Inverse Problems **16** (2000) 1561–1583. Printed in the UK PII: S0266-5611(00)12356-1

On relaxation-spectrum estimation for decades of data: accuracy and sampling-localization considerations

J Ross Macdonald

Department of Physics and Astronomy, University of North Carolina, Chapel Hill, NC 27599-3255, USA

E-mail: macd@email.unc.edu

Received 7 March 2000, in final form 2 August 2000

Abstract. In the inversion of data using the relaxation-spectrum model of electrical and rheological phenomena and processes, an important goal is to obtain a comprehensive and accurate approximation of the distribution of relaxation times (DRT) associated with the phenomenon or process under investigation. The estimation of such a DRT poses theoretical as well as experimental challenges. In the latter, it is often necessary to measure many decades of frequency-response data before sufficient information has been collected to allow an adequate estimate of the structure of the required spectrum to be calculated. At the theoretical level, among other things, one must solve the inverse problem of the relaxation-spectrum estimation, along with the frequent need to work with many decades of data. Some consequences of these two aspects of the problem are examined in this paper.

In order to address this problem, accurate (or noisy), wide-range frequency-response data sets derived from the Kohlrausch–Williams–Watts (KWW) continuous DRT function (with a value of its β_0 parameter of 0.5) are inverted numerically using a complex nonlinear least-squares method with variable τ free parameters and variable quadrature weighting, a method superior to Tikhonov regularization for data of the present type when inversion accuracy need not be sacrificed for increased resolution. The relative errors in the resulting DRT point estimates are investigated for various frequency ranges for both frequency-response data derived from the usual KWW DRT and from such a DRT which is abruptly cut off at its low τ end. Although the inversions are ill-posed, DRT errors were found to be very small over appreciable τ ranges, but the effects of cutoff of the range of the DRT and of a limited frequency range led to rapid relative-error increases at the ends of the DRT τ range, and allowed separation and quantification of the effects of these two limitations.

Important resolution differences are illustrated between the results of inversions of accurate and of noisy data by the present method and by Tikhonov regularization. Finally, the temporal response was calculated very simply from frequency-response inversion DRT estimates and compared with exact stretched-exponential response function values associated with the KWW dispersion model. Extremely small relative errors were found, but they nevertheless showed the effects of the above limitations at short and long times. Transformation to the time domain of the kind illustrated here for wide-range data is far simpler, more accurate, and more convenient than Fourier transformation.

1. Introduction

The estimation of a distribution of relaxation times (DRT) implicit in given frequency-response or temporal-response measurements is important for a wide variety of experimental situations because it can help elucidate the various physico-chemical processes occurring in the material studied. Here we shall consider inversion of data associated with small-signal, linear electrical response, immittance spectroscopy, although the methods and results considered apply equally

0266-5611/00/051561+23$30.00

to rheological situations. Immittance spectroscopy often requires one to work with data that extend over many decades of frequency (or time), from as low as 10^{-5} Hz to as high as 10^{12} Hz. The data obtained from most experiments often involve only 4–8 decades; however, the data points are usually distributed with constant, or approximately constant, spacing on a logarithmic scale. Consequently, if only 10–50 points are measured per decade, DRT estimation will only involve a few hundred data points or less.

Historically, when estimating relaxation spectra, it has been conventional wisdom that, for measured frequencies in the range $\omega_{\min} \leqslant \omega \leqslant \omega_{\max}$, the associated relaxation spectrum is completely determined within the temporal range defined by $\tau_{\min} \leqslant \tau \leqslant \tau_{\max}$, where

$$\tau_{\min} = \exp(\pi A/2)/\omega_{\max}, \tag{1}$$

and

$$\tau_{\max} = \exp(-\pi A/2)/\omega_{\min}, \tag{2}$$

with $A = 0$, e.g. [1–3]. This choice seems reasonable because it defines the range $(\omega_{\max})^{-1} \leqslant \tau \leqslant (\omega_{\min})^{-1}$, one which uses the full ω data set. This is a plausible assumption when those $\tau_{\min}$ and $\tau_{\max}$ values are unknown which define the range over which the actual DRT is significant, the usual situation for the estimation of an initially unknown DRT.

Recently, however, it has been established by Davies and Anderssen [4, 5] that $A = 1$ is a more conservative choice. Then, given $\omega_{\min}$ and $\omega_{\max}$, the relaxation-time range over which the DRT is well defined by the data is predicted to be shorter than the above $A = 0$ estimate by a factor of $\exp(\pi)$, about 1.364 decades. A principal purpose of the present work is to compare these two different estimates and to arrive at explicit conclusions about their relevance and consequent limitations on the accuracy of relaxation-spectrum recovery. Such conclusions are quantified below by accurate inversion of various data situations.

Note that to the degree that the $A = 1$ restrictions apply, they imply that if the actual DRT extends beyond the $A = 1$ range limits, it cannot be accurately estimated there from the specified frequency data. An important advantage of the $A = 0$ and 1 quantitative restrictions is that they show what needs to be measured if one wishes to recover a particular piece of information by a specific algorithm, or alternatively they allow one to work out what information can be recovered from measurements which have already been performed. Because of the importance of these results, it is worthwhile to verify them in practice. Therefore, an object of this paper is to investigate the applicability of the present limitations by numerical analysis of essentially exact frequency-domain data associated with a continuous DRT. In doing so, one must use as accurate an inversion method as possible in order to obtain meaningful estimates of DRT points and their errors. Thus, an allied aim of this work is to evaluate limiting inversion accuracy as well as possible, for both noiseless and noisy data.

When the appropriate range of a DRT is initially unknown, what frequency range is appropriate for use in determining the DRT? One might start with a range which one hopes will include the frequency of the peak of the magnitude of the imaginary part of the data. When that is found, one finds the frequencies where the magnitude has dropped to about 1% of the peak on either side of it. These two frequencies are then reasonable initial choices for $\omega_{\min}$ and $\omega_{\max}$. For the present data model, this procedure leads to a total range of about seven decades, although five or six decades is often sufficient for a good DRT determination from accurate data, and useful results are demonstrated below for a range of less than four decades.

A potential problem arises when one wishes to apply conventional inversion techniques, such as the Tikhonov regularization inversion approach [6, 7] or the moving-average approach of Anderssen and Davies [5, 8], to obtain estimated points of the associated relaxation spectrum

which may also involve a range, τ_{max}/τ_{min}, of many decades. These techniques often employ equal arithmetic-scale spacing of M τ points, resulting in the need for 10^6 or more points for wide-range data, an unrealistic and often impractical requirement. There are, however, other inversion approaches available for wide-range data that avoid the above many-point difficulty by using variable point spacing in the τ domain [9–12]. A detailed comparison of the utility of the present parametric inversion approach (described in the appendix) with Tikhonov regularization inversion appears in [12] and demonstrates the superiority of the parametric approach for noisy wide-range immittance data. The present results show, however, that for limited-range noisy data, regularization is useful.

Most inversion procedures for wide-range data involve M logarithmically distributed data points with equal intervals on a logarithmic scale. For noisy data involving a continuous DRT, however, one generally finds that the larger M, and so the higher the expected resolution, the poorer the accuracy of the resulting DRT point estimates obtained with the present method. Even with exact synthetic data, however, for which much larger useful values of M are possible, M_{max} remains finite because of the limited computational accuracy inherent in usual numerical computations.

Because experimental data points are discrete, a DRT estimated from them will itself involve discrete points, thus involving ill-conditioning. But if the DRT involved in the data includes only discrete Debye-type relaxations, it turns out that the inversion problem reduces to a direct-fitting one, and the DRT points may be estimated as accurately as the data itself, except for possible truncation errors [11, 12]. For this situation, a given DRT point, say τ_i with $1 \leqslant i \leqslant M$, may be accurately estimated from a few ω values whose inverses may be close to or far from the point if the original data are essentially error free. In this case, no localization restrictions, such as those of equations (1) and (2), apply, contrary to the conclusions in [4, 5]. Even with noisy data, it has been found that discrete-DRT estimation of this type is extremely robust [12, 13].

This paper deals, therefore, with continuous distributions represented by discrete points, distributions where inversion is intrinsically ill-posed as well as ill-conditioned, even in the absence of noise. A consequence is that no matter how accurately a given inversion solution can yield the original frequency-response data, one must always expect that there will be errors in the DRT estimates, no matter how noise-free the original data. Although inversion has been used for both experimental electrical and mechanical dispersion data and for synthetic data with random errors [8, 11–13], little attention has been devoted to quantitative estimation of errors associated with the ill-posed character of the problem for wide-range data. To assess the ultimate accuracy associated with an appropriate inversion procedure and thus to be able to examine the domain of applicability of equations (1) and (2), we deal here with frequency-response data correct to seven or more significant figures.

To do so, one must derive such data from a known analytic expression for the DRT of interest. It is reasonable to choose one which has been widely used and found useful for analysing both mechanical and electrical dispersion data: the Kohlrausch–Williams–Watts (KWW) distribution [14]. For this paper, we pick the usual KWW relaxation spectrum, the normalized KWW0 (that form of the KWW associated with stretched-exponential temporal response) DRT [15]† and [16]‡, $G_0(\tau_N)$, with its parameter β_0 taken equal to 0.5, since an analytic expression is available for it for this value of β_0, both with low τ cutoff ($\tau_{min} > 0$) and without it [17]. Low τ cutoff is physically necessary for the KWW0 and other similar

† Equations (40)–(43) in this work are less general expressions than the present equations (A.5) and (A.6).

