langle y,\tilde{y}\rangle_Y=y^*G_Y\tilde{y}$ with the Gram matrix $G_Y=(2m+1)^{-1}\mathrm{diag}(\delta_1^{-2},\ldots,\delta_{2m+1}^{-2})$ and the corresponding norm $\|y\|_Y^2=\langle y,y\rangle_Y$, i.e., we measure the data misfit by the weighted sum $\|y^{(k)}-y^\delta\|_Y^2=\Sigma_{j=1}^{2m+1}\frac{(y_j^{(k)}-y_j^\delta)^2}{(2m+1)\delta_j^2}$.

The choice of the inner product $\langle\phi,\tilde{\phi}\rangle_X=\phi^*G_X\tilde{\phi}$ in X should reflect the *a priori* information on the solution. From a Bayesian perspective this is called prior modeling. According to theorem 1 we expect uniqueness if the largest variations of ρ' are known and included in the initial guess ρ_0' represented by ϕ_0. Therefore, $\|\rho'-\rho_0'\|_{L^2}^2$ should be small, and we can assume $\rho'(0)=\rho_0'(0)$ and $\rho'(a)=\rho_0'(a)$.

Approximating the derivative by finite differences and using the boundary conditions leads to the Gram matrix

$$\tilde{G}_X = \frac{n}{a}\begin{pmatrix} 2 & -1 & & \\ -1 & 2 & -1 & \\ & \ddots & \ddots & \ddots \\ & -1 & 2 & -1 \\ & & -1 & 2 \end{pmatrix}.$$

Moreover, one typically has good *a priori* knowledge of ρ' in a subintervall $[\alpha,\beta]\subset[0,a]$, and it is important to incorporate this *a priori* information into the inversion method. This can be done by choosing the Gram matrix of the form $G_X=\tilde{G}_X+\mathrm{diag}(w_1,\ldots,w_n)$ where $w_j=1$ if $ja/n \in [\alpha,\beta]$ and $w_j=0$ else.

Since the solution of the nonconvex minimization problem (6) may be difficult and since the evaluation of F using fast Fourier transform is very fast, we have used the Newton conjugated-gradient (CG) method as an alternative to Eq. (6). It consists of solving the Newton equations

$$F'[\phi^{(k)}]\Delta\phi^{(k)} = y^\delta - F(\phi^{(k)}) \tag{7}$$

for the update $\Delta\phi^{(k)}=\phi^{(k+1)}-\phi^{(k)}$ by the conjugate gradient method applied to the normal equation (in the following referred to as "CGNE") using early stopping as regularization both for the inner and the outer iteration. The convergence and regularizing properties of the Newton-CG method were shown by Hanke [29]. For convenience we recall that the CGNE method for minimizing the quadratic functional $x \mapsto \|Tx-y\|_Y^2$ or equivalently the CG method for solving the normal equation $G_X^{-1}T^*G_YTx=G_X^{-1}T^*G_Yy$ can be coded as shown in Fig. 1. T^* denotes the transposed matrix, i.e., the adjoint with respect to the standard inner products, whereas the adjoint with respect to the inner products $\langle\ldots,\ldots\rangle_X$ and $\langle\ldots,\ldots\rangle_Y$ is given by $G_X^{-1}T^*G_Y$. $\rho \in (0,1)$ should be considered a regularization parameter for CGNE with useful value

function $f_{\mathrm{CGNE}}(y,T,T^*)$

$j=0;\ x_0=0;\ s_0=y;\tilde{r}_0=T^*G_Y s_0;\ r_0=G_X^{-1}\tilde{r}_0;\ d_0=r_0;$

while $(\|s_j\|_Y>\rho\|s_0\|_Y$ and $j\le 50)$

$\alpha_j=(r_j^*\tilde{r}_j)/\|Td_j\|_Y^2;\quad \%r_j^*\tilde{r}_j=\|r_j\|_X^2$

$x_{j+1}=x_j+\alpha_j d_j;$

$s_{j+1}=s_j-\alpha T d_j;$

$\tilde{r}_{j+1}=T^*G_Y s_{j+1};\ r_{j+1}=G_X^{-1}\tilde{r}_{j+1};$

$\beta_j=(\tilde{r}_{j+1}^*r_{j+1})/(r_j^*r_j);$

$d_{j+1}=r_j+\beta_j d_j;$

$j=j+1;$

return x_j;

FIG. 1. Inner CGNE iteration in the Newton-CG algorithm.

