Exact and approximate nonlinear least-squares inversion of dielectric relaxation spectra

J. Ross Macdonald
Department of Physics and Astronomy, University of North Carolina, Chapel Hill, North Carolina 27599

Three weighted, complex nonlinear least-squares methods for the deconvolution of dielectric or conducting system frequency-response data are described and applied to synthetic data and to dielectric data of n-pentanol alcohol, water, and glycerol. The first method represents a distribution of relaxation times or transition rates by an inherently discrete function. Its inversion accuracy and resolution power are shown to be limited only by the accuracy of the data when the data themselves arise from a discrete distribution involving an arbitrary number of spectral lines. It is shown that those inversion methods employed here which allow the relaxation times to be free variables are much superior to those where these quantities are fixed. Furthermore, free-τ methods allow unambiguous discrimination between discrete and continuous distributions, even for data with substantial errors. Contrary to previous conclusions, discrete distributions were determined for both n-pentanol alcohol and water. A complex, continuous distribution estimate was obtained for glycerol. Algorithms for all approaches are incorporated in a readily available computer program. Serious problems with some previous dielectric inversion methods are identified. Finally, several possibilities are mentioned that may allow greater inversion resolution to be obtained for complex nonlinear least-squares estimation of continuous distributions from noisy data. © *1995 American Institute of Physics.*

I. INTRODUCTION AND ANALYSIS METHODS

A. General background

Given some small-signal AC complex dielectric constant, frequency-response data, $\epsilon(\omega) = \epsilon'(\omega) - i\epsilon''(\omega)$, it is often useful to estimate the discrete or continuous distribution of relaxation times (DRT) function,[1,2] $G(\tau)$, from the data. Here, $i \equiv \sqrt{-1}$, ω is the angular frequency, and τ is a relaxation time.

Knowledge of the $G(\tau)$ associated with the $\epsilon(\omega)$ data can frequently be helpful in improving one's understanding of the physico-chemical processes occurring in the experimental material during measurement. Such knowledge can be particularly useful in initial stages of an investigation when no detailed response theory is available. In particular, the unambiguous determination of whether a distribution involves a sum of discrete spectral lines or is a continuous function of τ can be particularly valuable. The capability to do so is an important feature of two of the present inversion methods but not of earlier ones.

The estimation of $G(\tau)$ is an inverse problem involving a Fredholm integral equation of the first kind, one which is often ill posed.[3–6(a)] Then, no unique solution is possible and special approaches, such as Tikhonov regularization,[3–6(a)] are often used to obtain an approximate estimate of the distribution. Here, a weighted, nonlinear least-squares inversion method is described and illustrated, one which avoids some of these inversion problems and can lead to unique, accurate estimates under some conditions. Although the present method is applied here only to dielectric data, it is also valuable for conducting system DRT estimation[7–13] and even to inverse-diffusion problems.[14]

The dielectric response $\epsilon(\omega)$ may be expressed in terms of a general, normalized frequency-response function, $I(\omega)$, as[1,2,7]

$$\epsilon(\omega) = \epsilon_\infty + (\epsilon_0 - \epsilon_\infty)I(\omega), \tag{1}$$

where $\epsilon_\infty \equiv \epsilon'(\infty)$ and $\epsilon_0 \equiv \epsilon'(0)$. Therefore, $I(0)=1$ and $I(\infty)=0$.

For the present discrete or continuous DRT situation, we may replace the general $I(\omega)$ function by[1,2]

$$I_G(\omega) \equiv \int_0^\infty [G(\tau)/(1+i\omega\tau)]d\tau, \tag{2}$$

where normalization requires that $I_G(0)=1$. Note that $G(\tau)$ is not dimensionless when the relaxation time, τ, has its usual dimension of time.

