Information content of lineshapes

R. N. Silver, D. S. Sivia, and R. Pynn
Theoretical Division and Los Alamos Neutron Scattering Center
Los Alamos National Laboratory
Los Alamos, New Mexico 87545
U.S.A.

ABSTRACT: We examine the question of figures-of-merit for optimizing the lineshapes of neutron scattering sources, instruments and experiments. Using maximum entropy deconvolution of simulated data, we test the effects of various features of lineshapes including intensity, resolution (FWHM), shape, and background. We demonstrate that conventional figures-of-merit are of limited validity, and we suggest that bandwidth is an important criterion for optimization.

Introduction

An outstanding problem in the development of neutron scattering sources, instruments and experiments is how to establish figures-of-merit (FOM) for optimizing designs. This is central to the wise allocation of the billions of dollars which have been, or are proposed to be, invested in neutron scattering facilities for condensed matter research [1]. Applications [2] would include comparing the relative performances of pulsed spallation and reactor neutron sources, choosing the poisoning of moderators at pulsed sources, making decisions to trade intensity for resolution or noise for signal in the conduct of an experiment, etc. Prior approaches to optimization have been primarily intuitive "seat-of-the-pants" judgments based on empirical experience in neutron scattering research. Although the goal should be to maximize the information gained in neutron scattering experiments, until now there has been little effort to address the problem from an information theory viewpoint. Optimization should consider the statistical problem of data analysis. In the present paper, we begin to remedy this oversight by providing some simulations of the ability to recover information from differing instrument response functions (or lineshapes). We wish to motivate the application of information theory to neutron scattering science and facility design.

Let us consider a one-dimensional neutron scattering experiment in which the data, $D(x)$, are a convolution of an instrument response function (or resolution function), $R(x)$, with the neutron scattering law, $S(x)$. Then

$$D(x) = \int_{-\infty}^{\infty} R(x-x')\, S(x')\, dx' + B(x) + \Sigma(x) \quad , \qquad (1)$$

where $B(x)$ is the background, and $\Sigma(x)$ is the noise. The goal of any neutron scattering experiment is to infer the neutron scattering law, $S(x)$, from the data, $D(x)$.

In a representative neutron scattering experiment, the neutron scattering law, $S(x)$, we wish to measure may be as shown in Fig. 1(a). As a first approximation $R(x)$ is usually assumed to be Gaussian, as the Central Limit Theorem suggests. Traditionally, an experimenter would select the Full Width Half Maximum (FWHM) of $R(x)$ to be of the same magnitude as the width of the structure expected in $S(x)$. The corresponding typical data set, $D(x)$, is shown in Fig. 1(b), where there is broadening due to the instrument (of FWHM = 2.5 pixels as shown by the bar), noise governed by Poisson statistics, and background (chosen to be flat and equal to 5% of the peak signal). That is, experiments are typically optimized so that the raw data, $D(x)$, resembles a slightly broadened and noisy version of the $S(x)$ neutron scatterers expect to measure. This "what-you-see-is-what-you-get" philosophy of optimization is based on the widely-held perception that, if the FWHM of $R(x)$ is very broad, much sharper structure in $S(x)$ is not recoverable.

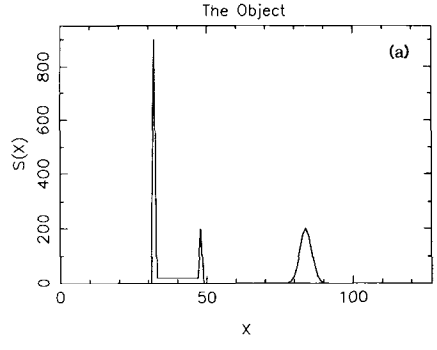


FIG. 1(a) – A representative neutron scattering law, $S(x)$, as a function of pixel (or channel) number, x.