$\rho=0.8$. The exterior Newton iteration is given by

$$\Delta\phi^{(k)}=f_{\mathrm{CGNE}}(y^\delta-F[\phi^{(k)}],F'[\phi^{(k)}],F'[\phi^{(k)}]^*)$$

with $k=0,1,2,\ldots$, and $\phi^{(k+1)}=\phi^{(k)}+\Delta\phi^{(k)}$. Here the derivative of F at $\phi^{(k)}$ in direction $h\in X$ is given by

$$F'[\phi^{(k)}]h=2J\,\mathrm{Re}[(\overline{\mathcal{F}_N L\phi^{(k)}})*(\mathcal{F}_N Lh)],$$

where $[v_j]*[w_j]=[v_jw_j]$ denotes element-wise multiplication of two vectors and $\mathrm{Re}([v_j])=[\mathrm{Re}(v_j)]$ ($[\overline{v_j}]=[\overline{v_j}]$) means taking the real part (complex conjugate) of all entries of a vector. Moreover,

$$F'[\phi^{(k)}]^*y=2L^*\,\mathrm{Re}\{\mathcal{F}_N^*[(\mathcal{F}_N L\phi^{(k)})*(J^*y)]\}.$$

Note that in the Newton-CG method we never have to set up the typically large and full Jacobian matrices $F'[\phi^{(k)}]$, since the matrix-vector products $x\mapsto F'[\phi^{(k)}]x$ and $y\mapsto F'[\phi^{(k)}]^*y$ can be implemented as routines using the last two formulas. The reconstruction algorithm presented here has been implemented as an open source MATLAB [33] code [34].

V. NUMERICAL RESULTS

For the reconstruction of an unknown structure, the influence of two factors on the quality of the reconstructions is of special interest to the experimentalist: The maximum allowable noise level of the data and the minimum amount of *a priori* information needed to obtain a physically meaningful and unique profile. The influence of these factors will be illustrated for selected examples in the following.

In order to elucidate the problem of possible ambiguity in the obtained reconstructions, the algorithm was first applied to simulated data, prior to applications to experimental data from reflectivity measurements of biomolecular films [5,6]. The simulated data sets were calculated from an exactly defined scattering length density profile $\rho(z)=r_0\rho_e(z)$, where r_0 is the classical electron radius and $\rho_e(z)$ the electron number density. Recall that the Fresnel reflectivity behaves as

$$R_F(q)=I_0\left|\frac{1-\sqrt{1-\chi_q}}{1+\sqrt{1-\chi_q}}\right|^2\sim I_0\frac{(4\pi)^2(\Delta\rho_e r_0)^2}{q^4} \tag{8}$$

as $\chi_q=(q_c/q)^2=4\pi\Delta\rho_e r_0/q^2\to 0$, where q_c is the critical angle of total external reflection and $\Delta\rho_e=\Delta\rho/r_0$ denotes the total electron density difference between the substrate and the bulk phase on top of the sample structure. Note that the kinematic model of reflectivity (1b) is only valid for angles $q>q_c$. However, for typical density differences $\Delta\rho_e$, e.g., with silicon as the substrate, often used in biomolecular applications, and air as the bulk phase, one arrives at $q_c^{(\mathrm{Si/air})}\approx 0.031$ Å^{-1}, substantially smaller than the highest q- value typically obtainable even with a laboratory source ($q_{\max}\approx 0.4$ Å).

The primary beam intensity I_0 was chosen in the range of 10^6 and 10^9 photons impinging on the whole sample in all, thus covering an intensity range from typical fixed-anode laboratory sources to undulator beamlines at third-generation synchrotrons [35]. Synthetic data were generated by the formula $y_j^\delta=Y_j/R_F(q_j)$ where the count data $Y_j\in\mathbb{N}_0$ were drawn from a Poisson distribution with mean $R(q_j)(1+0.02\epsilon_j)$ with independent $N(0,1)$-distributed random variables ϵ_j accounting for several inevitable steps in the analysis of reflectivity data (illumination correction, determination of the critical angle, etc.). Because of Eq. (8) the variances δ_j of y_j^δ grow rapidly with q_j (see Fig. 4, top right).

The model used here represents a typical density profile of a supported lipid bilayer membrane with an aqueous subphase on top of the thin film [5]. As visible in Fig. 2 (top left graph, dashed blue line), the silicon substrate is covered with a native oxide layer of reduced density and a thickness of about 15.5 Å. The prominent features of the lipid bilayer are the dense phospholipid headgroup regions, represented as minor maxima at 20 and 57 Å, as well as the lower-density hydrocarbon core region around the profile minimum at around 40 Å. The headgroup-headgroup distance is thus given as $d_{hh}=37$ Å, representing a rough estimate for the thickness of the bilayer.

