electronic reprint

Journal of
Applied Crystallography
ISSN 0021-8898

Determination of the regularization parameter in indirect-transform methods using perceptual criteria

D. I. Svergun

J. Appl. Cryst. (1992). **25**, 495–503

J. Appl. Cryst. (1992). **25**, 495–503

Determination of the Regularization Parameter in Indirect-Transform Methods Using Perceptual Criteria

BY D. I. SVERGUN*†

GKSS Research Center, GKSS-WS, 2054 Geesthacht, Germany

(*Received* 1 *July* 1991; *accepted* 4 *February* 1992)

Abstract

A method is proposed for the determination of the optimum value of the regularization parameter (Lagrange multiplier) when applying indirect transform techniques in small-angle scattering data analysis. The method is based on perceptual criteria of what is the best solution. A set of simple criteria is used to construct a total estimate describing the quality of the solution. Maximization of the total estimate is straightforward. Model computations show the effectiveness of the technique. The method is implemented in the program *GNOM* [Svergun, Semenyuk & Feigin (1988). *Acta Cryst.* A**44**, 244–250].

1. Introduction

Indirect methods are widely used in small-angle scattering (SAS) data treatment. Using a few *a priori* assumptions about the object being investigated, they often allow reliable results to be obtained, even with poor experimental data.

Indirect approaches in SAS are based on the assumption that the experimentally measured set of intensities $\mathbf{J} = J(s_i)$, $i = 1, \ldots, N$, can be related by some integral transform $\mathbf{K}$ to a distribution function in real space $\mathbf{p} = p(r)$. Here, N is the number of experimental points, s is the modulus of the scattering vector $\mathbf{s}$ [$s = (4\pi/\lambda)\sin\theta$, λ is the wavelength, 2θ the scattering angle] and the distribution function is assumed to have a finite support, *i.e.* $p(r)$ differs from zero only in the interval $D_{\min} < r < D_{\max}$. This can be written in the form $\mathbf{J} = \mathbf{Kp}$ or

$$J(s) = \int_{D_{\min}}^{D_{\max}} K(s, r)p(r)\,\mathrm{d}r. \qquad (1)$$

The principle of the indirect approach, first proposed by Glatter (1977), is to solve (1) with respect to $p(r)$. Here, $p(r)$ can be a pair distribution function (monodisperse systems) or a size distribution (polydisperse systems); the operator $\mathbf{K}$ includes corresponding Fourier transform and smearing effects, if necessary.

The advantages in using indirect transforms have been extensively described in the literature (Glatter, 1982; Feigin & Svergun, 1987). The main difficulty is that the solution of (1) is an ill-posed problem, *i.e.* small errors in $J(s_i)$ may lead to large errors in $p(r)$. Several algorithms are known (Glatter, 1977; Moore, 1980; Taupin & Luzzati, 1982; Provencher, 1982; Svergun, Semenyuk & Feigin, 1988; Mangani, Puliti & Stefanon, 1988; Hansen & Pedersen, 1991). In most of them (1) is solved by using the regularization method (Tikhonov, 1943; Tikhonov & Arsenin, 1977), *i.e.* by minimizing the functional

$$T_\alpha[\mathbf{p}] = \|\mathbf{J} - \mathbf{Kp}\|_\mathbf{J}^2 + \alpha\Omega[\mathbf{p}]. \qquad (2)$$

Here, $\|\mathbf{J} - \mathbf{Kp}\|_\mathbf{J}$ denotes the function norm in reciprocal space which is normally taken in the form

$$\|\mathbf{J} - \mathbf{Kp}\|_\mathbf{J} = \left\{(N-1)^{-1}\sum_{i=1}^{N}\sigma_i^{-2} \times \left[J(s_i) - \int_{D_{\min}}^{D_{\max}} p(r)K(s_i, r)\,\mathrm{d}r\right]^2\right\}^{1/2}, \qquad (3)$$

σ_i being the standard deviation at the ith data point. $\Omega[\mathbf{p}]$ is a stabilizer containing *a priori* information about the solution. A frequently used stabilizer is, for example,

$$\Omega[\mathbf{p}] = \int_{D_{\min}}^{D_{\max}} \left[\frac{\mathrm{d}p(r)}{\mathrm{d}r}\right]^2 \mathrm{d}r, \qquad (4)$$

which requires $p(r)$ to be a smooth function. A non-negative regularization parameter α plays the role of the Lagrange multiplier.

Such an approach allows a stable solution of (1) to be obtained. The procedure is, however, not straightforward, because some auxiliary parameters have to be specified. If one supposes that the range $[D_{\min}, D_{\max}]$ and the stabilizer $\Omega[\mathbf{p}]$ are fixed, the value of α becomes the crucial parameter. Indeed, with $\alpha = 0$, one obtains the unstable solution produced by

* On leave from Institute of Crystallography, Academy of Sciences of Russia, Leninsky prospekt 59, 117333 Moscow, Russia.

† Present address: European Molecular Biology Laboratory, Hamburg Outstation, EMBL, c/o DESY, Notkestrasse 85, D-W2000 Hamburg 52, Germany.