‡ The symbol σ_0 should be removed from the right end of equation (12) in this work, and the word 'frequency' in the third line from the bottom of the second column of p 111 should be 'temperature'.

response models, and its effects have been observed at very high frequencies [15] and [18]†. The normalized DRT with cutoff may be expressed in dimensionless form [15] as

$$G_W(\tau_N) \equiv \tau_N G_0(\tau_N) = R(\tau_N/4\pi)^{0.5} \exp(-\tau_N/4), \tag{3}$$

where $G_W(\tau_N)$ is zero for $\tau_N < \tau_{N\,\mathrm{min}}$, and for $\beta_0 = 0.5$

$$R \equiv \Gamma(0.5)/\,\Gamma(0.5, \tau_{N\,\mathrm{min}}/4). \tag{4}$$

Here, $\Gamma(a, x)$ is the incomplete gamma function; $\tau_N \equiv \tau/\tau_W$, where τ_W is the characteristic relaxation time of the KWW0 response; and R is a scale parameter used in fitting inversion estimates to this equation. The units of τ_W are seconds, but they will be omitted hereafter. In the absence of cutoff, $\tau_{N\,\mathrm{min}} = 0$ and $R = 1$, which is the usual result. But when $\tau_{N\,\mathrm{min}} = 10^{-3}$, the value which will be considered herein, $R \simeq 1.018\,163\,79$, which is the value required for normalization. For $\tau_N = \tau_{N\,\mathrm{min}} = 10^{-3}$, $G_W(\tau_N)$ is about 9×10^{-3}.

$G_W(\tau_N)$ without cutoff is non-zero within the full range $0 < \tau_N < \infty$ but it decreases rapidly for large τ_N and reaches a value of less than 4×10^{-11} at $\tau_N = 100$. Thus, for practical purposes, one may consider it zero for larger values. Now in order to test the above localization restrictions adequately it is important to deal with a DRT of finite and known extent. The theoretical stretched-exponential temporal response (an autocorrelation function) associated with $G_0(\tau_N)$ is [15, 18]

$$\phi_T(t_N) = \phi(0) \exp\{-(t_N)^{\beta_0}\}, \tag{5}$$

with $\beta_0 = 0.5$ here and $t_N \equiv t/\tau_{\mathrm{W}}$. It turns out that $\phi(0) \equiv R$, so $\phi(t_N)$ does not go to exactly unity when DRT cutoff is present unless it is renormalized.

A brief description of the LEVM‡ variable τ, complex nonlinear least-squares (CNLS) inversion/fitting method used herein [11, 12, 19, 20§] and [21]‖ is included in the appendix. In order to assess nearly ultimate available accuracy, we apply the method to synthetic frequency-response data sets accurate to eight or more significant figures over most of their ranges. These data sets are derived from the exact G_W DRT expression of equation (3) and lead through equation (A.1) of the appendix to normalized (resistive or dielectric) $I(\omega)$ response. The effects of random error in data of the present type have been investigated elsewhere [12]; here, although we show some inversion results for data with random errors, we concentrate primarily on the consequences of ill-conditioning for essentially exact data. We consider both no-cutoff and cutoff situations with differing ω ranges. Finally, we use our DRT estimates to obtain the temporal response associated with the original frequency-response data without employing Fourier transformation, an inaccurate approach for discrete data [22], and one that again generally requires the use of equal arithmetical intervals and thus a very large number of required points for wide-range data. See [23], however, for an alternative Fourier transform approach.

† The word 'out' in the third line from the bottom of the first column on p 820 should be 'but'.

‡ The latest version of the LEVM complex nonlinear least squares fitting/inversion computer program, V. 7.11, may be obtained at no cost from Solartron Instruments, Victoria Road, Farnborough, Hampshire, GU147PW, UK. Both the extensive manual, of more than 160 pages, and the program (including Fortran source code) may be downloaded by accessing http://www.physics.unc.edu/∼macd/. The earliest version of the predecessor of LEVM was developed in the period 1979–81, and it has been greatly expanded and improved since then.

§ This work describes an earlier version of LEVM, then called LOMFP.

‖ This work provides more background on aspects of LEVM and its capabilities, and the LEVM manual provides further definitive information.

2. Analysis results

2.1. Background

The continuous-distribution points of nearly all previous inversion approaches have been presented in graphs with the estimated c_i distribution-strength values (see the appendix) shown on a linear ordinate axis, and they have rarely been compared directly with the predictions of a known distribution. Such plotting, when it is so compared, provides a useful qualitative measure of the adequacy of the inversion procedure, particularly for the largest values, but it can show little or nothing of differences for DRT values less than 1 or 2% of the peak. Alternatively, plotting with a logarithmic ordinate scale can show results and comparisons over the full range of the estimated DRT but with reduced resolution. Here we avoid all these difficulties by using both a logarithmic ordinate scale and also by showing separately the relative errors of the estimated points.

All the inversion calculations discussed here were carried out using the LEVM CNLS computer program. It can invert temporal response data or frequency-response data. In the latter case, real, imaginary, or complex data may be inverted. Many inversion procedures use only one or the other, and particularly the imaginary part alone, but when practical it is better to analyse full complex data, a procedure that provides more error averaging than one using only a single part of the data. Note that the ability to use either part of complex data allows an inversion program to estimate one part from knowledge of the other, a simpler and more accurate approach than using ordinary Kronig–Kramers analysis. The moving-average inversion approach [5,8] has not been used here for several reasons: first, it involves sampling localization [4,5] and so it is not logically appropriate for testing such localization; second, it does not yield as accurate fits of the original data as does the LEVM approach; and finally, it leads to small non-physical negative values for some DRT estimates.

Most DRT estimates used here are for $M = 19$, the maximum available in the LEVM CNLS computer program. For the present essentially noiseless ω data, a larger value of M could be used but is unnecessary here. In this paper, no comparisons between the original frequency-response data and the fit results are presented because with such data the fits are so close that the standard error of the relative residuals, for each fit, S_F, is of the order of 10^{-5}–10^{-8}. All comparisons shown will therefore deal with exact DRT parameters and those estimated from the inversion procedure. For $M = 19$, the present LEVM CNLS procedure involves trying to find the absolute minimum of an objective function (least-squares) in a computational space of 38 dimensions, a daunting task. Present procedures do not guarantee the invariable discovery of such a minimum, but we believe our solutions are all at least satisfactorily close to it, as indicated by the results of repeat fitting with different initial parameter choices.

2.2. Frequency- or time-response inversion → DRT estimation

2.2.1. No low τ cutoff: $\tau_{N\,\mathrm{min}} = 0$. Figure 1 includes a solid curve illustrating the behaviour of exact values of $G_W(\tau_N)$ over more than nine decades of τ_N variation. Here $R = 1$, $\tau_W = 1$ and $\beta_0 = 0.5$. Inversion results are shown for two choices of $\{\log(\omega_{N\,\mathrm{min}}), \log(\omega_{N\,\mathrm{max}})\}$ values where $\omega_N \equiv \omega\tau_W$. The $\log(c_i)$ inversion results for the logarithmic range of $\{-3, 7\}$ are indicated by open circles, showing ones whose centres should fall exactly on the $G_W(\tau_N)$ line if there were no inversion errors. A crude indication of the accuracy of the inversion estimates is provided by showing, as solid circles, the exact DRT c_i values corresponding to the τ_{Ni} values found in the inversion. If an open-circle point seems to encircle the nearest solid one uniformly, the relative error in that estimate must be of the order of 1% or less. Inversion results for $\{-4, 5\}$ are indicated on the graph by open triangles and show that the τ_{Ni} values

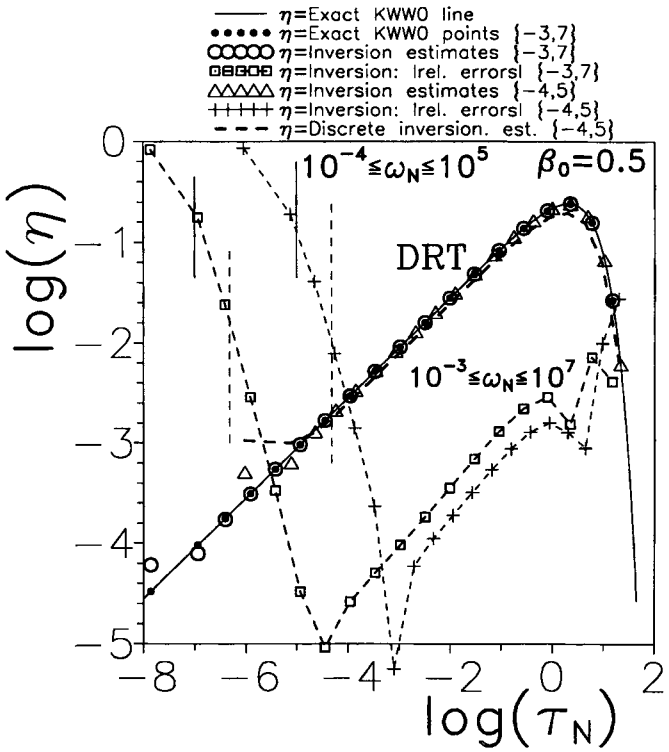

Figure 1. Comparison of the $\beta_0 = 0.5$ KWW0 DRT without DRT cutoff with inversion estimates for two frequency ranges denoted by $\{\log(\omega_{N\,\min}, \log(\omega_{N\,\max})\}$. The solid symbols are exact KWW0 points and the open circles are the corresponding inversion estimates for the $\{-3, 7\}$ range. The open triangles show the inversion estimates for $\{-4, 5\}$, but the corresponding exact points are omitted for this situation. The magnitudes of the exact relative errors of the inversions are shown by dashed curves. Here $\tau_N \equiv \tau/\tau_W$ and $\omega_N \equiv \omega\tau_W$, where τ_W is the characteristic relaxation time of the KWW0 model. Vertical solid and dashed curves show the positions of the $A = 0$ and 1 sampling localization limits of equations (1) and (2).

are different, even in their overlap range, for the two $\omega_{\max}$ choices. In addition, the heavy dashed curve for the $\{-4, 5\}$ situation shows the results of an inversion where it was assumed that the data involved an intrinsically discrete (not continuous) distribution. The results are clearly much more inaccurate than those shown for the continuous DRT approach. See the appendix for a discussion of these two procedures.

Much more instructive error assessment is provided by the lighter dashed curves in figure 1. These curves and those in the following figures were calculated by comparing the estimated c_i from an inversion with exact values, c_{Ti}, for the same τ_{Ni} obtained from equation (3) using the value of R calculated from equation (4) with $\tau_{N\,\min} = 0$ for no cutoff and $\tau_{N\,\min} = 10^{-3}$ for all the cutoff results discussed below. We define the relative error here as $E_{Ri} \equiv (c_{Ti} - c_i)/c_{Ti}$. For simplicity, the logarithms of $|E_{Ri}|$ are shown, but some sign changes appear in the neighbourhood of the 'V'-shaped regions of the curves. The error curves indicate that for both ω_N ranges the exact relative errors are appreciably less than 1%, except near the two ends of the DRT. Since the rising parts of the two error curves on the left are spaced apart by two decades, reflecting the differences in their $\omega_{N\,\max}$ values, it is clear that even without cutoff of the DRT $\omega_{N\,\max}$ plays a crucial role in determining the smallest value of τ_N for which useful estimates of c_i can be obtained.