At first, the algorithm was applied to a problem with a high signal-to-noise ratio (a synchrotron source of $I_0=7\times 10^9$ is assumed), see Fig. 2. As the initial input the true, simulated profile on a small interval around the sharp decrease in electron density at the substrate-film interface was used. This density gradient represents the maximum density change $(\delta\rho)_0$ referred to in corollary 2. Although the uniqueness condition is not strictly fulfilled here, it is shown in the following that the algorithm yields a very good approximation of the whole true profile even for wrong initial values of $(\delta\rho)_0$. As mentioned in Sec. III, an interval $[\alpha,\beta]$ can be chosen, in which the algorithm is forced to predominantly follow the initial profile. Here it was placed around the oxide layer and the beginning of the gradient $(\delta\rho)_0$ (see the shaded region in Fig. 2, top left graph).

As visible in the top left graph of Fig. 2 the reconstruction (solid red line) fits the true profile used for simulation of the data (dashed blue line) extremely well, given the nonvanishing noise in the simulated data. Similarly, the corresponding reflectivity curve (top right graph, red line) matches the simulated data points very precisely. To illustrate the effect

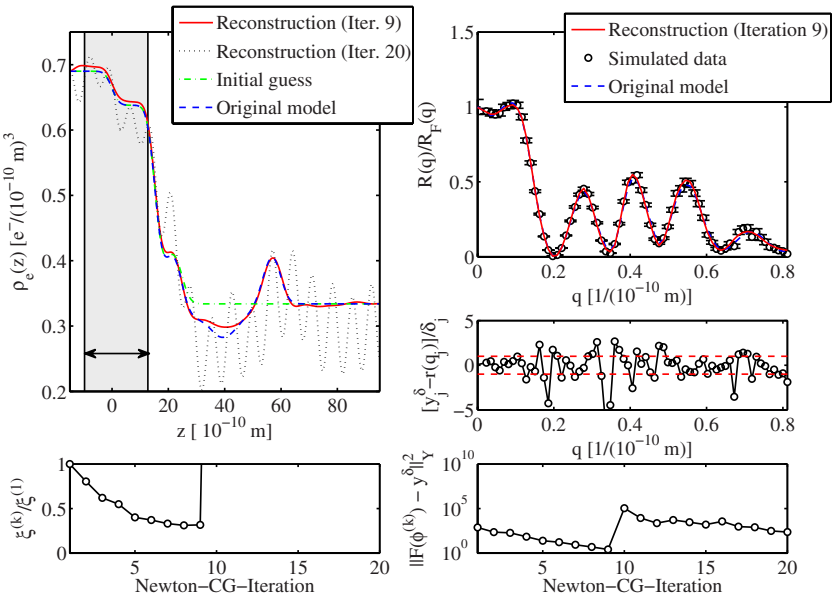

FIG. 2. (Color online) Results from a reconstruction of simulated data assuming a synchrotron radiation beam intensity ($I_0=7\times10^9$ photons). (Top left) Several electron density profiles are shown: The original model used for data simulation (dashed line, blue), resembling a phospholipid bilayer on a solid silicon support, furthermore the initial guess for the reconstruction (dashed-dotted line, green). While for one reconstructed profile (solid line, red) the algorithm was stopped according to the discrepancy principle after 9 iterations, for the other reconstruction (dotted line, black) it was continued for illustrative reasons up to iteration 20. Obviously, overfitting leads to unphysical oscillations in the profile here. The shaded area marks the interval $[\alpha,\beta]$, in which the initial profile was assumed to be a correct representation of the true one. (Top right) Reflectivity curves corresponding to profiles on the left, normalized by Fresnel reflectivity, i.e., the simulated data points with added noise (circles with error bars, black), the simulated reflectivity based on the same model, but without added noise (dashed line, blue) and the reconstruction after 9 iterations, obtained according to the discrepancy principle (solid line, red). (Center right) Residuals of the reconstructed reflectivity, normalized to the expectation value of the simulated errors. The majority of points lies within the error bars. (Bottom left) Misfit $\xi^{(k)}/\xi^{(1)}$ of the real space profile at iteration k with the original model, normalized for clarity to the misfit of the initial profile. Here $\xi^{(k)}$ is defined as $\xi^{(k)}=\|\phi^{(k)}-\phi_{\text{model}}\|_2$, where ϕ_{model} denotes the original model. (Bottom right) Data misfit $\|F(\phi^{(k)})-y^\delta\|_Y^2$. The misfit after iteration 9 is about 2.6, representing a good fit to the data.