0021-8898/92/040495-09$06.00

conventional least-squares methods. Too large values of α yield solutions that fulfil the *a priori* requirements in real space [*e.g.* $p(r)$ is very smooth], but correspond to strong deviations from the experimental data in reciprocal space. According to the regularization theory (Tikhonov & Arsenin, 1977), an optimal value of α should exist somewhere in between these extremes, provided the assumptions leading to (1) are valid.

The problem of choosing α has been considered by various authors. One way to estimate it, given by the χ^2 criterion (Bevington, 1969), is to select a value such that $\|\mathbf{J} - \mathbf{K}\mathbf{p}_\alpha\|_\mathbf{J} \simeq 1$, where $\mathbf{p}_\alpha = p_\alpha(r)$ is the solution corresponding to the given α. This, however, requires reliable estimates of the standard deviations, which are not always available in practice. Moreover, this criterion leads sometimes to oversmooth solutions (Hofmann, 1986, ch. 4.3.1), corresponding to α values that are too large. A second criterion is given by the point-of-inflection (Glatter, 1977) or quasi-optimality methods (Tikhonov & Arsenin, 1977; Hofmann, 1986, p. 96), which assume that the best solution is most stable with respect to changes in α. This criterion may also fail in practice especially as unacceptable solutions obtained as $\alpha \to 0$ and $\alpha \to \infty$ are also very 'stable'.

There is thus no universal recipe to choose the regularization parameter and such a recipe may not even exist in principle. Often the users of the indirect methods rely upon their expectations, that is they just visually compare results obtained for different α and select the most acceptable solution. This is, in principle, a reasonable although very tedious way of solving ill-posed problems. In this paper, this perceptual approach is formulated in mathematical terms and an algorithm to select the regularization parameter is proposed.

2. Perceptual criteria

We first mathematically formulate criteria that are often used to select the solution visually in real and reciprocal space. In the following, the notation $\|f\|$ is used for the norm of an arbitrary function $f(r)$ in real space,

$$\|f\| = \left[\int_{D_{\min}}^{D_{\max}} f^2(r)\,\mathrm{d}r\right]^{1/2}. \tag{5}$$

Oscillations. It is generally assumed in visual interpretation that the best distribution function $p_\alpha(r)$ is smooth. A measure of smoothness is provided by the ratio $\|p'_\alpha\|/\|p_\alpha\|$, where $p'(r) = \mathrm{d}p(r)/\mathrm{d}r$ denotes the derivative of the function. This estimate can be put on an absolute scale by comparison with the same ratio for a smooth reference function. For example, a sine function $f(r) = \sin(\pi n r/\Delta D)$, $\Delta D = D_{\max} - D_{\min}$, with n an integer, gives $\|f'\|/\|f\| = n\pi/\Delta D$. Therefore, if the criterion

$$\mathrm{OSCILL} = (\|p'_\alpha\|/\|p_\alpha\|)/(\pi/\Delta D) \tag{6}$$

is close to 1, $p_\alpha(r)$ is a smooth monomodal function (*e.g.* the pair distribution function of a sphere gives OSCILL = 1.1), OSCILL $\simeq 2$ corresponds to either a smooth bimodal or an oscillating monomodal distribution, and so on.

Systematic deviations. The most important criterion in reciprocal space is related to the systematic deviations of the restored function $\mathbf{J}_\alpha = \mathbf{K}\mathbf{p}_\alpha$ from the experimental data set. This can be estimated by analyzing the residuals $\Delta_i = J(s_i) - J_\alpha(s_i)$, for example, by counting the number N_s of successive residuals Δ_i that change sign (Hamming, 1971). In the absence of systematic deviations, the probability of two successive differences having the same sign is 0.5, whereas this probability increases if there are systematic deviations. N_s is thus randomly distributed around half the total number of points $N/2$ with a dispersion $N^{1/2}$ in the absence of systematic deviations and it decreases when such deviations are present. In the absence of systematic deviations in reciprocal space, the value of the criterion

$$\mathrm{SYSDEV} = N_s/(N/2) \tag{7}$$

must be close to 1.

Discrepancy. This criterion measures whether the residuals obtained correspond to the experimental errors. It can be taken in the standard form of the so-called generalized discrepancy (Tikhonov & Arsenin, 1977, ch. 2)

$$\mathrm{DISCRP} = [\|\mathbf{J} - \mathbf{K}\mathbf{p}_\alpha\|_\mathbf{J}^2 - \mu^2(\mathbf{K}, \mathbf{p})]^{1/2}. \tag{8}$$

Here, $\mu^2(\mathbf{K}, \mathbf{p}) = \inf_{\alpha>0} \|\mathbf{J} - \mathbf{K}\mathbf{p}_\alpha\|_\mathbf{J}^2$ is the measure of inconsistency of the problem, *i.e.* the χ^2 value of the best obtainable fit to the experimental data points. The value of $\mu^2(\mathbf{K}, \mathbf{p})$ should be equal to zero when the $p(r)$ function is described by a number of parameters N_p equal to the number of experimental points N and increases with increasing ration N/N_p. In practice, $\mu^2(\mathbf{K}, \mathbf{p})$ is never zero due to the rounding errors. According to the χ^2 criterion (see *Introduction*), DISCRP for the optimum solution should be somewhat lower than 1. For practical application, values in the range $0.7 < \mathrm{DISCRP} < 0.95$ are recommended (see, for example, Hofmann, 1986, p. 129).