The four short vertical solid ($A = 0$) and dotted ($A = 1$) lines on the left-hand side of the curves of figure 1 indicate the positions of the two types of sampling localization limits. We see that at the $A = 1$ positions the relative estimation error is of the order of 1%, while it is somewhat larger than 10% for the less stringent $A = 0$ choice. For the present DRT inversion procedure, we have thus quantified the effects of the two localization error choices in the low τ region. There is clearly no abrupt or more rapid increase in error to the left of the $A = 1$ lines, as implied by the work of Davies and Anderssen [4, 5]. Note that the error curves also show a rise near their right-hand sides, but the region where these rises occur is well within the $A = 0$ and 1 limits associated with $\omega_{N\,\min} = 0.001$. It thus seems clear that such increasing errors arise from the effective cutoff (rapid decay towards zero) of $G_W(\tau_N)$ for $\tau_N > 1$. Since it is possible that the CNLS inversion method may yield less accurate estimated results for such a rapid-decay region, it is particularly germane to consider the effects of an abrupt $\tau_{\min}$ cutoff, such as that discussed in section 1 for $G_{\rm W}(\tau_N)$. Although most of the rest of this paper deals with such situations, the following two figures demonstrate some of the strengths and weaknesses of the present inversion method for limited-range data without explicit DRT cutoff.

The rapid increase in relative error magnitude as τ_N approaches $\omega_{N\,\max}^{-1}$ or $\omega_{N\,\min}^{-1}$ is readily explained. The present frequency-response data are themselves cut off and do not extend beyond $\omega_{N\,\max}$ or below $\omega_{N\,\min}$, but the smallest (largest) τ_{Ni} in the fit model is associated with a simple Debye frequency response (see equation (A.5)), a response which is not abruptly terminated at $\omega_{N\,\max}$ ($\omega_{N\,\min}$) as are the data points. It is the inability of the present inversion model to adequately represent the abrupt cutoffs of the frequency-response data that necessarily leads to the errors shown here at the extremes of the τ_N range. Since it is important to distinguish between limited frequency-range effects and those associated with τ_N cutoffs, in order to test sampling-localization effects, we need to consider situations where one or the other dominates.

Figure 2 compares inversion errors for several data and quadrature-weighting choices. All except the curve with star symbols are for inversion of frequency-response data for the range shown on the figure. Further, all the results except the curve with triangle symbols involve variable quadrature weighting, appropriate for data involving either equally spaced fixed points on a logarithmic scale or free or fixed non-uniformly spaced points [11, 24]. The curve with triangular symbols involves both fixed, uniformly spaced $\log(\omega_N)$ points and a constant quadrature weight, e.g. [2, 3]; clearly a very poor choice here.

The results with + symbols are the same as those in figure 1 for the present range and involve inversion of the complex data: simultaneous fitting of both its real and imaginary parts. In the middle region, the results shown in the figure indicate that the real-part fit is appreciably more accurate than the complex one, and the imaginary-part fit is appreciably worse. These differences are not necessarily similar for other DRT estimates. A comparison of the complex fits with all τ points free to vary (+ symbols) with those involving fixed, uniform logarithmic spacing (open square symbols) shows that in this instance the results are very similar, with the latter one slightly superior. Finally, essentially exact temporal data were generated using equation (5), and LEVM was used for inversion to estimate the DRT, again with variable weighting and all τ points free to vary. Although there are appreciable differences in the accuracies of the bottom five curves of the figure, in the middle region their accuracies are all sufficiently high that the differences are not of major importance.

Values of 10^5 $S_{\rm F}$† found for the original fitting/inversion for the six situations involved in the present figure were about 2.7, 1.1, 5.8, 3.0, 1.6, and 0.28 in the order listed in the caption. The $S_{\rm F}$ of the fit which led to the very inaccurate open-triangle curve here is not substantially

† $S_{\rm F}$ = relative deviation of frequency-response data residuals.

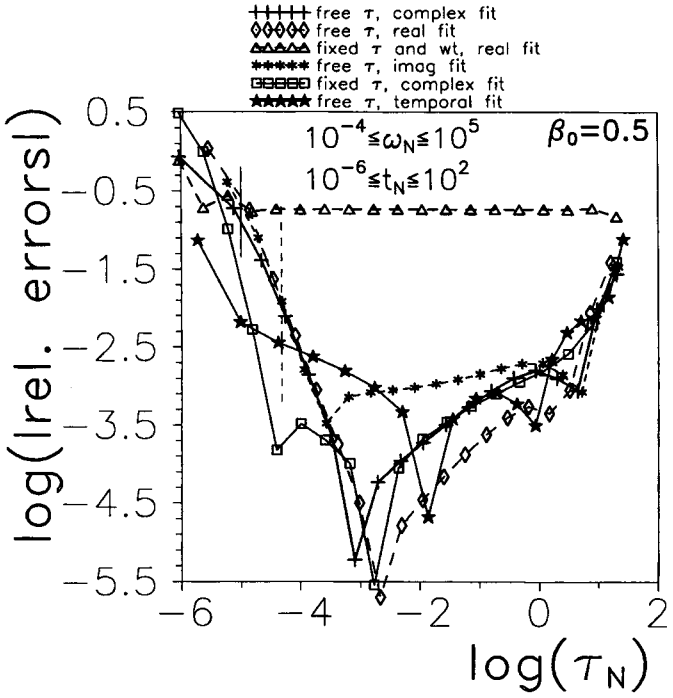

Figure 2. Comparison of the magnitudes of exact relative errors of no-cutoff DRT estimates for inversions of various different types of data without any added errors. The first five fits in the listing involve frequency-response data, and the last one is for stretched-exponential temporal response data. Results involving open-triangle symbols involve fixed, equally spaced τ points with a fixed quadrature weight. All the other results involve the usual choice of variable quadrature weights appropriate for all LEVM inversions of continuous distributions.

larger than those of the other very much more accurate curves, and the very small S_F value for the star-symbol temporal inversion did not lead to appreciably better accuracy of the estimated DRT points shown than that of the other variable-weighting curves. Therefore, the value of S_F is not a good predictor of the accuracy of the associated DRT estimate, a result to be expected when ill-conditioning and ill-posed effects are present.

Although this paper is concerned primarily with DRT inversion errors arising from the effects of limited-frequency ranges and cutoff of the DRT used in calculating frequency-domain data sets containing no added errors, details of the dependence on M of errors in such inversion estimates have not been previously presented. Therefore, figure 3 shows such no-cutoff results for the $\{-4, 5\}$ situation of figure 1. Results are presented for various values of M, both for the original, virtually exact frequency-domain data and for data to which independent, normally distributed, zero-mean proportional errors have been added to both the real and imaginary parts [12, 13].

These results for data with 1% random errors well illustrate the effect of ill-posed inversion. Notice that as M is increased for the no-added-error situation, the relative errors of the estimated DRT points generally decrease. On the other hand, with added 1% proportional errors in the data, the relative inversion errors generally increase in magnitude as M increases from 9 to 13 estimated points. Even with 11 points, however, the relative error magnitude of most points is still close to 1%, but when $M = 13$ that of some middle-range points begins to approach 10%. The ubiquitous tradeoff between accuracy and resolution is evident here and suggests that for

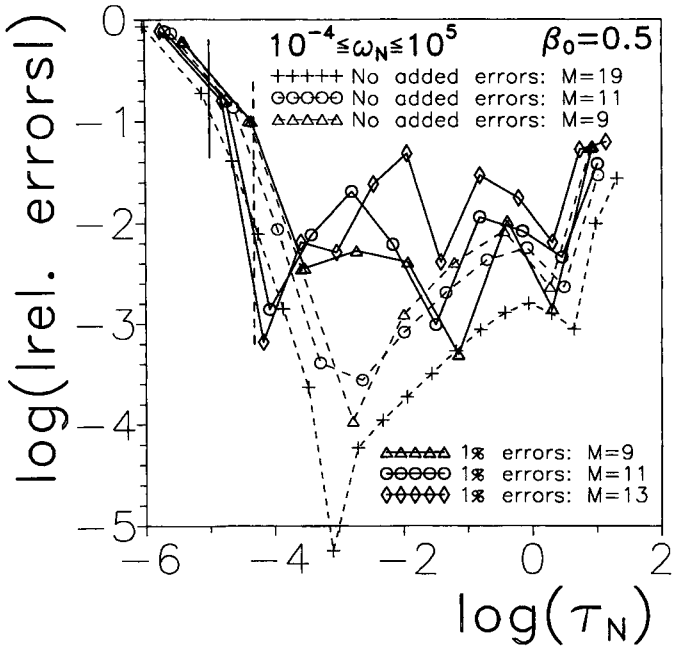

Figure 3. Comparison of the magnitudes of exact relative errors of no-cutoff DRT inversion estimates for frequency data without added errors and those with 1% random proportional errors added for various numbers of estimated DRT points, M.

the present 1%-error situation the value of M should be limited to 12 or less. Additional detailed inversion results for noisy data have been discussed in [12] and [13] and further inversion of noisy data for a more stringent data situation is presented below.

2.2.2. *Low τ cutoff:* $\tau_{N\,\min} = 10^{-3}$

No added noise. Most of the subsequent inversion results are for data involving the equation (3) KWW0 DRT with cutoff so $R \simeq 1.018\,16$. Figure 4 is included to illustrate further the difference between fitting the data assuming a discrete DRT and that obtained by (correctly) assuming a continuous DRT. It shows results for two $\{-4, 4\}$ cutoff situations, ones where the localization restrictions have no effects in the range shown. We compare the results of estimating the continuous $G_W(\tau_N)$ by both a discrete approach (letter 'D' in the graph), where the distribution is assumed, *ab initio*, to be composed of discrete points, and for the continuous-distribution inversion approach (letter 'C'). Because the variable weighting is different for these two calculations, the estimated peak value of the discrete distribution is smaller than that of the continuous one by about a factor of 1.7. The direct difference curve is not shown, but its logarithmic absolute values decrease rapidly from about -0.01 at the left to about -0.4 over the rest of the range.