of "overfitting" the data (see below), a second reconstruction is shown (top left graph, dotted black line), which was obtained after 20 iterations without applying the discrepancy principle. As shown in the bottom right graph of Fig. 2, the algorithm converged to a good approximation with a geometric decrease of $\|F(\phi^{(k)})-y^\delta\|_Y^2$, and was stopped according to the discrepancy principle after $k=9$ iterations ($\tau=2.7, \rho=0.8$). Further application of the algorithm leads to a discrete jump in the data and real space profile misfit. While the data misfit is then recovered (bottom right graph), the profile misfit does not improve any more (bottom left graph). Note that the fit converges to the correct solution, although the *a priori* interval $[\alpha,\beta]$ of the profile (gray shaded area in Fig. 2, top left graph) has not been extended to the end of the maximum gradient $(\delta\rho)_0$ in the profile.

In a real experiment, the maximum density gradient $(\delta\rho)_0$ is often known with limited accuracy only, i.e., usually it is only known that $|(\delta\rho)_0|$ is smaller than the total gradient $|\Delta\rho|$ between substrate and bulk phase. In fact, $|\Delta\rho|$ can be determined very easily even at laboratory sources by measuring the critical angle q_c and utilizing the equality $q_c=2\sqrt{\pi\Delta\rho_e r_0}$. To address the problem of the usually unknown maximum gradient $(\delta\rho)_0$, the stability of the reconstructions with respect to changes in $(\delta\rho)_0$ was examined. All conditions for reconstruction were the same as before, except a change of the density gradient $(\delta\rho)_0$ between the oxide layer on top of the silicon substrate and the water layer usually present between the substrate and a lipid double layer [5]. In addition, the stopping parameter τ was adjusted to prevent overfitting. As shown in Fig. 3 (top graph) the quality of the fits in Fourier space is not substantially decreased by a change in the gradient $|(\delta\rho)_0|$. More importantly, the real space reconstructions do not substantially decrease in quality as well (bottom graph). For clarity, the plots of all reconstructed profiles corresponding to the data and profile misfit values shown in Fig. 3 are given as supplemental online material [34]. This suggests that the algorithm leads to very accurate solutions under realistic and experimentally relevant conditions.

Using the same parameters (except a value of $\tau=2.5$ and an increased length 33 Å of the interval $[\alpha,\beta]$) the algorithm was also applied to very noisy data, typically obtained with a laboratory source (simulated $I_0=7\times10^6$). The simulated data were generated from the same model parameters as used above, except the lower total intensity. The results are shown in Fig. 4. Considering the reduced amount of data ($q_z \leq 0.4$ Å^{-1}) and the high noise level, the quality of the reconstructed profile is still quite good. Even the positions of the maxima and minima and thus also the bilayer thickness are reproduced very well.

We emphasize that a stopping rule is an essential part of the algorithm since the reconstructions deteriorate after a cer-

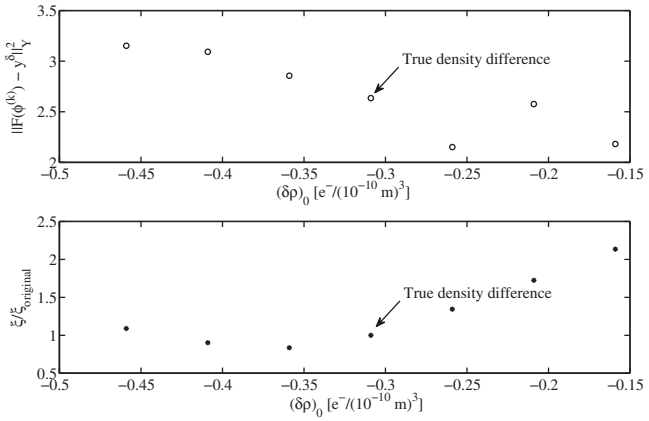

FIG. 3. (Top) Maximum gradient change $(\delta\rho)_0$ in initial profile versus misfit of data and reconstructed reflectivities obtained after stopping of the algorithm by the discrepancy principle. The same simulated data as presented in Fig. 2 were used. The data point corresponding to the true gradient in the initial profile is marked by an arrow and corresponds to the fit result presented in Fig. 2. (Bottom) Gradient change $(\delta\rho)_0$ in initial profile versus misfit of the reconstructed real space profiles corresponding to the reflectivity curves evaluated in the top graph. For comparison, the points are normalized to the misfit ξ_{original} resulting from reconstruction with the true gradient $(\delta\rho)_0$ in the initial profile.

tain number of iterations. (The best fit in q space does often, but not necessarily correspond to the most accurate profile reconstruction.) In Figs. 4 and 2 the discrepancy principle always selected close-to-optimal stopping index as expected from theory [29]. The value of the stopping parameter τ must be chosen >2, otherwise the Newton iteration may never terminate or the data may be overfitted leading to oscillations in the reconstructed profile as shown in Fig. 2 (top left).