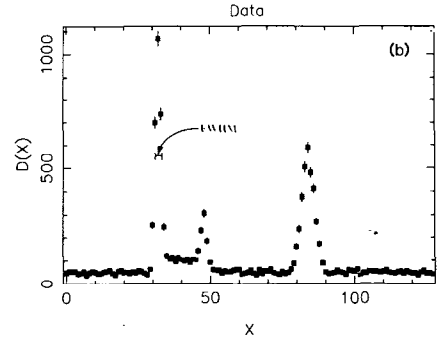


FIG. 1(b) – Typical neutron scattering data, $D(x)$, corresponding to the neutron scattering law in Fig. 1(a). The data are broadened by a Gaussian with FWHM of 2.5 pixels as shown, and with noise and 5% background added.

While neutron scattering is in general a Poisson process, in the limit of large numbers we can make an independent Gaussian approximation for $\Sigma(x)$ in which

$$< \Sigma(x) > \; = \; 0 \quad , \qquad (2)$$

and

$$< \Sigma(x) \; \Sigma(x') > \; = \; \delta(x - x') \left[\int_{-\infty}^{\infty} R(x - x') \; S(x') \; dx' + B(x) \right] \quad . \qquad (3)$$

Here < > denotes an average over all such experiments. A common data analysis procedure would be to estimate the parameters of a model for $S(x)$ by minimizing

$$\chi^2 = \frac{1}{N}\sum_i \Delta x \, \frac{\left(D(x_i) - \sum_j \Delta x \; R(x_i - x_j) \; S(x_j) - B(x_i)\right)^2}{D(x_i)} \quad , \qquad (4)$$

where we have broken the integral up into pixels of width Δx, and N is the number of pixels in the range of the experiment. The errors on the parameters would be determined from the variation of χ^2.

Another popular perception is that the FOM for the design of spectrometers should be

$$F.O.M. \; = \; \frac{Total\ Intensity}{FWHM^2} \quad . \qquad (5)$$

Instruments with the same FOM are supposed to have comparable performance for neutron scattering experiments. This FOM is rigorously correct for the problem of determining the position of a δ–function broadened by a Gaussian $R(x)$ by minimizing χ^2. It has been proposed as a more general FOM for the problem of optimizing spectrometers for pulsed neutron sources [3].

We contend that both these popular perceptions are demonstrably false. A counterexample is provided by the Be Filter Difference Spectrometer at LANSCE, where features in $S(x)$ orders of magnitude sharper than the FWHM have been recovered, using both direct inversion [4] and maximum entropy deconvolution [5], by taking advantage of the sharp leading edge of $R(x)$. The present paper generalizes this example to the overall problem of the optimization of neutron scattering experiments.

Approach

We formulate the problem of inferring $S(x)$ from $D(x)$ in terms of Bayes' theorem [6], upon which all data analysis procedures are at least implicitly based. This states that the conditional probability of $S(x)$ given $D(x)$ is

$$P[S(x)|D(x)] \; \propto \; P[D(x)|S(x)] \times P[S(x)] \quad . \qquad (6)$$

Here, $P[D(x)|S(x)]$ is the probability of the measured data for a given scattering law, which is referred to as the *Likelihood* function. In our limit of independent Gaussian statistics, this reduces to the familiar form

$$P[D(x)|S(x)] \Rightarrow \exp\left(-\frac{\chi^2}{2}\right) \quad . \qquad (7)$$

$P[S(x)]$ represents our state of knowledge about $S(x)$ (or the lack of it) before we have any data, and it is referred to as the *Prior*. Eq. (6) states that the product of the Likelihood and the Prior is proportional to the *Posterior*, or our state of knowledge after we have measured the data. The best estimate of $S(x)$ from the data is given by the maximum of the Posterior, and the errors in this estimate are given by its width. The data analysis procedure of parameter estimation by minimizing χ^2 (i.e. *maximum Likelihood*) is equivalent to maximizing the Posterior if the Prior is taken to be a uniform function of the parameters of the model. A less familiar procedure to neutron scatterers is deconvolution, in which an alternative form for the Prior is chosen. For example, in the maximum entropy method [7] the Prior is taken to be the exponential of the entropy of $S(x)$ relative to a starting default model.