When one compares the estimated DRT-strength points at the lowest τ_{Ni} values in figure 4 and in subsequent figures involving DRT cutoff at $\log(\tau_N) = -3$, one sees that although these points do not agree with the corresponding exact ones, they lie appreciably below the exact DRT curve and indicate an approach towards zero caused by the abrupt cutoff. As discussed earlier, the present inversion approach cannot model a discontinuous cutoff change exactly, but it can

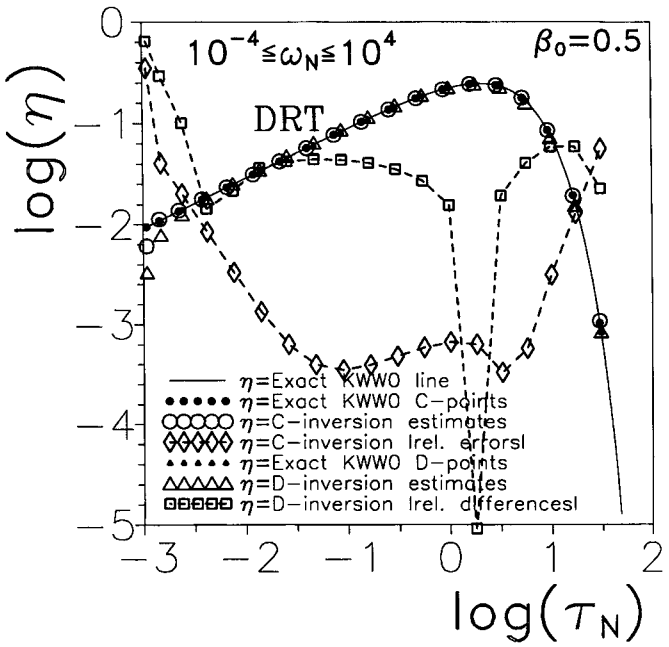

Figure 4. Inversion results for frequency-response data constructed from the exact continuous KWW0 DRT with a lower cutoff at $\tau_N = 10^{-3}$. Results are shown for the CNLS inversion method which assumes the distribution is continuous (C) and for the approach which assumes the DRT is discrete (D). The discrete points and the relative difference curve (between exact results and the discrete-inversion estimates) were scaled by normalizing the peak discrete estimate to equal that for the continuous DRT.

nevertheless suggest the presence of such an effect, one quite different from that appearing at the left extreme of the no-cutoff, limited-frequency range results depicted in figure 1 where the smallest τ_{Ni} point lies above the DRT curve.

When the discrete-approach points of figure 4, indicated by the open triangles, are normalized to have the same peak value as the continuous ones, the logarithm of the absolute values of the resulting differences, plotted with open-square symbols, shows that even then the differences are much larger than the C ones, except near the edges of the graph and near the zero crossover point at the peak. It is thus clear that even with the same peak normalization there are significant differences between the results of the D and C approaches, indicating that it is a poor approximation to use the former for continuous-distribution situations [11, 12]. Incidentally, it has been shown elsewhere [11–13] that comparison of the results of the two approaches, both with two or more different values of M, allows unequivocal identification of the type of distribution present.

Figure 5 presents more details about fitting results for the $\{-4, 4\}$ range choice with cutoff. Although this choice does not lead to any sampling localization limits within the τ_N range shown, we nevertheless see characteristic rises in the estimated relative error at the τ_N extremes. Just as sampling localization involves a finite frequency range which sets limits on the τ range over which inversion results may be significant, it is evident that a DRT that is only non-zero over a finite τ range leads to analogous limits on the significant inversion range, limits which seem to be, from figure 5, somewhat less than one-half a decade at either end of the DRT. It thus appears that cutoff limitations may dominate sampling-localization limits, and

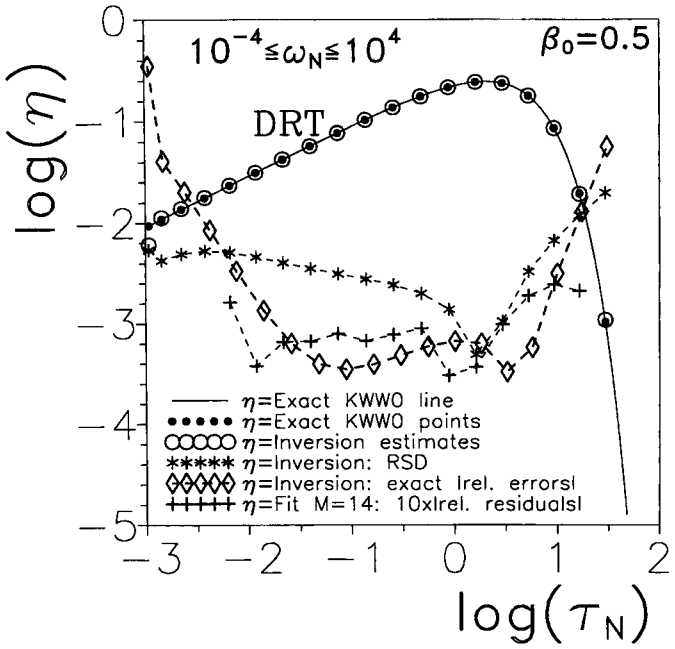

Figure 5. Comparison of various uncertainty/error measures for the inversion of the $\{-4, 4\}$ cutoff DRT. The RSD associated with the inversion are compared with the magnitudes of the actual relative errors of the inversion and to the magnitudes of the relative residuals of a fit of estimated inversion points to the exact model. Three low-end points and two high-end ones were omitted, and the fit relative residuals have been scaled up by a factor of 10.

when they do, the range over which a good inversion estimate of the DRT is possible is nearly a decade shorter than the extent of the DRT itself, at least for DRTs similar to the KWW0.

A curve showing the relative standard deviations (RSD) of the estimated c_i DRT strength parameters of the frequency-response inversion fit is included in figure 5. In addition, a relative-residual curve is presented for a direct fit of the 14 middle-range parameter values to the KWW0 model with cutoff, the one used to generate the original frequency-response data. For this fit, we allow the R and τ_W parameters of the model to be free variables. Here the ordinate values have been increased by a factor of 10 and the three smallest τ points and the two largest τ ones were omitted. We see that for fitting with most of the boundary-error points omitted, the fit relative residuals are much smaller than the original relative errors and the RSDs of the inversion estimates.

The parameter estimates obtained from the $M = 14$ LEVM fit to the KWW0 model involving both equations (3) and (4) with cutoff were $R \simeq 1.000\,81$ and $\tau_W \simeq 0.9981$, values which were both unity for the exact KWW0 cutoff model used to calculate the frequency response. We define $S_{F\tau}$ as the RSD of the residuals for a fit of DRT points obtained from inversion to a DRT model, here that of equation (3). The present fit also led to an estimate of the mean over the distribution, $\langle \tau_N \rangle_{CO}$, of 2.036 321 and to a value of $S_{F\tau}$ of about 0.0022. When the mean value is divided by 2, the value for the DRT without cutoff, one obtains 1.018 162, essentially identical with the R value following from equation (3) with $\tau_{N\,\mathrm{min}} = 10^{-3}$, thus indicating a means of calculating $\langle \tau_N \rangle$ for cutoff situations from non-cutoff results. A fit of the same 14 c_i estimates to the model without cutoff (equation (3) alone) led to the same τ_W estimated value as with cutoff but to $R \simeq 1.018\,989$. When this value is divided by 1.000 81,

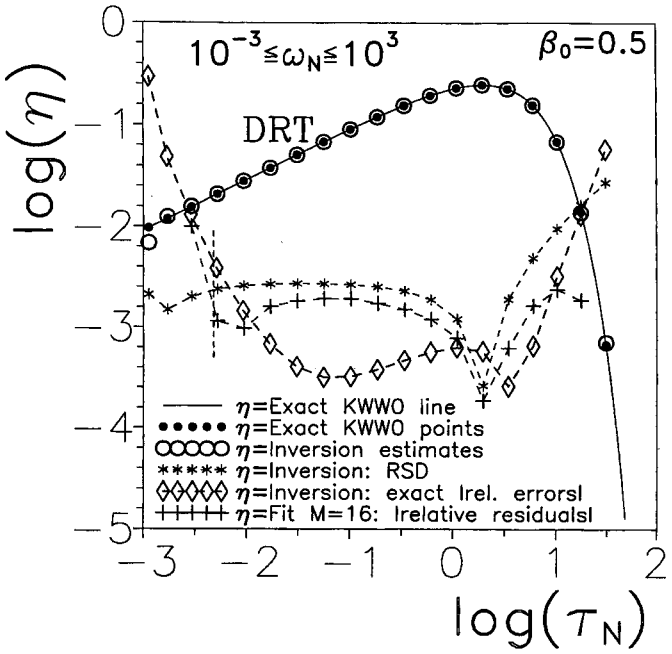

Figure 6. Inversion results for the $\{-3, 3\}$ cutoff DRT for comparison with the larger-range ones of figure 5. An $A = 1$ localization limit is shown here by the vertical dashed line.

one obtains 1.018 164, the correct value, showing the consistency of the results.

Figure 6 for $\{-3, 3\}$ involves an $A = 0$ localization limitation at the left axis and an $A = 1$ one shown by the dashed vertical line. The relative error curve is quite close to that of figure 5, where no such limitations apply, but its points are actually somewhat smaller here for $\log(\tau_N) < -1.5$. It thus appears that there is no localization limitation here, and the rise at the left is dominated by τ cutoff. The KWW0 model-fit curve here involves omission of only two left points and one right one, and its relative residuals are consequently larger than those of figure 5 because the present fit includes more of the high-error boundary-region points than the figure 5 one.