Two other important parameters are the support length a and the sampling width a/n of the reconstructed profile in real space. We found that it is advisable to choose a so that $|a-A|/A<0.2$, where A is the true support, in order to yield good results. As discussed in Sec. III n should be chosen so that $aq_{\max}/n<1$. In our experiments $aq_{\max}/n\approx 0.25$ was sufficient, the resulting profiles are quite stable against a further increase of n. However, increasing n might increase the accuracy of the profile reconstruction.

The reconstructions with simulated data discussed so far may give an idea of the minimum quality of the reflectivity data and the necessary *a priori* knowledge of the density profile for the regularization technique to yield reliable and good results. In the following, we will compare the algorithm to a published inversion routine named MOTOFIT [30], which is based on a matrix method, i.e., a dynamic description of reflectivity. Beforehand, a remark about the different existing profile modeling approaches used for reflectivity inversion may be in order. At least two different techniques are presently used: While in the first approach a limited number of boxes or a repeating box pattern is applied to describe physically distinguished layers [5], in a second approach, the transition to a very large number of boxes is made, i.e., the transition to a sampling description of the electron density profile [31,32]. Each approach can be applied within the framework of either the full dynamic or the kinematic theory. As the scheme presented here uses a large number of sampling points to describe the density profile, it clearly belongs to the latter so-called "free-form methods."

The results of the comparison of the regularization scheme and the genetic algorithm are shown in Fig. 5. The data used for reconstruction here was simulated on the basis of the same original model profile as used before, i.e., a lipid bilayer on a solid support. In performing the comparison of the different methods, we aimed at greatest possible objectivity: To this end, a density model was generated in MOTOFIT with the same support length as used for the reconstruction by regularization ($a=100$ Å). The support interval was divided into 18 layers of fixed widths to keep the number of free parameters low, but yet to achieve a free-form description of the density rather than a traditional box-model description. This would need much more *a priori* information to converge due to the reduced variability in profiles, which can be described by the model. The density profiles used as an initial guess were the same in both reconstruction schemes. As visible in Fig. 5 (top left), the reconstruction by the genetic algorithm differs substantially from the true

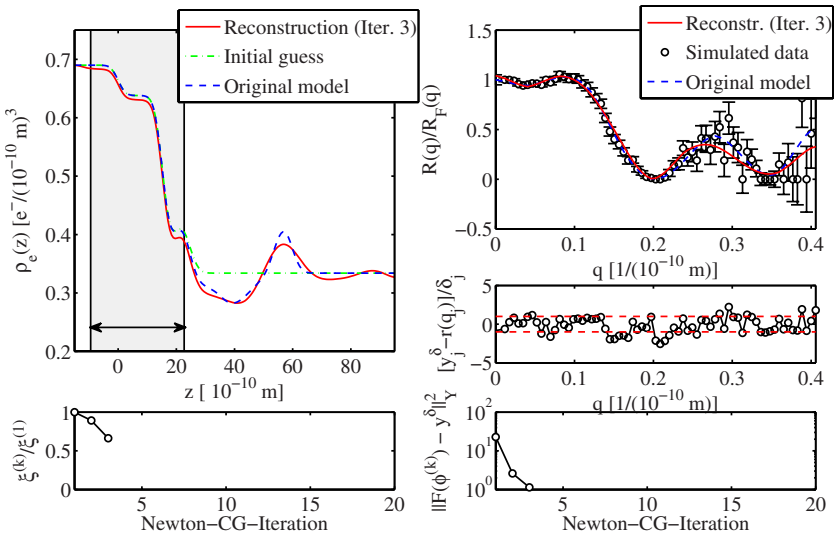

FIG. 4. (Color online) Results from a reconstruction of simulated data assuming a laboratory source beam intensity ($I_0=7\times10^6$ photons). For nomenclature refer to Fig. 2.