The present paper describes simulations of the effect of differing instrument responses ($R(x)$, B, $\Sigma(x)$), on the ability to infer the scattering law, $S(x)$, from the data, $D(x)$. Altering the instrument response only alters the Likelihood function and not the Prior, so that our general conclusions will be independent of whether we attempt parameter estimation or deconvolution to infer $S(x)$. Similarly, since different deconvolution procedures only alter the Prior, our conclusions are also qualitatively independent of the choice of deconvolution method.

To be specific, we take the test $S(x)$ (termed the *object*) shown in Fig. 1(a), and we create simulated data resulting from various instrument responses. Although such simulations could be attempted for any physical experiment, in this paper the choice of instrument responses will be limited to what we consider to be typical of neutron scattering spectrometers and sources. Unless otherwise stated, we set the background level at 5% of the peak signal. We perform a maximum entropy (MaxEnt) deconvolution of the data to recover an inferred $S(x)$ (termed the *image*). Maximum entropy has been shown to be the preferred method for the deconvolution of positive additive distribution functions [7]. The MaxEnt deconvolution procedure will use a flat default model with a stopping criteria of $\chi^2=1$. We shall compare the image to the object, and we will exercise our subjective judgement about which images are more faithful to the object.

Our purpose is to be provocative rather than definitive. Therefore, we attempt to keep the argument simple by presuming that the instrument response parameters, $R(x)$ and $B(x)$, are accurately known, although this is often not true in practice. Data analysis procedures exist for the cases where these are imperfectly known including: *calibration experiments* on a known sample, $S(x)$, which can be used to infer $R(x)$; *blind deconvolution* which can be used to infer both $R(x)$ and $S(x)$; and, *two-channel deconvolution* which

can be used to infer $B(x)$. In principle a straightforward extension of our approach can include such complications, but we do not attempt that here.

Simulations

First, it is important to establish some experience with deconvolution. We begin with the ubiquitous Gaussian instrument response function, $R(x)$. Figure 2(a) shows typical simulated data when the object of Fig. 1(a) is convoluted with a broad Gaussian $R(x)$ of FWHM of 25 pixels, which is ten times the broadening of Fig. 1(b) and much broader than the structure in $S(x)$. Figure 2(b) shows the MaxEnt images corresponding to the data in Fig. 2(a) (dashed line), with 100 times the counts (solid line), and 10,000 times the counts (dotted line). Figure 2 demonstrates that deconvolution can resolve peaks which are not evident in the raw data, and that increasing statistical accuracy can improve the resolution of the image.

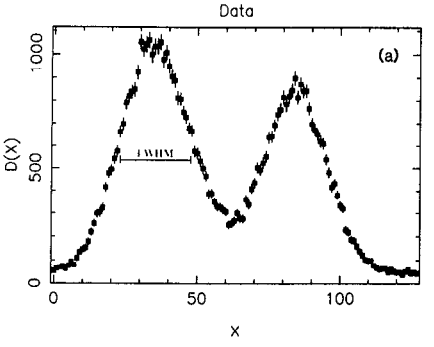


FIG. 2(a) – Gaussian broadened data with FWHM = 25 pixels, for the $S(x)$ of Fig. 1(a).

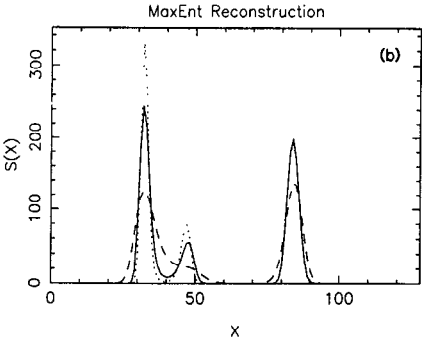


FIG. 2(b) – Images of $S(x)$ calculated by maximum entropy (MaxEnt) deconvolution of the data in Fig. 2(a) (dashed line), with 100 times the counts (solid line), and with 10,000 times the counts (dotted line).