Figure 7 involves an ω range that leads to sampling localization limits which fall mostly to the right of the left region where cutoff is significant. It is evident that for this case quite small error estimates are obtained for the region between the low τ $A = 0$ and 1 limits. Thus the $A = 1$ limit is not dominant here, as it is for the two no-cutoff cases shown in figure 1. The present results suggest that for abrupt cutoff, the limitation arising from such cutoff is again of primary importance, as it is in the results of figures 4–6. Nevertheless, comparison of the error curves of figures 6 and 7 show that in roughly the region $-2.5 < \tau_N < -1$ the figure 7 relative errors are larger than those of figure 6.

Further, the figure 7 RSD curve lies appreciably higher than that for figure 6 and its values are much poorer estimates of the actual relative error than those in figure 6. The open-square curve shows RSD results when the left DRT point parameters, $\{\tau_{N1}, c_1\}$, were held fixed during the inversion (but not eliminated). Only for $\log(\tau_N)$ less than about -1.3 do the RSD estimates fall below the original inversion RSD predictions, thus showing the effect of uncertainty in the lowest DRT point. Incidentally, after inversion with all $\{\tau_{Ni}, c_i\}$ points taken as free parameters of the fit, the usual case here, one finds that if all the resulting τ_{Ni} estimates are

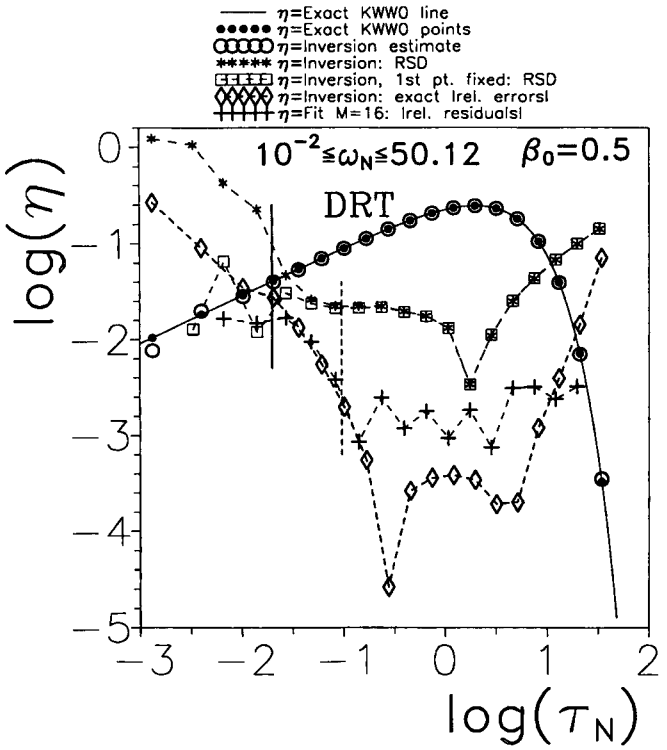

Figure 7. Inversion results for the cutoff DRT for the narrow range $\{-2, 1.7\}$. Two sampling-localization vertical lines are shown.

held fixed, a further inversion shows that the RSD estimates for the full c_i set all fall to the order of 10^{-5}, although the c_i estimates themselves remain essentially unchanged. Thus, RSD estimates should not always be trusted! Although the $A = 1$ criterion is not directly applicable here, the present limited ω range increases the figure 7 errors very considerably compared with those of figures 5 and 6. If the τ region over which a DRT is expected to be significant is at least roughly known, it therefore seems prudent to make measurements over a range of about $(10\tau_{max})^{-1} \leqslant \omega \leqslant (\tau_{min}/10)^{-1}$ if convenient, or at least over $(\tau_{max})^{-1} \leqslant \omega \leqslant (\tau_{min})^{-1}$, the $A = 0$ criterion.

Added noise. Although the effects of added noise on inversion for a situation without DRT cutoff are discussed above for figure 3, the data and DRT-inversion problem are much more stringent for the figure 7 results just considered. Thus, it is of interest to investigate inversion for this situation when the data contain added random errors. As in the noisy results shown in figure 3, we use normally distributed random errors proportional to the magnitudes of the real and imaginary components of the $I(\omega)$ data [5,25]. We shall use noise magnitudes, α, of both 1% [12] and 4% [5]. Then the population mean of S_F should approach α as the number of independent-error fits increases [12,25]. Another method for inversion of noisy data, one which has some elements of similarity to that of this paper, appears in [26]. Since it has been found that a KWW response model can fit some measured immittance spectroscopy data with an S_F appreciably smaller than 1% [16], the present KWW0 fit/inversions of synthetic frequency-response data with 1% or greater random errors provide a realistic picture of what

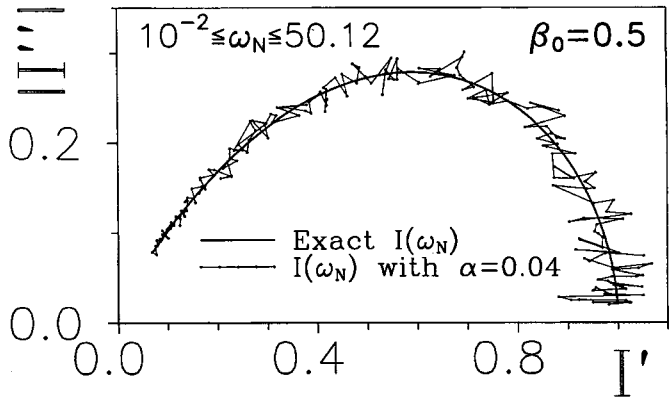

Figure 8. Complex plane plots of exact $I(\omega_N)$ model response and of $I(\omega_N)$ with 4% random Gaussian errors. For the latter, the points are connected by a line to guide the eye.

might be possible for comparable experimental data.

Figure 8 is a complex-plane plot of the imaginary and real parts of the $I(\omega)$ noisy data with $\alpha = 0.04$. As in figure (7), we use $10^{-2} \leqslant \omega_N \leqslant 50.12$, a narrow data range compared with that of the cutoff DRT. In this figure, which involves different vertical and horizontal scales in order to show the errors more clearly, lines connecting the 186 data points are included. It is evident that the errors are appreciable, especially at the right, low-frequency region, and that those for the real and imaginary points appear to be independent, as they should.

Figure 9 shows inversion results for $\alpha = 0.01$ and $M = 7$ and 8. For the present set of random errors, S_F decreased from about 0.012 for $M = 5$ to about 0.009 59 for $M = 6$ and reached a minimum of about 0.009 29 for $M = 8$. The figure suggests that the $M = 7$ results are somewhat better than those for $M = 8$. Note that the smallest τ point of both inversions lies above the correct value, an indication that the limited high-frequency cutoff dominates that of the DRT, different from the no-noise results of figure 7. Actual fits of these inversion results to the DRT model, both with the lowest τ point eliminated, led to $S_{F\tau} \simeq 0.017$† and 0.048 for $M = 7$ and 8, respectively. The parameter estimates for the six $M = 7$ points were $R \simeq 1.001$, $\tau_W \simeq 1.04$ and $\beta_0 \simeq 0.509$, showing some error amplification arising from inversion. When an inversion fit with $M = 9$ was attempted, two τ became equal, a degeneracy that precludes any useful value of $M > 8$ for this data set when the τ_N parameters are free. Although $M = 9$ or larger M results could be found with fixed rather than free τ_N values (see below), no improvement in estimating DRT model parameter values resulted.

Figure 10, for the $\alpha = 0.04$ situation, presents a log–log plot of the noisy $|I''(\omega_N)|$ data, as well as $G_W(1/\tau_N)$ DRT results, where the relation $\omega_N = 1/\tau_N$ has been employed. An approximate equation

$$G_W(1/\tau_N) = (2/\pi)|I''(\omega_N)|$$

was proposed long ago [27], and variants of it have been employed in the present area [28], so it is of some interest to compare these quantities for the present situation. We see that neither the peak values, after accounting for the $\log(2/\pi) \simeq -0.2$ factor, nor the shapes of the curves are very close, indicating that equation (6) is a poor approximation for data of the present type.

For $\alpha = 0.04$, S_F values of 0.0493, 0.0436, 0.0434, and 0.0433 were obtained as M increased from 4 to 7 with free τ_N parameters. Degeneracy set in for $M > 7$, but fits could be

† S_{Fx} RSD of DRT-fit residuals ($x = \tau$), or those of temporal data-fit residuals ($x = t$).

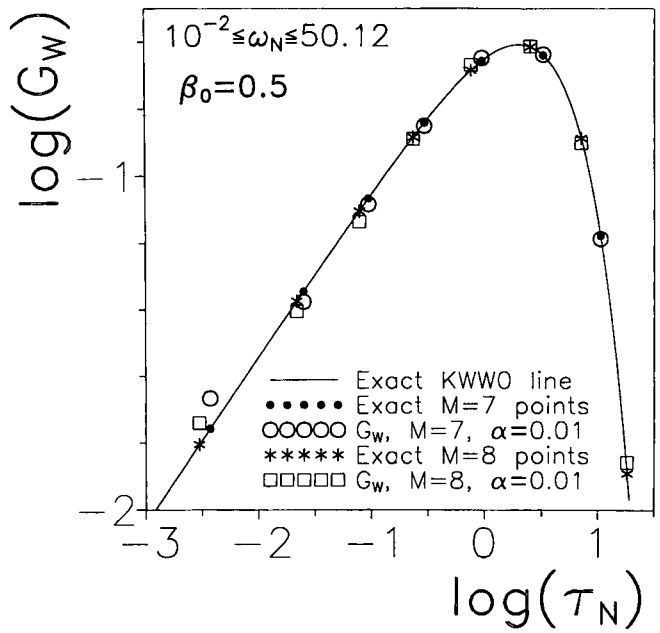

Figure 9. Log–log comparison of exact values for the KWW0 DRT with $M = 7$ and 8 inversion estimates obtained using limited-range frequency-response data containing 1% errors. The DRT is cut off at $\tau_N = 10^{-3}$ as in figure 7 and also for the results of figures 10 and 11 below.