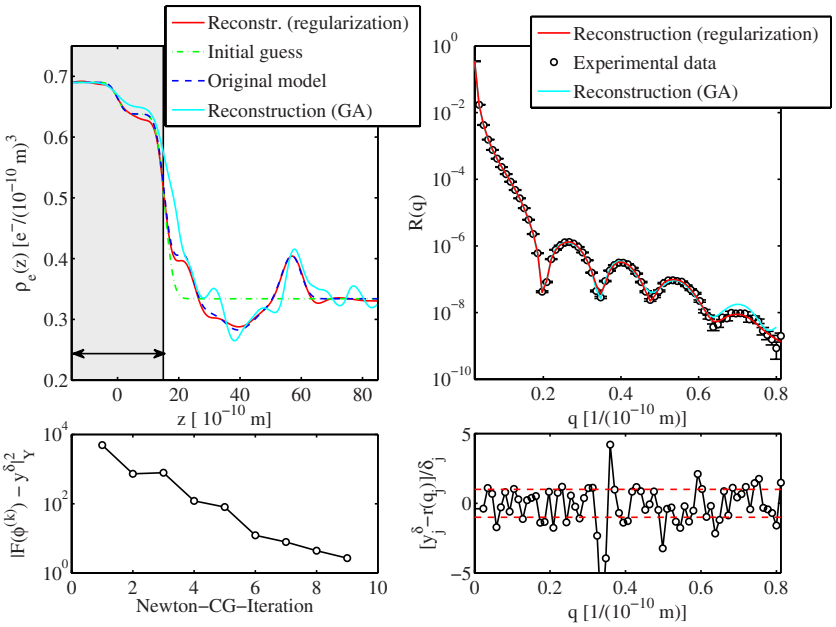

FIG. 5. (Color online) Comparison of the regularization method with a reconstruction scheme based on a genetic algorithm (MOTOFIT). (Top left) Several density profiles are shown, more importantly (i) the reconstructed density profile resulting from the regularization scheme with abort due to the discrepancy principle (solid line, red), and (ii) reconstruction by application of a genetic algorithm in MOTOFIT (solid line, cyan). In addition, the profile used as the start profile for both reconstruction schemes is shown ("initial guess," dashed-dotted line, green). Finally, the profile used for simulation of the data with added noise, i.e., the "true" solution is shown ("original model," dashed line, blue). (Top right) Reflectivity curves—not normalized to Fresnel reflectivity here—corresponding to profiles on the left, i.e., simulated data with added noise ("experimental data"), reconstructed reflectivity by regularization (solid line, red) and by application of the genetic algorithm (solid line, cyan). (Bottom left) Data misfit versus iteration number for the reconstruction by regularization. (Bottom right) Residuals of reconstructed reflectivity curve by regularization with respect to the theoretical reflectivity curve without added noise.

profile, while the regularization scheme succeeds to provide a very accurate representation of the true model here. It is emphasized that the genetic algorithm—due to its nondeterministic nature—may lead to better profiles by evaluating even the same start parameters many times. Furthermore the reconstruction might improve by increasing the number of sampling points in the profile description, which however substantially increases computation time.

Summarizing, it may be concluded that the presented reconstruction scheme can be advantageous over or complement existing free-form approaches in some respects and situations: Results are produced within seconds, so that adjustments in the initial parameters can be performed very effectively. Furthermore, the amount of necessary *a priori* information is very low in many cases, e.g., no lower and upper limits for fit parameters have to be provided as it was also the case for the genetic-algorithm example presented here.

To give a last example, the algorithm was applied to experimental data obtained from reflectivity measurements of supported di-palmitoyl-phosphatidyl-choline (DPPC) monolayers on a hydrophobized silicon support [7]. The data were collected at the undulator beamline ID1 at the European Synchrotron Radiation Facility in Grenoble, France. During the preparation procedure the silicon wafer was first silanized by coverage with a covalently bound hydrocarbon layer of octadecyl-tri-chloro-silane (OTS). Second, a monolayer of DPPC was added by vesicle fusion. The data have then been analyzed by fitting the parameters of a box-model (3) with six boxes. In the nomenclature described before, this is a "conventional" box model, where usually each box can be identified with a part in the experimental system. Here, this result only serves as a reference for the reconstruction based on the regularization technique. The reference fit presented here was generated by successive runs of a genetic, global optimization technique and a deterministic, local optimization scheme, with many manual adjustments of the fit results, including a considerable degree of physical knowledge into the reconstruction.