Stability. This criterion, describing how much the solution changes with changing α, can be formulated in relative terms as

$$\mathrm{STABIL} = (\|p_\alpha - p_{\alpha+\delta\alpha}\|/\|p_\alpha\|)/(\delta\alpha/\alpha). \tag{9}$$

As discussed in the *Introduction*, according to the point-of-inflection and the quasi-optimality methods, a value of STABIL $\ll 1$ can be expected in the vicinity of the correct solution.

Positivity. In many cases, it is known that $p(r)$ should be a non-negative function. The natural measure of non-negativity of $p_\alpha(r)$ is

$$\text{POSITV} = \|p_\alpha^+\|/\|p_\alpha\|, \qquad (10a)$$

where $p^+(r) = p(r)$ if $p(r) > 0$, otherwise $p^+(r) = 0$. For a non-negative function, POSITV = 1.

Validity in the central part of $p(r)$. Normally, when the range $[D_{\min}, D_{\max}]$ is correctly specified (*i.e.* corresponds to the actual range of sizes in the given system), most information is contained in the central part of the distribution $p(r)$, which corresponds to the largest values of $p(r)$. These values must not only be the largest but should also be reasonable from the physical point of view. To avoid unstable solutions oscillating around zero, the criterion that measures the validity of the central part of $p(r)$ can be formulated as

$$\text{VALCEN} = \max\{\|p_\alpha^*\|\}/\|p_\alpha\|, \qquad (10b)$$

where $p^*(r) = p(r)$ if $D_{\min} + \Delta D/4 < r < D_{\max} - \Delta D/4$, otherwise $p^*(r) = 0$ (here $\Delta D = D_{\max} - D_{\min}$). The maximum value is taken from the integrals along the portions of $p(r)$ with a constant sign. For example, for the $p(r)$ of a solid sphere, VALCEN = 0.95.

The characterization of the solution in terms of the criteria just introduced is best achieved with a few examples. Fig. 1 illustrates some possible types of real-space distribution functions together with the corresponding values of OSCILL, POSITV and VALCEN. The value of $D_{\min}$ is taken as zero for simplicity. Among the functions shown, the curves (*a*), (*c*), (*e*) and (*g*) are the ones that a trained observer would most probably select as plausible solutions. Curve (*a*) could correspond to a globular particle, (*c*) to an anisometric particle, (*e*) is a typical profile of a bilayer and (*g*) is a bimodal size-distribution function. Using these examples, one can estimate the plausible range of the criteria (namely, OSCILL < 3, VALCEN > 0.5; the meaning of POSITV is clear). Curve (*b*), which might correspond to an understabilized solution, would probably not be selected. It has a