Figure 3 shows images for the narrow Gaussian (FWHM of 2.5 pixels) broadening obtained by deconvoluting the data in Fig. 1(b) (solid line), and for the same experiment with 100 times the counts (dotted line). The sharp peaks on the left of the object are more highly resolved in the image with better statistical accuracy. The intrinsically broad peak on the right is unchanged. Figure 3 demonstrates that, even with a sharp $R(x)$, deconvolution can be useful because it can determine whether a peak is narrower than the instrument broadening.

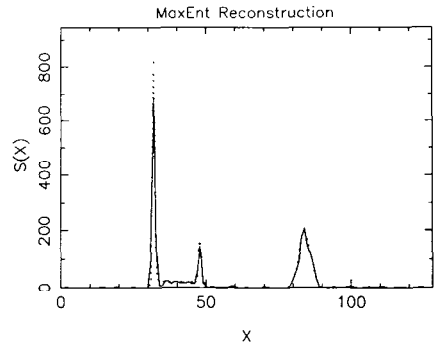


FIG. 3 – MaxEnt images of $S(x)$ for the narrow Gaussian broadened data in Fig. 1(a) (solid line), and with 100 times the counts (dashed line).

Moreover, Fig. 3 should be compared with Fig 2(b), which are the images created using data from the broad Gaussian $R(x)$. According to the popular FOM, Eq. (5), the solid line in Fig. 2(b)) should be comparable to Fig. 3, and the dotted line should be much sharper. In fact, even with 10,000 times the peak intensity, a ten times broader Gaussian $R(x)$ produces poorer images than the narrow Gaussian. This provides our first counterexample to this FOM.

Next we consider an $R(x)$ constructed by convoluting the narrow Gaussian (of FWHM = 2.5 pixels) with a wide exponential of $1/\tau$ = 1/15 pixels with identical peak intensity to Fig. 1(b). This lineshape is common to the *unpoisoned* moderators of pulsed neutron sources. This $R(x)$ is shown (solid line) in Fig. 4(a). Our initial expectation may be that the quality of the image would be severely degraded. The simulated data shown in Fig. 4(b) (pluses) do not resemble the object. Nevertheless, the quality of the image shown in Fig. 4(c) obtained by deconvolution (solid line) is almost equal to that obtained with the narrow Gaussian, shown as the solid line in Fig. 3 (note the change in vertical scale). We conclude that the correct FOM may not be very sensitive to the FWHM of $R(x)$, which also contradicts Eq. (5).

Figure 4 also shows a simulation analogous to the poisoning of moderators for pulsed neutron sources. Figure 4(a) shows a *poisoned* $R(x)$ (dashed line) constructed by convoluting a narrower exponential of $1/\tau$= 1/2 pixels with the same narrow Gaussian. Both the poisoned and unpoisoned $R(x)$ have the same peak intensities, although in practice there would be some decrease in peak intensity with poisoning. In the conventional neutron scatterer's view, poisoning moderators has the desirable effect of making the data much more closely resemble the object, as shown in Fig. 4(b) (dots), and therefore improving the apparent resolution. Fig. 4(c) shows the corresponding image (dashed line). Both the poisoned and unpoisoned moderators in fact have almost identical resolving power. Poisoning moderators is questionable from an information content viewpoint! Moreover, if the object being measured is broader than the FWHM of $R(x)$, neutrons are lost by poisoning.