The reconstruction shown in Fig. 6 was generated from an initial guess with a minimum amount of *a priori* knowledge about the profile. Just the density difference $\Delta\rho$ of the substrate and the bulk phase as well as an estimated representation of the native oxide layer on the silicon substrate were included into the initial profile. Obviously, the main features of the profile (DPPC headgroup region, hydrocarbon core region) are reproduced in the reconstruction very well. There are only minor differences between the profiles generated by regularization and with the help of the box model, especially in the hydrocarbon region of the OTS layer. Given the poor initial guess and the complicated and time-consuming procedure that led to the box model fit, the performance of the regularization technique in this case is remarkably good. In addition, it is pointed out that the true profile cannot be known here with absolute accuracy as the data have been collected experimentally, so that the box model reference also has some uncertainty.

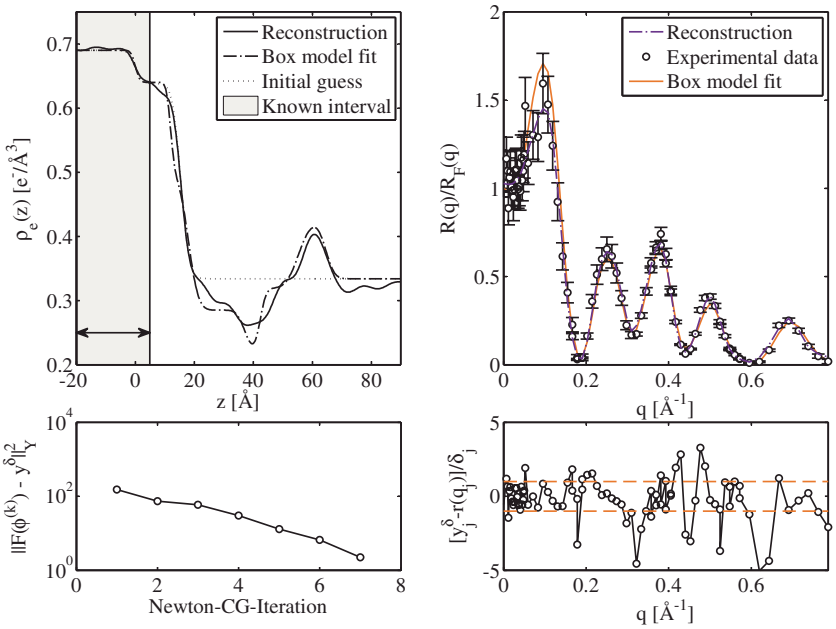

FIG. 6. (Color online) (Top left) reconstructed profile for a DPPC monolayer on OTS-silicon support (solid line). For comparison, a box-model fit to the same data is shown (dashed-dotted line), as well as the initial guess for the reconstruction (dotted line). (Top right) experimental data with errors and reflectivity simulations corresponding to the profiles shown on the left. (Bottom left) Misfit $\|F(\phi^{(k)}) - y^\delta\|^2_Y$. (Bottom right) Residuals for the reconstructed profile.

VI. DISCUSSION

We have presented an iterative method for the reconstruction of a refractive index profile from phaseless reflectivity measurements, which is more flexible in handling incomplete data and valuable *a priori* information on the profile and the covariance structure of the noise than reconstruction methods based on explicit inversion formulas [19,21,22]. In particular, bounded and nonequidistant grids of q values, varying uncertainties of the measured values, and *a priori* information on parts of the profile can be incorporated in the algorithm in a natural way. The convergence is very fast, usually a reconstruction takes only a few seconds on a standard desktop PC.

As a main disadvantage of our method we mention that it requires a sufficiently good initial guess. However, the physical knowledge about the sample is usually good enough to give a more than sufficient initial guess. In our numerical experiments the incorporation of *a priori* information on the behavior of the solution in a small interval was always sufficient.

As opposed to widely used fitting techniques, our algorithm can in principle reconstruct arbitrary profiles and does not use the assumption that the unknown profile can be described by a small number of parameters in a box model. Moreover, the speed of convergence is much faster than for global optimization techniques. The necessary condition that at least part of the profile is known at the level of a good guess, is given in most practical cases, where thin layers are deposited or adsorb to a generally well-known substrate.

ACKNOWLEDGMENTS

We would like to thank Professor Rainer Kress for helpful discussions. The support of the DFG/SFB 755 Nanoscale Photonic Imaging is gratefully acknowledged. Moreover, the ESRF is acknowledged for beamtime, Peter Boesecke and Hartmut Metzger from ID1 are thanked for their support during the beamtime.