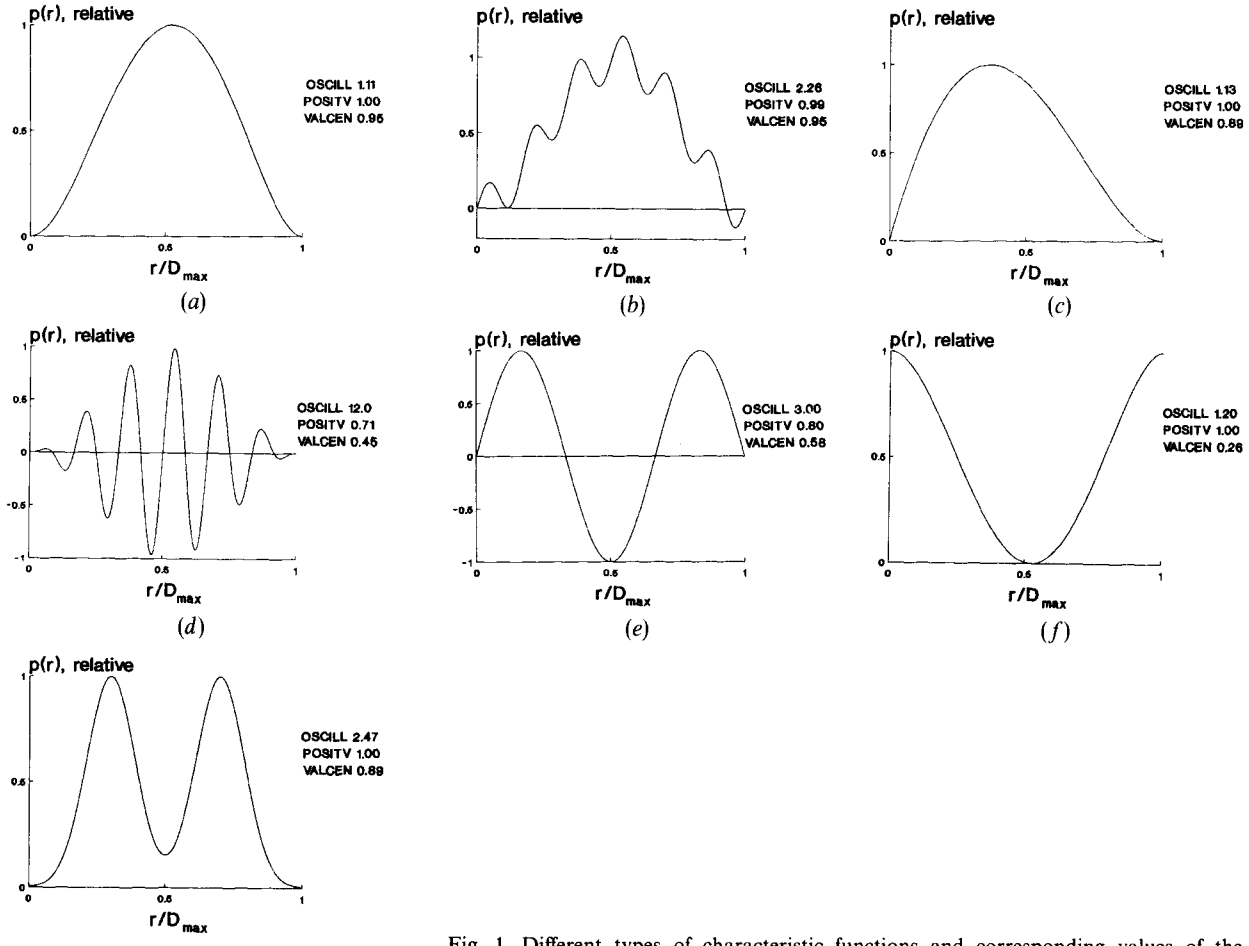

Fig. 1. Different types of characteristic functions and corresponding values of the perceptual criteria.

value of OSCILL that is too large for a monomodal function. Curves (*d*) and (*f*) are definitely to be rejected. The former represents an unstable solution, characterized by high OSCILL and small VALCEN. Curve (*f*), which could be obtained when the range $[D_{min}, D_{max}]$ is incorrect, has an unacceptable value of VALCEN.

In Fig. 2, the reciprocal-space criteria are illustrated using generated model data. Plot (*a*) shows the 'normal' case: DISCRP and SYSDEV are close to 1. The points in this pattern were obtained from the smooth curve by a random-number generator with a relative error of 5%. Plot (*b*) simulates an understabilized solution: DISCRP is somewhat smaller than 1. Plots (*c*) and (*d*) illustrate the result of incorrect estimates of the standard deviations. Plot (*c*) is identical to plot (*a*), but the error bars are halved, therefore DISCRP is too large. In plot (*d*), a constant is added to the smooth curve and the error bars are doubled. Although an acceptable DISCRP is obtained, SYSDEV indicates that this fit is incorrect.

The parameter STABIL describes the solution dynamics resulting from the use of the regularization procedure. To get its estimate on the absolute scale, the following example can be given: STABIL = 0.01 means that, when doubling α, the solution $p_\alpha(r)$ changes on average by $\exp(0.01 \ln 2) = 1.007$ times, *i.e.* there is a difference of 0.7%. Thus, if STABIL $\simeq 10^{-2}$, the solution can be considered to be stable.

The behavior of these criteria with changing α is demonstrated by example 1, a smeared scattering curve from a solid sphere with statistical noise (not shown here), using the method of Svergun, Semenyuk & Feigin (1988). Fig. 3 illustrates dependence of the six criteria on α, which correspond well to the qualitative expectations. The left arrow indicates the optimum value of α providing the best solution, the right arrow indicates the theoretical maximum value of α: $\alpha_{max} \simeq \|\mathbf{K}\|_J^2/\|\mathbf{J}\|_J$ (Tikhonov & Arsenin, 1977, ch. 2). For higher α values, the fit to the experimental data is poor and therefore DISCRP $\gg 1$, SYSDEV $\ll 1$. In the lower α range [unstable solution: $p(r)$ oscillating around zero], OSCILL $\gg 1$ and POSITV $\simeq 0.5$. VALCEN is also small for the low α values, shows a maximum in the vicinity of the correct solution and then decreases again, because the $p_\alpha(r)$ function degrades to a constant for very large α. STABIL shows a minimum in the range of plausible α, being, however, small for very low and very high α as well. The dependencies illustrated here are typical. A similar behavior has been observed for numerous theoretical and experimental examples.

3. Total estimate

The perceptual criteria reflect different properties of the solution and they are fulfilled in the different ranges of α. In general, acceptable values of the reciprocal-space criteria DISCRP and SYSDEV require lower α values in order to fit the data. However, good values OSCILL and POSITV require larger values of α to avoid oscillating $p_\alpha(r)$ functions. In addition, in the vicinity of the correct solution,