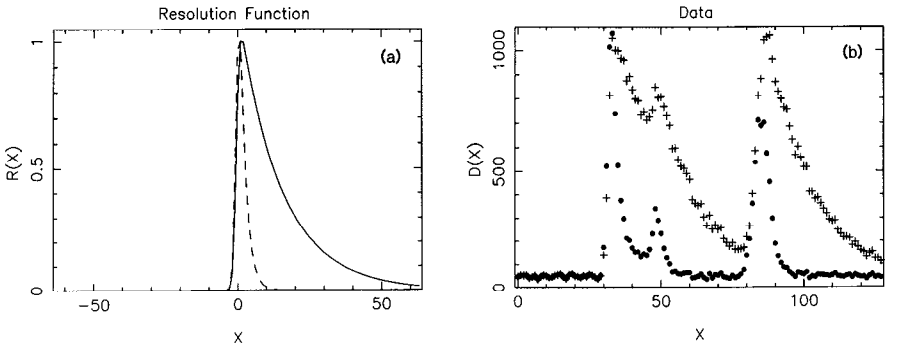
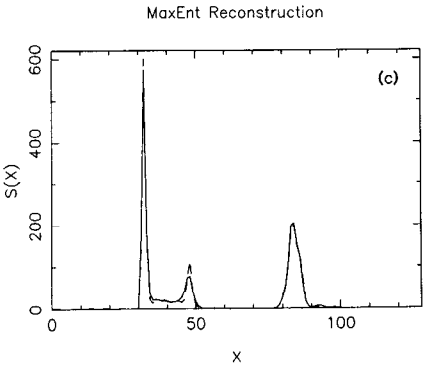


FIG. 4 – (a) Instrument response functions, $R(x)$, for a narrow Gaussian convoluted with a broad exponential (solid) and a narrow exponential (dashed); (b) corresponding data, $D(x)$, for broad (pluses) and narrow (dots) $R(x)$; (c) corresponding MaxEnt images of the neutron scattering law, $S(x)$, for broad (solid) and narrow (dashed) $R(x)$.



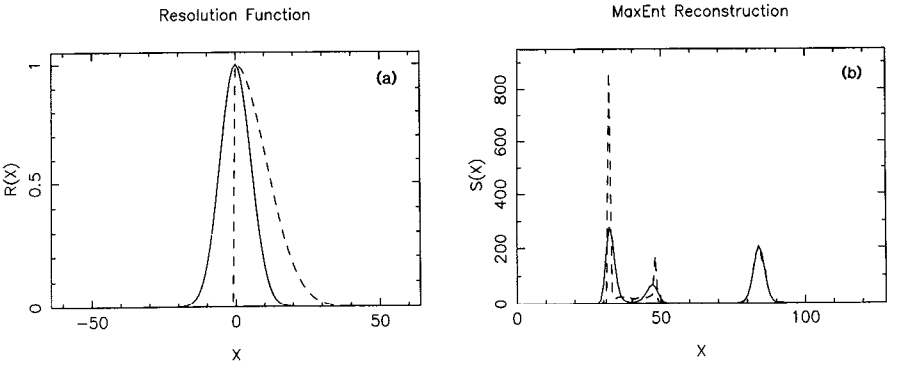


FIG. 5(a) – Two instrument response functions, $R(x)$, with the same figure-of-merit according to Eq. (5). The solid line is a Gaussian, and the dashed line is a half-Gaussian with the same FWHM.

FIG. 5(b) – MaxEnt images of $S(x)$ for the two $R(x)$ in Fig. 5(a). Solid line corresponds to the Gaussian $R(x)$, and the dashed line to the half-Gaussian $R(x)$.

The insensitivity of the FOM to the FWHM is further illustrated in Fig. 5. Figure 5(a) shows two different $R(x)$ with identical FWHM and total intensity, and therefore identical FOM according to Eq. (5). One is a Gaussian (solid line) and the other is a half Gaussian (dashed line). Figure 5(b) shows the corresponding MaxEnt images. It is clear that the half Gaussian $R(x)$ has much better resolving power.