[1] J. Als-Nielsen and D. McMorrow, *Elements of Modern X-Ray Physics* (Wiley, Chichester, 2001).
[2] M. Tolan, *X-ray Scattering from Soft-Matter Thin Films*, *Springer Tracts in Modern Physics*, Vol. 149 (Springer, New York, 1999).
[3] T. P. Russell, Mater. Sci. Rep. **5**, 171 (1990).
[4] A. Braslau, P. S. Pershan, G. Swislow, B. M. Ocko, and J. Als-Nielsen, Phys. Rev. A **38**, 2457 (1988).
[5] C. E. Miller, J. Majewski, T. Gog, and T. L. Kuhl, Phys. Rev. Lett. **94**, 238104 (2005).
[6] E. Novakova, K. Giewekemeyer, and T. Salditt, Phys. Rev. E **74**, 051911 (2006).
[7] K. Giewekemeyer and T. Salditt, Europhys. Lett. **79**, 18003 (2007).
[8] L. G. Parratt, Phys. Rev. **95**, 359 (1954).
[9] I. M. Tidswell, B. M. Ocko, P. S. Pershan, S. R. Wasserman, G. M. Whitesides, and J. D. Axe, Phys. Rev. B **41**, 1111 (1990).
[10] A. Braslau, M. Deutsch, P. S. Pershan, A. H. Weiss, J. Als-Nielsen, and J. Bohr, Phys. Rev. Lett. **54**, 114 (1985).
[11] V. Bargmann, Rev. Mod. Phys. **21**, 488 (1949).
[12] M. Braun, S. A. Sofianos, and R. Lipperheide, Inverse Probl. **11**, L1 (1995).
[13] K. Chadan and P. Sabatier, *Inverse Problems in Quantum Scattering Theory* (Springer, Berlin, 1989).
[14] Z. S. Agranovich and V. A. Marchenko, *The Inverse Problem of Scattering Theory* (Gordon and Breach, New York, 1963).
[15] C. F. Majkrzak and N. F. Berk, Phys. Rev. B **58**, 15416 (1998).
[16] M. V. Klibanov, P. E. Sacks, and A. V. Tikhonravov, Inverse

Probl. **11**, 1 (1995).
[17] E. Wolf, Proc. Phys. Soc. London **80**, 1269 (1962).
[18] F. N. C. P. V. Petrashen, Sov. Phys. Dokl. **34**, 1269 (1989).
[19] W. L. Clinton, Phys. Rev. B **48**, 1 (1993).
[20] M. V. Klibanov and P. E. Sacks, J. Math. Phys. **33**, 3813 (1992).
[21] M. V. Klibanov and P. E. Sacks, J. Comput. Phys. **112**, 273 (1994).
[22] G. Reiss and R. Lipperheide, Phys. Rev. B **53**, 8157 (1996).
[23] K.-M. Zimmermann, M. Tolan, R. Weber, J. Stettner, A. K. Doerr, and W. Press, Phys. Rev. B **62**, 10377 (2000).
[24] J. Strzalka, E. DiMasi, I. Kuzmenko, T. Gog, and J. K. Blasie, Phys. Rev. E **70**, 051603 (2004).
[25] M. V. Klibanov, Adv. Differ. Equ. **22**, 1232 (1986).
[26] B. Blaschke, H. W. Engl, W. Grever, and M. V. Klibanov, Nonlinear World **3**, 771 (1996).
[27] H. W. Engl, M. Hanke, and A. Neubauer, *Regularization of Inverse Problems* (Kluwer Academic Publisher, Dordrecht, 1996).
[28] J. P. Kaipio and E. Somersalo, *Statistical and Computational Inverse Problems* (Springer, New York, 2004).
[29] M. Hanke, Numer. Funct. Anal. Optim. **18**, 971 (1997).
[30] A. Nelson, J. Appl. Crystallogr. **39**, 273 (2006).
[31] E. Politsch and G. Cevc, J. Appl. Crystallogr. **35**, 347 (2002).
[32] C. F. Laub and T. L. Kuhl, J. Chem. Phys. **125**, 244702 (2006).
[33] MATLAB is a registered trademark of The MathWorks, Inc.
[34] See EPAPS Document No. E-PLEEE8-77-120805 for open source MATLAB code of the reconstruction algorithm, and plots of all reconstructed profiles. The MATLAB code is free to be used and modified, if the source is cited. For more information on EPAPS, see http://www.aip.org/pubservs/epaps.html
[35] It is assumed that this number of photons is collected within a reasonable period of time, i.e., 1 s.