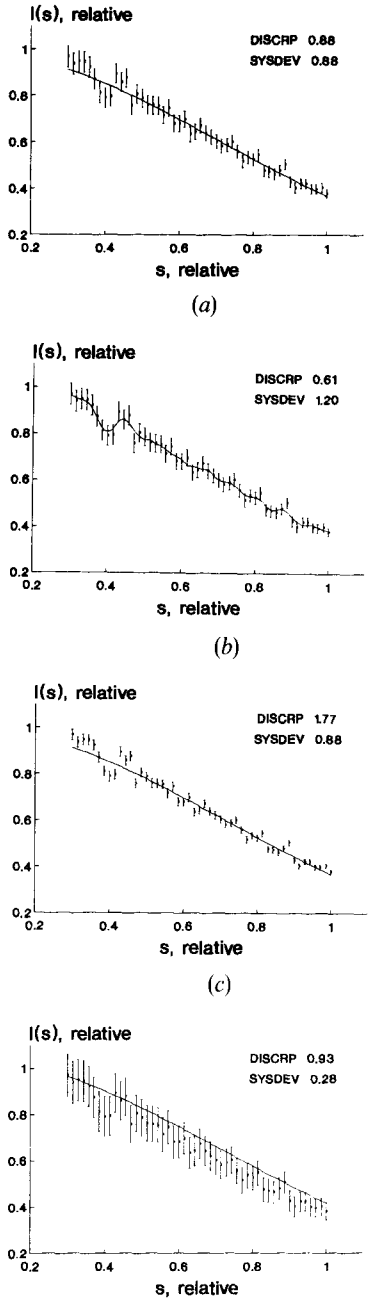

Fig. 2. Types of fit in reciprocal space. Smooth curves correspond to $J_\alpha(s)$, points with error bars correspond to $J(s_i)$. For details, see text.

VALCEN is likely to have a maximum, whereas STABIL should be small. A combination of these parameters can then be used to construct an estimate that summarizes the impressions of an expert and describes the quality of the solution.

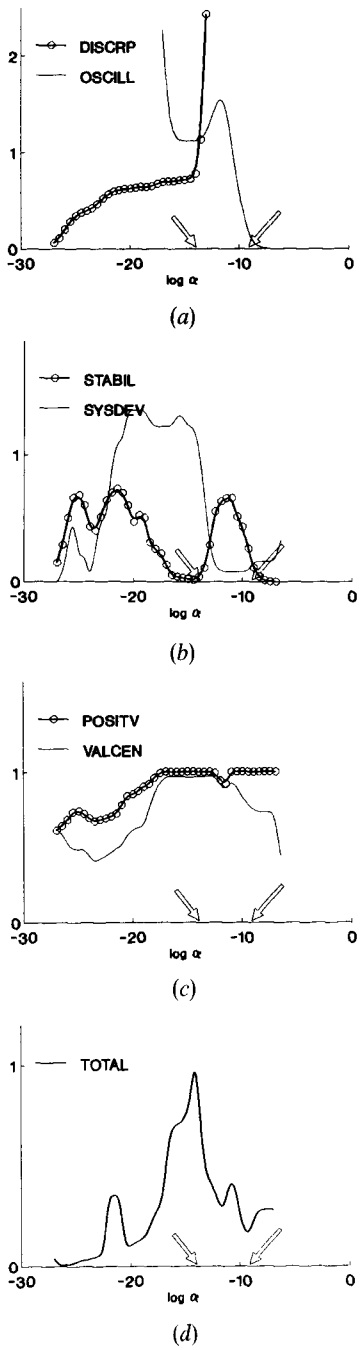

Fig. 3. Dependence of the perceptual criteria and the total estimate on the regularization parameter α for example 1. In each plot, the left arrow indicates the value of α corresponding to the best solution and the right arrow the theoretical maximum value of α.

To relate the numerical values of the criteria to a common scale, 'probability' estimates

$$P_j = \exp\{-[(A_j - B_j)/C_j]^2\}, \quad j = 1, \ldots, 6 \quad (11)$$

can be introduced, where A_j is an expected ('ideal') value of the jth criterion. B_j its actual value and C_j specifies the width of the distribution. Using this representation, all the estimates are between 0 and 1, the constants C_j being 'scale factors' for the criteria presented in Fig. 3.

A compromise between these probabilities can be found by constructing a total estimate in the form

$$\text{TOTAL} = \sum_j W_j P_j \Big/ \sum_j W_j, \quad (12)$$

where W_j describes the weight (significance) of the jth estimate. The total estimate also ranges between 0 (totally unacceptable) and 1 (ideal solution).

Numerous test calculations have been carried out with various sets of parameters A_j, C_j and W_j. The set presented in Table 1 provides the optimum balance between the criteria. Here, the optimum values of OSCILL, POSITV and VALCEN are those of the solid sphere. The allowed deviations are largest for DISCRP and OSCILL because the values of these criteria may vary significantly with α. OSCILL, STABIL and SYSDEV are taken as the most important criteria. This choice will be discussed in more detail in the next section.

The dependence of TOTAL corresponding to this parameter set is also presented in Fig. 3. It displays a pronounced maximum allowing easy determination of the optimal α. The characteristics of the solution are given in Table 1.

4. Model examples

Example 1 is trivial and the solution need not be shown. Since the real-space perception criteria are based on a solid sphere, the corresponding scattering curve should be handled perfectly. The resulting TOTAL = 0.99 could not be obtained in a practical case.