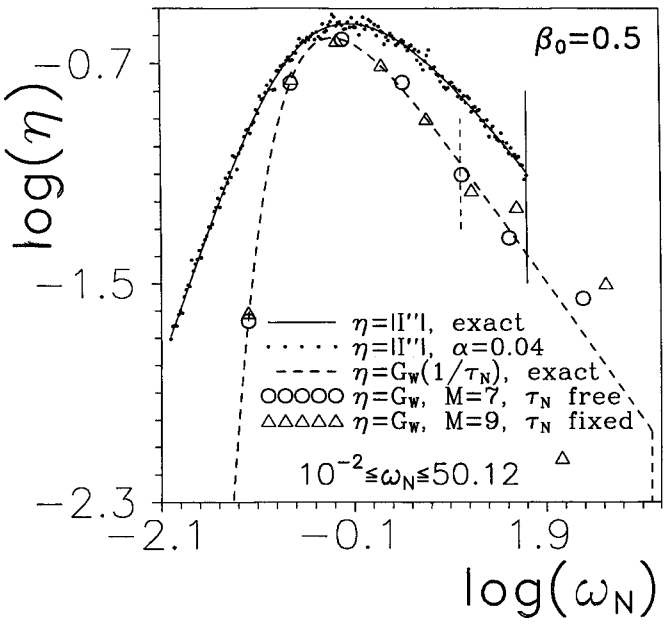

Figure 10. Comparison of log–log $|I''(\omega_N)|$ response containing 4% random errors with the corresponding exact KWW0 $G_W(1/\tau_N)$ DRT and with $M = 7$ and 9 inversion estimates for the noisy data. These DRT results were converted to the ω_N scale using $\tau_N \rightarrow \omega_N = 1/\tau_N$.

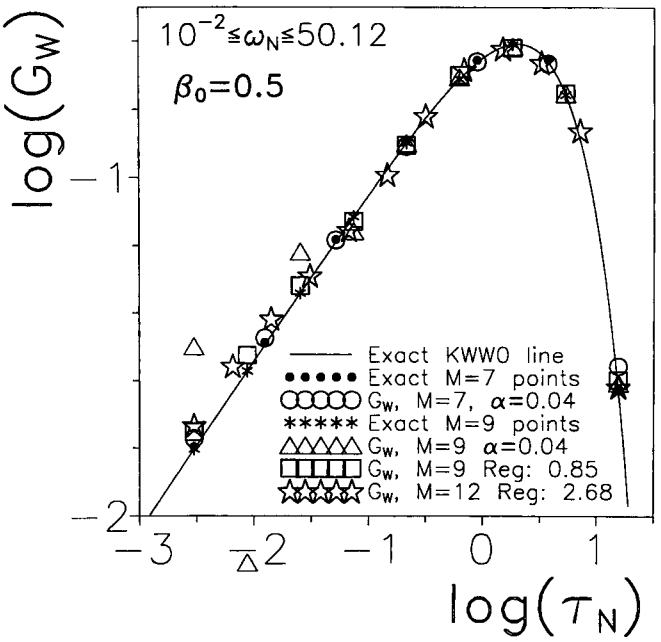

Figure 11. Log–log comparisons using data with 4% errors for fixed τ_N inversions with $M = 7$ and 9. Tikhonov regularization inversion points are also shown for $M = 9$ and 12.

obtained with larger M values for fixed τ_N values distributed evenly on a log scale. Figure 11 includes such $M = 7$ and 9 results. LEVM fitting of the inversion points to the KWW0 DRT with R, τ_W, and β_0 free, led to $S_{F\tau}$ values of 0.139 and 0.564 for $M = 7$ and 9, respectively. When the smallest two τ_N points were omitted, these values reduced to 0.022 and 0.039. But note that the first of these is for only five points and the second for 7. Further, another fit with $\alpha = 0.04$ but with a different choice of random numbers led to an appreciably larger $S_{F\tau}$ value for $M = 7$ than that given above.

It is of interest to compare $M = 7$ inversion results, with all the τ_N points fixed rather than free, as in the above $M = 9$ situation and to see if it is possible to obtain a larger number of significant inversion points in order to provide increased resolution. Figure 11 compares fixed τ_N $M = 7$ and 9 results for $\alpha = 0.04$. We see that here the $M = 7$ points are appreciably closer to the exact results than those shown in figure 10. With all seven points, $S_{F\tau} = 0.021$, and the R, τ_W and β_0 estimates for the $M = 7$ fits were 1.018, 0.960, 0.495, and 1.024, 0.991, 0.492, for the free and fixed τ_N inversions, respectively. Evidently, the estimates of the DRT parameters are both quite good and are comparable for these fits.

The $M = 9$ open-triangle points in figure 11 are only reasonable for the larger five τ_N values, so using $M = 9$ has led to fewer significant points than were obtained with the $M = 7$ fixed τ_N inversion. Since LEVM allows Tikhonov regularization to be carried out for fixed τ_N situations, however, we have applied it, using second-order regularization, a choice which proved to be superior to first- or third-order analysis. Results for the $M = 9$ optimum regularization parameter value, 0.85, are shown by the open-square symbols in figure 11. Fitting of these regularization results yielded $S_{F\tau} = 0.0447$, $R = 1.025$, $\tau_W = 0.960$ and $\beta_0 = 0.491$, slightly inferior to the $M = 7$ results. Figure 11 also shows regularized results for $M = 12$, ones involving the larger regularization parameter value of 2.68. Here, $S_{F\tau} = 0.069$,

$R = 1.028$, $\tau_W = 0.943$ and $\beta_0 = 0.490$. With the two lowest τ_N points omitted, the fit yielded $S_{F\tau} = 0.042$, $R = 1.016$, $\tau_W = 1.012$ and $\beta_0 = 0.500$, better 10-point DRT parameter estimates than obtained from the $M = 7$ inversions without regularization.

What do these noisy-data inversions tell us? First, that with $\alpha = 0.01$, we can expect for limited-range data to obtain seven to eight meaningful inversion points using free τ_N parameters, more than enough to identify the underlying DRT if it corresponds to one whose shape and analytical expression are known or can be determined. Earlier work with a different DRT and no appreciable data-range limitation showed that at least 12 significant points could be obtained for the present $\alpha = 0.01$ situation [12]. Surprisingly, one can also often obtain six or seven significant points for the present data situation even when $\alpha = 0.04$, an appreciably larger error level than is common in the immittance spectroscopy area.

Finally, for limited-range data of the present type the above results indicate that optimum Tikhonov regularization with a larger value of M than that used without it can lead to good estimation of KWW0 DRT parameters as well as higher resolution. Nevertheless, when the object is DRT identification and quantification, the main purpose of inversion, it is best to use the present free τ_N approach for the best accuracy until noise limits the maximum M that is practical with this inversion procedure. Using regularization will eventually reduce the accuracy of inversion estimates and increase the uncertainty of DRT identification, however, even for wide-range data.

2.3. DRT estimates → temporal response

Once an inversion of frequency-domain data has been carried out, the LEVM program may be used to calculate the associated temporal response by using the resulting DRT point estimates in equation (A.6) of the appendix. This simple and straightforward procedure, implemented in LEVM, allows one to generate temporal response over any range and with as many points as desired, but points outside the range $\tau_{min} \leqslant t \leqslant \tau_{max}$, either defined by the DRT range or by the $A = 0$ choice of equations (1) and (2), will generally be inaccurate, as illustrated below. Note that although inversion of a continuous distribution is an ill-posed problem, the use of the resulting DRT points in calculating temporal response is not ill-posed, and one may expect to obtain time-domain point estimates that may be more accurate than the DRT points themselves.

After one calculates the temporal response one may compare it with that expected from the DRT model. Here that is given by the stretched-exponential expression of equation (5) with $\beta_0 = 0.5$. For present purposes, we will not combine equations (4) and (5) but use equation (5) for fitting with $\phi(0) = R$, τ_W and β_0 all taken as free parameters. Exact DRT point values should then yield fit estimates of $R = 1$ for data without cutoff, $R \simeq 1.018\,163\,79$ for $\tau_{N\,min} = 10^{-3}$, and 1 and 0.5 for the other two parameters. We shall define the S_F of temporal fits of the present type as S_{Ft}.

The results of fitting to equation (5) are shown in figure 12 and lead to virtually the same relative residual curves for both the two disparate ω ranges, in spite of their appreciably different DRT relative-error curves. Only for $\log(t_N) \lesssim -2.4$ can one see infinitesimal differences. The parameter estimates found from the $\{-4, 4\}$ fit were $R \simeq 1.0159$, $\tau_W \simeq 1.0038$ and $\beta_0 \simeq 0.5004$, with $S_{Ft} \simeq 0.002$, while those for $\{-2, 1.7\}$ were 1.018 159, 1.000 01 and 0.500 001, respectively, with $S_{Ft} \simeq 4 \times 10^{-6}$, a very accurate fit. Note that the t_N range for this case has been extended on the low side by about 1.3 decades beyond the $A = 0$ $\tau_{N\,min}$ limit. It is significant that its exact relative error curve melds with that for the $\{-4, 4\}$ case for t_N values smaller than this limit.