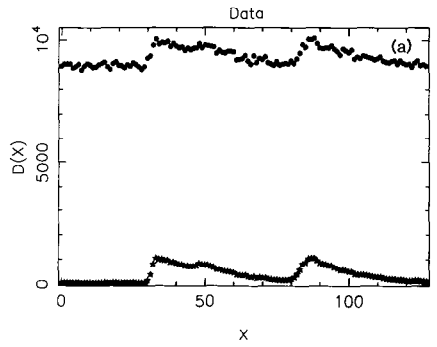


FIG. 6(a) – Study of the effect of background level, using the $R(x)$ in Fig. 4(a). The bottom data (stars) is with the usual 5% background. The top data (dots) has much higher background with the same signal.

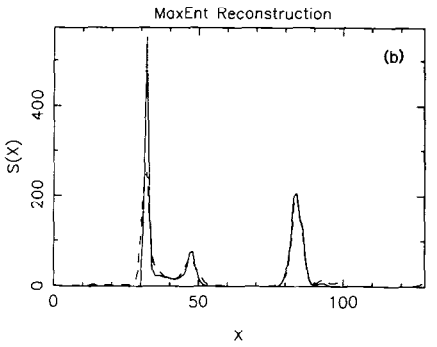


FIG. 6(b) – MaxEnt images of $S(x)$ for the two $D(x)$ in Fig. 6(a). Solid line corresponds to the low background data, and the dashed line to the high background data.

Finally, Fig. 6 illustrates that background is also important to the correct FOM for experiment optimization. We consider the same $R(x)$ as the solid line in Fig. 4(a), but with two different levels of background. The data are shown in Fig. 6(a) and the corresponding MaxEnt images are shown in Fig. 6(b). Higher background degrades the ability to recover information from the experiment, as expected.

Analysis

It should be clear from these simulations that the figure-of-merit is a much more complex object than suggested in Eq. (5). In particular, the sharpness of the structure in $R(x)$ appears to be far more important than the FWHM.

An argument to support this observation may be most easily developed by considering the direct inversion of the data by Fourier transform. We define the transform by

$$\overline{R}(k) \equiv \int_{-\infty}^{\infty} e^{ikx} R(x)\, dx \quad . \qquad (8)$$

Then the image, $S_I(x)$, formed by direct inversion is given by

$$S_I(x) = \int_{-\infty}^{\infty} \frac{dk}{2\pi} e^{-ikx} \left[\frac{\overline{D}(k) - \overline{B}(k)}{\overline{R}(k)} \right] \quad . \qquad (9)$$

Such an inversion would satisfy $\chi^2 = 0$. If we average over all such experiments in the same sense as Eq. (2) and (3), we find that the expectation value of the image is given by

$$< S_I(x) > \; = \; S(x) \qquad (10)$$

as desired. However, this inversion is poorly conditioned because of the noise term, $\overline{\Sigma}(x)$, in Eq. (1). Using Eq. (3), the variance of the image is given by

$$< \delta S_I(x) \; \delta S_I(x) > \; = \int \frac{dk}{2\pi} \int \frac{dk'}{2\pi} \frac{e^{-i(k+k')x}}{\overline{R}(k)\overline{R}(k')} \; \overline{D}_o(k+k') \quad , \qquad (11)$$

where

$$\overline{D}_o(k) = \overline{R}(k) \; \overline{S}(k) \; + \; \overline{B}(k) \quad . \qquad (12)$$

Since $\overline{R}(k)$ goes toward zero for k's larger than some critical value, call it k_c , direct inversion amplifies the noise and the variance is divergent. The image for any particular experiment would appear to be noisy.

The solution of this problem is to condition the inversion by asking instead for a broadened image. This can be crudely done by cutting off the limits on the integrals in Eq. (9) and Eq. (11) at some critical value, say k_c . The image for any particular experiment would have the noise suppressed at the expense of a broadened image; that is, the variance would be well behaved.