The question arises naturally as to what would happen if the characteristic function differed significantly from the assumed monomodal distribution. The present approach works well for such cases even without changing the parameters of the total estimate. If one assumes, for example, that the $p(r)$ function is actually trimodal, the resulting solution $p_\alpha(r)$ could only have OSCILL $\simeq 1$ for large α values, for which SYSDEV, DISCRP and VALCEN differ significantly from the expected values, leading to a low total estimate. On the other hand, in the vicinity of the correct solution, where OSCILL $\simeq 3$, the corresponding probability is nearly zero and thus does not

Table 1. *Set of default parameters and the resulting estimates for the model examples*

Parameter	DISCRP	OSCILL	STABIL	SYSDEV	POSITV	VALCEN	
W_j	1.00	3.00	3.00	3.00	1.00	1.00	
C_j	0.30	0.60	0.12	0.12	0.12	0.12	
A_j	0.70	1.10	0.00	1.00	1.00	0.95	
Example no.							TOTAL
1	0.66	1.08	0.02	1.10	1.00	0.92	0.99
2	0.81	5.82	0.11	1.06	0.83	0.55	0.44
3	6.02	1.10	0.20	0.53	1.00	0.94	0.43
4	0.43	3.06	0.05	1.14	0.99	0.35	0.58
5	1.19	3.87	0.03	0.94	0.91	0.79	0.55

influence the position of the maximum of the total estimate due to the other parameters. This maximum may be less pronounced and the value of TOTAL will certainly be smaller than for a monomodal distribution indicating that the solution is somewhat different from the expected one.

These qualitative considerations are illustrated below by a series of model calculations. The method has been implemented for automatic determination of the regularization parameter in the program package *GNOM* (Svergun, Semenyuk & Feigin, 1988; Svergun, 1991; Semenyuk & Svergun, 1991). Starting with the theoretical value α_{max}, the program performs a conventional golden-section search (Press, Flannery, Teukolsky & Vetterling, 1989) to maximize the total estimate. In all the following examples, the set of default parameters shown in Table 1 was used and the solutions presented here correspond in each case to the maximum total estimate found automatically. The corresponding values of the perception criteria and the total estimates are summarized in Table 1.

Examples 2 and 3 (see Fig. 4) use a simulated scattering curve from a densely packed polydisperse system of solid spheres (generated by J. Skov Pedersen). The number distribution function of spheres, $N(R)$, which specifies the number of particles of radius R, is described by a Gaussian centered at $R = 30$ Å with the full width at half-maximum of 20 Å (average radius of gyration $R_g = 31.8$ Å). The volume fraction of spheres is 0.2. The intensity is calculated using the Percus–Yevick approximation (van Beurten & Vrij, 1981) and 5% statistical noise is added. The interference effects are very strong and the correlation radius is approximately 250 Å. When treating the sytem as monodisperse with $D_{max} =$ 250 Å (example 2), a mixture of particle scattering and of the interference function is obtained. In this case, smooth and positive $p_\alpha(r)$ functions are simply impossible. Therefore, the total estimate, relying mainly on SYSDEV and DISCRP, provides the solution shown in Fig. 4 curves (*a*) and (*b*). The result is totally different if the same data are treated as a polydisperse system of spheres with $D_{min} = 0$, $D_{max} = 50$ Å, as illustrated in example 3 [Fig. 4 curves (*a*) and (*b*)]. In this case, it is not possible to find a reasonable solution that would fit the data and, therefore, DISCRP and SYSDEV are sacrificed in favor of the real-space criteria. As a result, the fit in reciprocal space is poor, whereas the obtained volume distribution function, $p(r) \equiv V(R) \sim R^3N(R)$, provides a reasonable average radius of gyration, $R_g = 24.8$ Å. Better restoration of the initial distribution function is obviously impossible due to the extremely large interference effects. It must be noted that both examples describe very special cases; this is reflected in low values of the maximum total estimates.

Another example (no. 4) was taken from the paper of May & Nowotny (1989) and presents the application of the method to a multidomain structure. The scattering intensity in Fig. 5(*a*) corresponds to the two-domain particle consisting of eight spheres, shown in Fig. 5(*b*). The corresponding calculated data contain 6% statistical noise. In this case, a monomodal function cannot fit the data, therefore, OSCILL plays no role. The quality of the resulting solution is fairly good compared to the results of other indirect methods (Hansen & Pedersen, 1991, example 3).

The application of the method to real experimental data (example 5) is illustrated using the scattering pattern of an aqueous solution of phospholipids (Somjen, Coleman, Koch, Wachtel, Billington, Towns-Andrews & Gilat, 1991). The experimental data shown in Fig. 6(*a*) were collected on the X33 camera (Koch & Bordas, 1983) of the EMBL at DESY in Hamburg. Since standard deviations were not known, a constant relative error of 5% was assumed. The sample contained 3.9% (weight percent) of egg lecithin, 2.2% sodium taurocholate and 93.9% aqueous solution of 0.05 *M* TRIS–HCl, 0.15 *M* NaCl and 0.0015 *M* EDTA. Under these conditions, disc-like phospholipid bilayers are formed. The resulting profile of the $p(r)$ function illustrated in Fig. 6(*b*) is typical for a bilayer.