Let us define the relative error for the temporal domain as $E_R(t_{Ni}) \equiv [\phi(t_{Ni}) - \phi_T(t_{Ni})]/\phi_T(t_{Ni})$, where $\phi(t_{Ni})$ is a value calculated from frequency-response DRT inversion

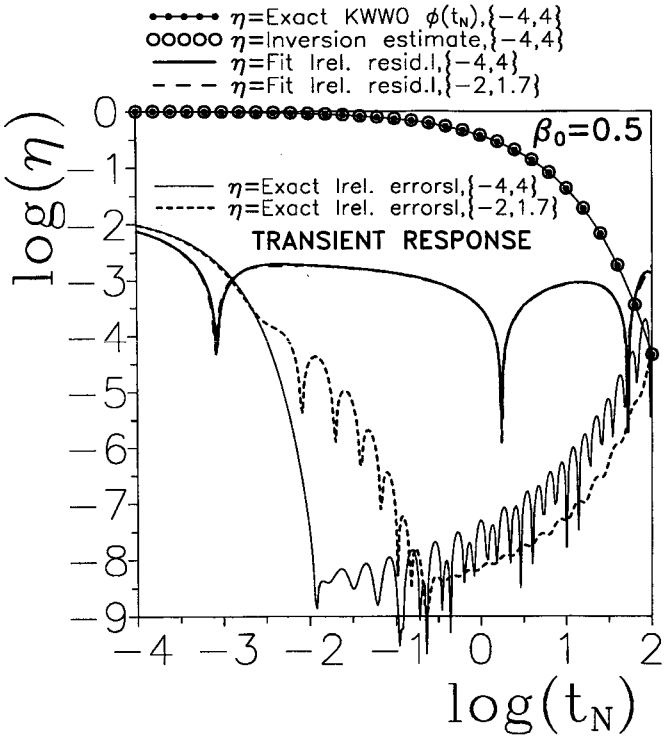

Figure 12. Temporal response, $\phi(t_N)$, calculated from the cutoff DRT estimates obtained from inversion of frequency-response data with ranges of $\{-4, 4\}$ and $\{-2, 1.7\}$. Relative residuals obtained from fitting $\phi(t_N)$ data to the exact stretched exponential model of equation (5) with free parameters $\phi(0)$, τ_W, and β_0 are shown, as well as exact relative residuals between the $\phi(t_N)$ data and the $\phi_T(t_N)$ model with exact parameter values. Note the extremely small values of errors except in the low t_N region where cutoff effects dominate.

results, as discussed above. Figure 12 shows that both relative-error curves are much smaller than the residual-fit curves over most of the range shown. The error and residual curves were all calculated for 301 points, with 50 points per decade evenly spaced in $\log(t_N)$. The resulting resolution is sufficient to show that many sign changes occur for the relative errors and that their magnitudes are remarkably low. Note that the error level for $\{-2, 1.7\}$ begins to exceed (in absolute value) that of the $\{-4, 4\}$ situation just at $\log(t_N) = -1.018$, just the pertinent $A = 1$ localization limit. Thus, although such localization does not preclude very small errors outside its limit, it nevertheless leads to larger errors than otherwise.

There is another way to show exact errors for the present situation, one appropriate when the proper value of β_0 is known. We use equation (5) with β_0 fixed at 0.5 and $\phi(0) = R = 1.018\,163\,79$ for the present cutoff cases and calculate the new error quantity, E, defined in figure 13. This should be zero for completely accurate $\phi(t_N)$ data. Figure 13 shows $E(t_N)$ curves for the two ranges of figure 12 and for the no-cutoff $\{-4, 5\}$ range of figure 1. The two curves for the cutoff ranges were calculated with the same points as those used to obtain the figure 12 results, and the no-cutoff $\{-4, 5\}$ one was calculated, using $\phi(0) = 1$, from the DRT inversion results shown in figure 1. The left-side rise of this no-cutoff error curve is small because neither of the A-choice localization limits falls within the range shown.

Although $E(t_N)$ is a different error measure from the exact relative error shown in figure 12,

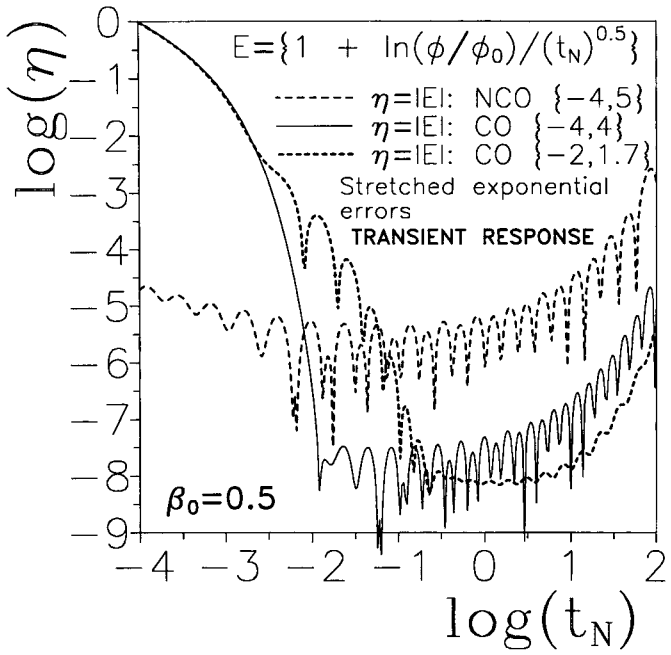

Figure 13. The error $E(t_N)$, as defined in the figure, calculated with exact $\phi(0)$, τ_W and β_0 values. Results are shown for one no-cutoff situation and two with cutoff.

it is remarkable how similar its values for the two cutoff situations are to the relative error in the region where $\log(t_N)$ is greater than about -1.5. In fact, it can be shown that for the present situation $E(t_{Ni}) = \ln\{1 + E_R(t_{Ni})\}/(t_{Ni})^{0.5}$, which reduces to $|E(t_{Ni})| \simeq |E_R(t_{Ni})|/(t_{Ni})^{0.5}$ when $|E_R(t_{Ni})| \ll 1$. Thus the two error magnitudes should be essentially equal at $t_{Ni} = 1$, as is indeed the case here.

The rise of the two cutoff E curves towards a value of $\log|E(t_N)|$ of about zero indicates that cutoff data and inversion results are unable to lead to accurate estimates of $\phi(t_N)$ in the region where t_N is smaller than about 10^{-2} or 10^{-3}. But the top curve of figure 12 indicates that $\phi(t_N)$ is so close to unity in this range anyway that the difference could not be distinguished experimentally in any case. Finally, the present DRT and time-domain results indicate that although the $A = 0$ and 1 localization limits are not associated with abrupt changes in errors outside the ranges they define, they nevertheless play important, if more subtle, roles in defining regions where inversion and temporal response errors become significant.

Acknowledgment

It is a pleasure to thank Dr Robert S Anderssen for valuable comments and suggestions concerning this paper and for broadening my knowledge of aspects of industrial inverse problems.

Appendix. Analysis methods

Let a complex quantity U represent the relaxation response of a dielectric system, a conductive system, or a rheological system. It is mathematically convenient and often physically

significant to express the normalized form of U, I, in terms of a DRT, say $g(\tau)$. Let $x \equiv \tau/\tau_o$, where τ_o is a characteristic response time of the fitting model, here taken as τ_W, so $\tau_N = x$, and we define $y \equiv \ln(x)$ and the dimensionless DRT $G(x) \equiv \tau_o\, g(\tau)$. We may then write [12, 15]

$$I(\omega) \equiv \frac{U(\omega) - U(\infty)}{U(0) - U(\infty)} = \int_{x_{\min}}^{x_{\max}} \frac{G(x)\,\mathrm{d}x}{[1 + \mathrm{i}\omega\tau_o x]} = \int_{y_{\min}}^{y_{\max}} \frac{F(y)\,\mathrm{d}y}{[1 + \mathrm{i}\omega\tau_o \exp(y)]}, \tag{A.1}$$

where

$$U(\omega) = U'(\omega) + \mathrm{i}\xi U''(\omega), \tag{A.2}$$

and thus

$$I(\omega) = I'(\omega) + \mathrm{i}\xi I''(\omega). \tag{A.3}$$

We follow the usual sign conventions and set the quantity $\xi = 1$ for conductive and rheological relaxation systems and to -1 for dielectric ones. Since the DRTs are taken normalized in the above, it follows that $I(0) = 1$ and $I(\infty) = 0$. The limits for the two integrals above are theoretically 0 and ∞ for the first one and $-\infty$ and ∞ for the second. In practice, however, one can take, with good accuracy, $y_{\min} = -40$ and $y_{\max} = 20$ for the no-cutoff KWW0 DRT situation. For low-end cutoff, the case also considered here, we set $x_{\min} = 0.001$, so $y_{\min} \simeq -6.91$. Incidentally, the F distribution above may be simply related to a distribution of activation energies [29] and is given by $F(y) \equiv xG(x)$, equivalent to the present $G_W(\tau_N)$ for the KWW0 situation.

The $G(x)$ DRT is defined at the impedance or complex resistivity level for a conductive system and at the complex capacitance or dielectric level for a dielectric system, one which involves no intrinsic dc conduction [11, 15, 16, 30]. Note that for mechanical relaxation, one usually deals with the modulus, M (not to be confused with the DRT point count, also denoted M herein), where $M = (\mathrm{i}\omega B)U(\omega)$ for non-dielectric systems. For electrical relaxation the constant B is either ϵ_V, the permittivity of vacuum or the cell capacitance associated with it. When proportional weighting is used, as here, CNLS fitting of data to a conductive-system model at either the original level or at the associated M level yields exactly the same parameter estimates and so the same DRT estimate for inversion.

The calculation of $I(\omega)$ for discrete data requires evaluation of the integrals in equation (A.1) by numerical quadrature. Consider discrete data extending over the range $\omega_{\min} \leqslant \omega_k \leqslant \omega_{\max}$ with $1 \leqslant k \leqslant N$. There is no problem for a purely discrete distribution, one involving delta-function lines, so that

$$G(x) = \sum_{i=1}^{i=M} d_i \delta(x - x_i), \tag{A.4}$$

where the d_i quantities specify the strength of the M lines of the distribution. But for a continuous distribution one must use approximate numerical quadrature. Let c_i be the strength parameter for a point τ_i of the continuous $F(y_i)$ or $G_W(\tau_{Ni})$ distribution, and take e_i as either d_i or c_i, as appropriate. Then equation (A.1) may be rewritten as

$$I(\omega_k) = \sum_{i=1}^{i=M} \frac{e_i w_i}{[1 + \mathrm{i}\omega_k \tau_o \mathrm{e}^{y_i}]}, \tag{A.5}$$

where $\omega_k \tau_o \mathrm{e}^{y_i} \equiv \omega_{Nk}\tau_{Ni} \equiv \omega_k \tau_i$. Note that the τ_i points need not be equally spaced on a linear or logarithmic scale, and both the e_i and τ_i quantities may be free parameters of a CNLS fit. Further, LEVM allows all e_i parameters to be constrained to be positive, a necessary requirement for realistic probability densities.

Equation (A.5) is exact for a discrete DRT with all the w_i weights equal to unity. But for continuous distributions, where $e_i = c_i$, this equation is only approximate because the

quadrature procedure is necessarily approximate for discrete data. Then, the quadrature weights are not unity but should be taken variable instead. Appropriate choices for them have been discussed elsewhere [11, 12, 24] and are used in the LEVM program.