The k_c may be chosen such that $\chi^2 \approx 1$ for the image. It determines the achievable resolution by

$$\Delta x \approx \frac{2\pi}{k_c} \quad , \qquad (13)$$

Sharp features in $R(x)$ produce high Fourier components of $\overline{R}(k)$ which makes k_c large. A broad $R(x)$ lacks high Fourier components and so k_c must be small. Figure 7 shows the $\overline{R}(k)$ of the resolution functions shown in Fig. 5(a). The full Gaussian does not have high Fourier components at large k while the half Gaussian has large Fourier components at high k. The corresponding MaxEnt images are shown in Fig. 5(b). The resolution of an experiment primarily depends on the Fourier spectrum of the instrument response function.

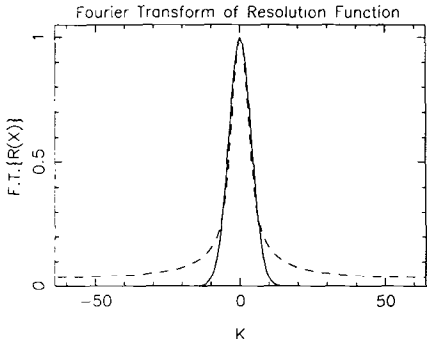


FIG. 7 – Fourier transforms of the instrument response functions, $R(x)$, for the Gaussian (solid line) and the half-Gaussian (dashed line) shown in Fig. 5(a).

Conclusions

We have provided counterexamples to several of the popular perceptions in the neutron scattering community regarding the relation between the instrument response function and the resolution of an experiment. We have shown that the figure-of-merit of an instrument response function strongly depends on its Fourier spectrum, in addition to other more traditional variables such as intensity, background, etc. Our Bayes' theorem argument suggests that this qualitative conclusion will remain valid regardless of whether deconvolution or parameter estimation is used to infer the neutron scattering law from the data. It will also remain true regardless of the specific choice of deconvolution procedure or fitting model.

We have not proposed a specific new figure-of-merit to replace Eq. (5). However, we suggest that the ultimate answer may have much in common with the theory of communication. The characteristic Fourier variable, k_c, which governs the resolution of an experiment, is analogous to the *bandwidth* of a signal processing circuit. In this sense, the design of neutron scattering experiments is related to the theory of communication [8], in which the capacity of a channel to transmit information is proportional to the bandwidth. Neutron scatterers should adapt the extensive knowledge and experience in information theory to the design of neutron scattering experiments and sources. A statistical theory for spectrometer optimization will be published elsewhere [9].

Acknowledgements

We thank R. Robinson for helpful discussions. This research was supported by the Office of Basic Energy Sciences of the U.S. Department of Energy.

References

[1] W. Seitz, D. Eastman, eds., *Major Facilities for Materials Research and Related Disciplines*, National Academy Press, Washington, D.C., 1984

[2] G. H. Lander, V. J. Emery, eds., *Scientific Opportunities with Advanced Facilities for Neutron Scattering*, Shelter Island Workshop, Octo-

ber, 1984, Argonne National Laboratory CONF-8410256; see also prior proceedings of the *International Collaboration on Advanced Neutron Sources* (ICANS).

[3] A. Michaudon, Reactor Science and Technology (Journal of Nuclear Energy Parts A/B), 17, 165–186 (1963); D. H. Day, R. N. Sinclair, Nuclear Inst. Meth. 72, 237–253 (1969); C. Windsor, *Pulsed Neutron Scattering*, Taylor & Francis, Ltd., London, 1981, p. 104ff

[4] P. Vorderwisch, F. Mezei, J. Eckert, J. Goldstone, Inst. Phys. Conf. 81, 161–168 (1986)

[5] D. Sivia, P. Vorderwisch, R. N. Silver, to be published; D. S. Sivia, this volume

[6] H. Jeffreys, *Theory of Probability*, Cambridge University Press, 1939

[7] S. F. Gull, J. Skilling, IEE Proc., 131F, 646 (1984); D. S. Sivia, ibid.

[8] C. E. Shannon, Bell Systems Tech. J., 27, 379–423, 623–656 (1948)

[9] R. N. Silver, D. S. Sivia, R. Pynn, to be published