5. Discussion

The examples illustrate that the method selects a plausible solution using the set of default parameters for different systems. Of course, any *a priori* information can be used to improve the reliability of the procedure. If, for example, no reliable error

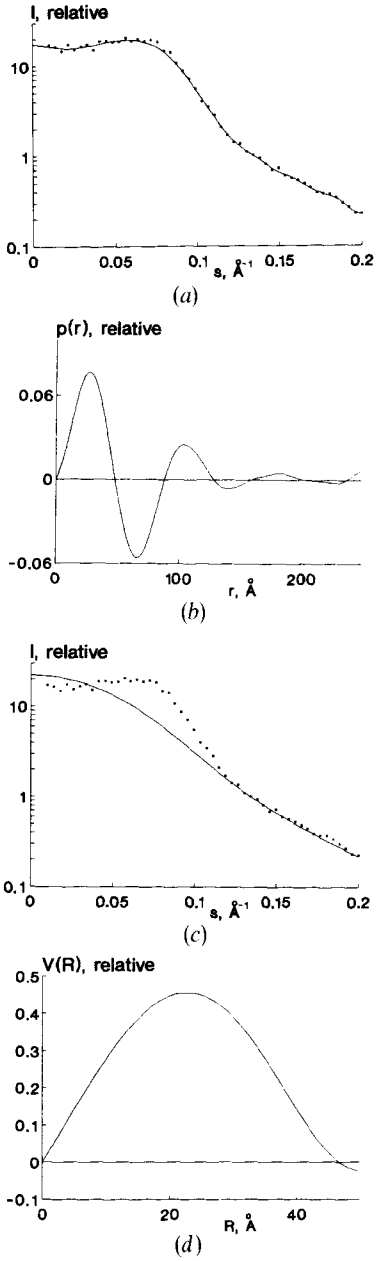

Fig. 4. Treatment of the simulated scattering from a densely packed polydisperse system of spheres. Example 2, curves (*a*) and (*b*): treated as a monodisperse system with $D_{max} = 250$ Å; example 3, curves (*c*) and (*d*): treated as a polydisperse system with $D_{min} = 0$, $D_{max} = 50$ Å. Dots indicate the simulated data points containing statistical errors, solid lines indicate the restored intensities and characteristic functions.

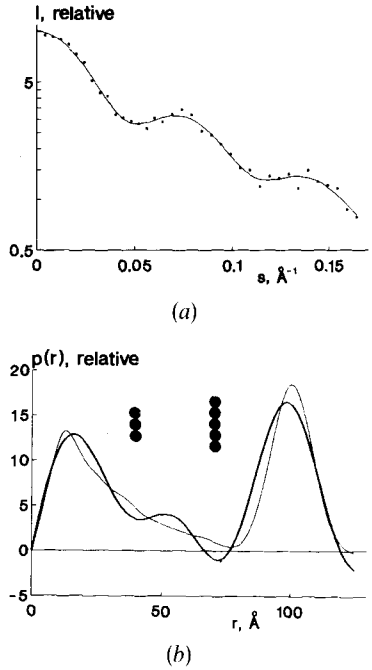

Fig. 5. Treatment of the two-domain particle scattering (example 4). Dots indicate the simulated data points containing statistical errors, solid lines indicate (*a*) restored intensities and (*b*) characteristic functions. The bold curve in (*b*) corresponds to the theoretical $p(r)$ function of the model, which is schematically shown as eight full circles.

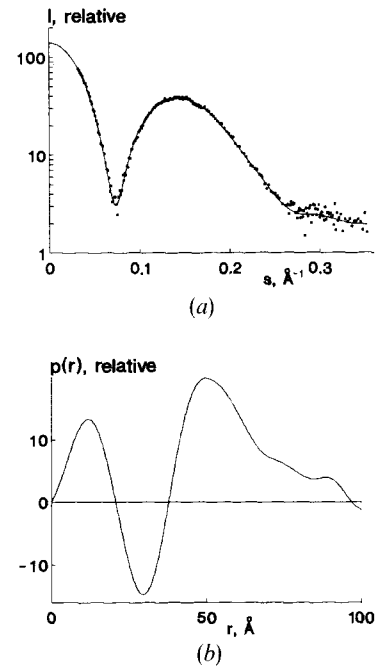

Fig. 6. Scattering from phospholipid bilayers (example 5). Dots: experimental data points; solid curves: (*a*) restored intensity and (*b*) pair distribution function.

estimates are available, it is better to switch off DISCRP by setting its weight equal to 0. If one expects a bilayer structure, the ideal values can be set for OSCILL $\simeq 3$ and POSITV $\simeq 0.6$. The program allows the modified set to be saved and used afterwards for similar objects.

The relationship between DISCRP and SYSDEV deserves attention. In principle, when the standard deviations are known, DISCRP can be more sensitive than SYSDEV. However, incorrect values of the standard deviations may result in incorrect estimates (*cf.* Fig. 2). Moreover, the ideal value of DISCRP is not well defined and may vary between 0.5 and 0.9. Therefore, SYSDEV has more significance in the default set, but this can easily be changed by the user.

Examples 1, 4 and 5 illustrate that in the normal cases the correct solution is really stable (criterion STABIL is low). In some special cases, *e.g.* examples 2 and 3, the stability is poor and, although the point-of-inflection or quasi-optimality methods may fail, the perceptual criteria are still able to provide the solution.