The temporal response associated with equation (A.1) becomes, for discrete temporal data [15],

$$\phi(t_{Nk}) = \sum_{i=1}^{i=M} e_i w_i \exp(-t_{Nk}/\tau_{Ni}), \tag{A.6}$$

where $t_{Nk} \equiv t/\tau_0$ and $t_{Nk}/\tau_{Ni} \equiv t_k/\tau_i$. Note that

$$\phi(0) = \sum_{i=1}^{i=M} e_i w_i, \tag{A.7}$$

a quantity which should be unity for no cutoff and R for left-end cutoff of a KWW0 DRT. For this DRT, equation (A.6) thus approximates the stretched-exponential response of equation (5) by a weighted sum of simple exponentials.

In this paper, we shall use U_k to denote values which involve the model $I(\omega_k)$ through equation (A.1), and U_{dk} to denote data values. The parameters involved in U_k are estimated by CNLS fitting of the data to the U_k points, a process which requires iteration when any parameters of the model appear nonlinearly in it. For a fixed value of M the objective function which is to be minimized, O_M, is of the form [11, 12, 20, 21]

$$O_M = \sum_{k=1}^{N} [(R'_k)^2 + (R''_k)^2], \tag{A.8}$$

where $R'_k \equiv [U'_{dk} - U'_k]/V'_k$ and $R''_k \equiv [U''_{dk} - U''_k]/V''_k$ are weighted residuals. The actual weights involved in the fitting are $1/V'^2_k$ and $1/V''^2_k$.

There are several ways the present CNLS method of analysis instantiated in LEVM differs from other continuous-distribution inversion approaches. It is a full CNLS fitting method, one that can constrain all e_i strength parameters to be positive; it allows e_i and τ_i values to be fixed or free variables of the fit; fixed τ_i values may be distributed uniformly or non-uniformly on a linear or logarithmic scale; and quadrature weights may be taken independent of τ or may depend on it for either fixed or free τ_i values.

Although a CNLS inversion approach was described as early as 1968 [9], it treated a continuous distribution as intrinsically discrete (quadrature weights all implicitly taken as 1), as did the work of [10], and, in practice, fitting used only the real part of the data. In both [2] and [3], nonlinear least-squares fitting was employed, with fixed τ_i points distributed uniformly on a logarithmic scale and all quadrature weights taken equal, a common approach. Further, neither treatment actually involved full CNLS fitting. As the present work demonstrates in figures 2 and 4, the use of variable quadrature weighting can lead to far more accurate estimates of the parameters of a continuous DRT for either kind of least-squares inversion with variable τ spacing than does the use of equal fixed weights. Nonlinear regression alone provides some regularization [2], but variable quadrature weighting clearly is a crucial element of the LEVM inversion procedure for continuous-distribution estimation.

LEVM allows many different types of V_k weighting (error models) to be chosen; see, e.g., [11, 20, 21]. Here we need only consider proportional weighting, where $V'_k = U'_{dk}$ and $V''_k = U''_{dk}$ for all k values. If we define $D \equiv 2N - M_T$ as the number of degrees of freedom for a fit of complex data, then the variance of the fit, S^2_F, is just given by O/D. This is a standard definition of the variance, generalized to complex-data fitting [15, 20, 21]. For complex-data fitting, M_T is the sum of $2M$ and the number of all other free parameters of the fit, especially including the high-frequency limiting dipolar dielectric constant, $\epsilon_{D\infty}$, for

electrical relaxation situations. Note that S_F is the standard deviation of the relative residuals for proportional weighting and is thus then the RSD of the fit.

For the present inversions, the most stringent possible criteria for iterative convergence were used in LEVM, with all calculations carried out in double precision on a desktop computer. The convergence criteria used were the same for both exact and noisy data. Several criteria were always employed, and the one which could be least well satisfied led to the cessation of iteration. Usually, iterative reduction in the objective function continued until it was impossible to change the free parameter estimates any further, either because of noise or limited machine accuracy. This generally resulted in no further reduction in the L2 norm of the fit residuals at the 12th significant figure or better. For less accurate and quicker convergence this number could be set to a smaller value. Most fits converged with less than 100 iterations but a few involved many more when the most stringent convergence was required. For some fits with a large number of free parameters, we found, nevertheless, that replication of the fit a number of times was desirable to obtain a somewhat smaller minimum value of S_F. Each such replication used as input the results of the previous fit, and S_F for each replication either decreased towards a final limiting value or remained constant.

References

[1] Uhlmann D R and Hakim R M 1971 Derivation of a distribution function from relaxation data *J. Phys. Chem. Solids* **32** 2652–5

[2] Cost J R 1983 Nonlinear regression least-squares method for determining relaxation time spectra for processes with first-order kinetics *J. Appl. Phys.* **54** 2137–46

[3] Morgan F D and Lesmes D P 1994 Inversion for dielectric relaxation spectra *J. Chem. Phys.* **100** 671–81

[4] Davies A R and Anderssen R S 1997 Sampling localization in determining the relaxation spectrum *J. Non-Newtonian Fluid Mech.* **73** 163–79

[5] Davies A R and Anderssen R S 1998 Sampling localization and duality algorithms in practice *J. Non-Newtonian Fluid Mech.* **79** 235–53

[6] Groetsch C W 1984 *The Theory of Tikhonov Regularization for Fredholm Equations of the First Kind* (Boston, MA: Pitman)

[7] Winterhalter J, Ebling D G, Maier D and Honerkamp J 1999 Analysis of admittance data: comparison of a parametric and a nonparametric method *J. Comput. Phys.* **153** 139–59

[8] Anderssen R S and Davies A R 1998 Simple moving-average formulae for direct recovery of the relaxation spectrum *Mathematics Research Report* MRR 016-98 Centre for Mathematics and its Applications, The Australian National University, Canberra, Australia pp 1–30

[9] Holmes R and Stott M A 1968 Analysis of multiple relaxation processes *Br. J. Appl. Phys.* D **1** 607–15

[10] Baumgärtel M and Winter H H 1989 Determination of the discrete relaxation and retardation time spectrum from dynamic mechanical data *Rheol. Acta* **28** 511–19

[11] Macdonald J R 1995 Exact and approximate nonlinear least-squares inversion of dielectric relaxation spectra *J. Chem. Phys.* **102** 6241–50

[12] Macdonald J R 2000 Comparison of parametric and nonparametric methods for the analysis and inversion of immittance data: critique of earlier work *J. Comput. Phys.* **157** 280–301

[13] Macdonald J R 1996 Analysis of immittance spectroscopy data: model comparisons, universality?, and estimation of distributions of activation energies *Electrically Based Microstructural Characterization: Proc. Fall meeting (Boston, MA, Nov. 1996)* vol 411 (Pittsburgh, PA: Materials Research Society) pp 71–83

[14] Kohlrausch R 1854 Theorie des elektrischen Rückstandes in der Leidener Flasche Pogg *Ann. Phys. Chem.* **91** 179–214

Williams G and Watts D C 1970 Non-symmetrical dielectric relaxation behavior arising from a simple empirical decay function *Trans. Faraday Soc.* **66** 80–85

Williams G, Watts D C, Dev S B and North A M 1971 Further considerations of non-symmetrical dielectric relaxation behavior arising from a simple empirical decay function *Trans. Faraday Soc.* **67** 1323–35

[15] Macdonald J R 1996 Analysis of dispersed, conducting-system frequency-response data *J. Non-Cryst. Solids* **197** 83–110

Macdonald J R 1996 *J. Non-Cryst. Solids* **204** 309 (erratum)

[16] Macdonald J R 1997 Accurate fitting of immittance spectroscopy frequency-response data using the stretched exponential model *J. Non-Cryst. Solids* **212** 95–116
[17] Lindsay C P and Patterson G D 1980 Detailed comparison of the Williams–Watts and Cole–Davidson functions *J. Chem. Phys.* **73** 3348–57
[18] Macdonald J R 1998 The Ngai coupling model of relaxation: generalizations, alternatives, and their use in the analysis of non-Arrhenius conductivity in glassy, fast-ionic materials *J. Appl. Phys.* **84** 812–27
[19] Macdonald J R, Schoonman J and Lehnen A P 1982 The applicability and power of complex nonlinear least squares for the analysis of impedance and admittance data *J. Electroanal. Chem.* **131** 77–95
[20] Macdonald J R and Potter L D Jr 1987 A flexible procedure for analyzing impedance spectroscopy results *Solid State Ionics* **23** 61–79
[21] Macdonald J R and Thompson W J 1991 Strongly heteroscedastic nonlinear regression *Commun. Stat. - Simul. Comput.* **20** 843–86
[22] Thompson W J and Macdonald J R 1993 Discrete and integral Fourier transforms: analytical examples *Proc. Natl Acad. Sci. USA* **90** 6904–8
[23] VanderNoot T J 1992 Hilbert transformation of immittance data using the fast Fourier transform *J. Electroanal. Chem.* **322** 9–24
[24] Macdonald J R 1993 Some new directions in impedance spectroscopy data analysis *Electrochim. Acta* **38** 1883–90 see the appendix of this paper
[25] Macdonald J R 1990 Impedance spectroscopy: Old problems and new developments *Electrochim. Acta* **35** 1483–92
[26] Landl G, Langthaler T, Engl H W and Kauffmann H F 1991 Distribution of event times in time-resolved fluorescence: the exponential series approach—algorithm, regularization, analysis *J. Comput. Phys.* **95** 1–28
[27] Böttcher C J F and Bordewijk P 1978 *Theory of Electric Polarization* vol 2, 2nd edn (Amsterdam: Elsevier) equation (9.24)
[28] Nowick A S, Vaysleyb A V and Liu W 1998 Identification of distinctive regimes of behavior in the ac electrical response of glasses *Solid State Ionics* **105** 121–8
[29] Macdonald J R 1962 Restriction on the form of relaxation-time distribution functions for a thermally activated process *J. Chem. Phys.* **36** 345–9
[30] Macdonald J R 1999 Dispersed electrical-relaxation response: discrimination between conductive and dielectric relaxation processes *Braz. J. Phys.* **29** 332–46