Other criteria could also be included in the total estimate. For example, the Fisher information (see, for example, Frieden, 1989)

$$I[p] = \int_{-\infty}^{\infty} \{[p'(r)]^2/p(r)\}\, \mathrm{d}r = \min \qquad (13)$$

may be used to describe the smoothness of $p(r)$. The difference from OSCILL is that in (13) one demands more smoothness from the less informative portions of $p(r)$ [*i.e.* where $p(r)$ is small]. This is theoretically reasonable, but in practice, since these parts make small contributions to the scattering intensity, $I[p]$ may become a hypersensitive criterion.

If a prior estimate of $p(r)$ is available, one can also use the maximum-entropy criterion (Skilling, 1989)

$$S[p, m] = \sum_k [-p_k \ln (p_k/m_k) + p_k - m_k] = \max, \qquad (14)$$

where $m(r)$ is the prior distribution. Unfortunately, with an incorrect (or neutral, for example, a constant) prior estimate, this criterion only degrades the solution. One should also note that both (13) and (14) formally require positivity of $p(r)$. Since they are not dimensionless, additional scaling procedures are required. In the current version of the program, criteria (13) and (14) are not implemented.

The present approach is not claimed to be a universal recipe. Cases may exist where the default set does not work and *a priori* information is to be taken into account for reliable estimation of the α value. In such cases, however, a low value of the maximum total estimate (say TOTAL < 0.5) indicates the poor quality of the solution in terms of the initial expectations. Normally, when dealing with single-domain particles or a monomodal or bimodal distribution, values of TOTAL $\simeq 0.7$–0.8 can be obtained using the set of default parameters.

Concluding remarks

The proposed method based on perceptual criteria allows the optimum value of the regularization parameter to be determined automatically and also allows a quantitative estimate of the quality of the solution to be obtained. Numerous calculations with model data as well as applications to real objects illustrate that the set of default parameters for the total estimate ensures a straightforward solution for various kinds of objects. This set can, if necessary, be altered to take into account the user's experience and/or *a priori* information. A similar approach could also be applied to the methods based on Shannon's theorem (Moore, 1980; Taupin & Luzzati, 1982), where the maximum particle diameter D_{max} becomes a variable.

The author would like to thank Dr M. H. J. Koch for valuable comments on the manuscript and Dr J. Skov Pedersen and Dr R. May for providing the test examples. He is also grateful for the support he received from the GKSS Research Center.

References

Beurten, P. van & Vrij, A. (1981). *J. Chem. Phys.* **74**, 2744–2748.
Bevington, P. R. (1969). *Data Reduction and Error Analysis for the Physical Sciences*, p. 85. New York: McGraw-Hill.
Feigin, L. A. & Svergun, D. I. (1987). *Structure Analysis by Small-Angle X-ray and Neutron Scattering*, pp. 310–319. New York: Plenum Press.
Frieden, B. R. (1989). *Am. J. Phys.* **57**, 1004–1008.
Glatter, O. (1977). *J. Appl. Cryst.* **10**, 415–421.
Glatter, O. (1982). In *Small-Angle X-ray Scattering*, edited by O. Glatter & O. Kratky, pp. 119–165. London: Academic Press.
Hamming, R. W. (1971). *Introduction to Applied Numerical Analysis*, ch. 11.6. New York: McGraw-Hill.
Hansen, S. & Pedersen, J. S. (1991). *J. Appl. Cryst.* **24**, 541–548.
Hofmann, B. (1986). *Regularization for Applied Inverse and Ill-Posed Problems*. Leipzig: BSB Teubner.
Koch, M. H. J. & Bordas, J. (1983). *Nucl. Instrum. Methods*, **208**, 461–469.
Mangani, M., Puliti, P. & Stefanon, M. (1988). *Nucl. Instrum. Methods*, A**271**, 611–616.
May, R. P. & Nowotny, V. (1989). *J. Appl. Cryst.* **22**, 231–237.
Moore, P. B. (1980). *J. Appl. Cryst.* **13**, 168–175.
Press, W. H., Flannery, B. P., Teukolsky, S. A. & Vetterling, W. T. (1989). *Numerical Recipes*, ch. 10.1. Cambridge Univ. Press.
Provencher, S. W. (1982). *Comput. Phys. Commun.* **27**, 213–227, 229–242.

SEMENYUK, A. V. & SVERGUN, D. I. (1991). *J. Appl. Cryst.* **24**, 537–540.
SKILLING, J. (1989). In *Maximum Entropy and Bayesian Methods*, edited by J. SKILLING, pp. 42–52. Dordrecht: Kluwer Academic Publishers.
SOMJEN, G. J., COLEMAN, R., KOCH, M. H. J., WACHTEL, E., BILLINGTON, D., TOWNS-ANDREWS, E. & GILAT, T. (1991). *FEBS Lett.* **289**, 163–166.
SVERGUN, D. I. (1991). *J. Appl. Cryst.* **24**, 485–492.
SVERGUN, D. I., SEMENYUK, A. V. & FEIGIN, L. A. (1988). *Acta Cryst.* A**44**, 244–250.
TAUPIN, D. & LUZZATI, V. (1982). *J. Appl. Cryst.* **15**, 289–300.
TIKHONOV, A. N. (1943). *Dokl. Acad. Nauk SSSR*, **39**, 5, 195–198. (In Russian.)
TIKHONOV, A. N. & ARSENIN, V. YA. (1977). *Solution of Ill-Posed Problems.* New York: Wiley.