Most dielectric and conducting system small-signal AC data extend over a wide frequency range, sometimes as great as 10 decades. It is then appropriate to consider logarithmic frequency variation and a dimensionless logarithmic distribution of relaxation times,[2,7,15] or activation energies.[7,16] Let us introduce into Eq. (2) the dimensionless quantities $y \equiv -\ln(\tau/\tau_0)$ and $x \equiv \ln(\omega/\omega_0)$, with the scaling values τ_0 and ω_0 satisfying the relation $\omega_0\tau_0=1$. Here τ_0 is arbitrary but will usually be taken as $1s$. Now denote the new distribution function as $F(y)$. Note that since conservation of probability requires that $G(\tau)|d\tau| = F(y)|dy|$, we can replace the $I_G(\omega)$ response function in Eq. (2) by

$$I_F(\omega) \equiv \int_{-\infty}^\infty \{F(y)/[1+i\ \exp(x-y)]\}dy, \tag{3}$$

now expressed in convolution form, with $I_F(0)=1$. This condition specifies that the area under the $F(y)$ curve is unity. Physical considerations also require that both $F(y)$ and $G(\tau)$ approach zero at their extremes.

The functions $G(\tau)$ and $F(y)$ are related by $F(y) = \tau G(\tau)$. Unfortunately, the same symbol, $G(\tau)$, has sometimes been used for both of them, tending to confuse the

distinction between them. Another common usage[2] is to denote the logarithmic distribution function as $G[\ln(\tau)]$, but the use of an entirely different symbol for the two different quantities eliminates the possibility of ambiguity and is strongly recommended. Further, it is worth mentioning that expressions such as $\ln(\tau)$, very commonly used in the present field,[2,17] are mathematically inadmissible when τ is not dimensionless. Finally, since experimental frequency-response data are discrete, the frequency variable ω must then be written as ω_n, where $1 \leq n \leq N < \infty$.

B. Numerical inversion approaches

1. Discrete distribution

In this case the unknown DRT is purely discrete and so is made up of M spectral lines. Then one can write

$$G(\tau) = \sum_{m=1}^{M} d_m \delta(\tau - \tau_m), \tag{4}$$

where the d_m's are amplitude coefficients which define the relative strength of each $\{d_m, \tau_m\}$ line. The distribution is then fully defined by the values of the set of parameters. Now define $I_{\text{mod}}(\omega_n)$ as a theoretical or model response. Substitution of Eq. (4) into Eq. (2) leads to the exact result

$$I_G(\omega_n) \equiv I_{\text{mod}}(\omega_n) = \sum_{m=1}^{M} d_m / (1 + i\omega_n \tau_m), \tag{5}$$

a sum of Debye responses which can be readily shown to be independent of whether τ or y is used as the relaxation-time variable. Notice that normalization requires that $I_{\text{mod}}(0) = 1$, a requirement which sets the scale of the d_m parameters.

2. Continuous distribution

The situation is somewhat different if the unknown distribution is continuous. Then one must use numerical quadrature to obtain estimates of $I_{\text{mod}}(\omega_n)$. Such quadrature approximates the distribution by means of discrete points. It is then fully defined by the parameter set $\{c_m, \tau_m\}$, where here the continuous character of the distribution is recognized by replacing the d_m parameters by the c_m ones. It is important to remember that although the $\{d_m, \tau_m\}$ set defines isolated, unconnected points, the $\{c_m, \tau_m\}$ one specifies connected points along the distribution. If quadrature weights are defined as w_m, then a general approximation for the $I_F(\omega)$ result of Eq. (3) may be written as

$$I_{\text{mod}}(\omega_n) = \sum_{m=1}^{M} (w_m c_m) / (1 + i\omega_n \tau_m), \tag{6}$$

a result similar to that of Eq. (5), but where normalization now requires that the sum of the products $w_m c_m$ be unity. Some specific choices for the w_m quantities will be discussed later.

3. Inversion procedure

Most previous inversion methods, whether they use Tikhonov regularization or other procedures,[3–6(a),18–22] start by defining a set of fixed points over the expected range of the distribution, and usually take their spacing constant. For the dielectric area, the usual choice[17,23,24] has been to define a range for τ by setting $\omega_{\text{max}}\tau_{\text{min}} = \omega_{\text{min}}\tau_{\text{max}} = 1$. Then, for equal spacing in the variable $\ln(\tau/\tau_0)$, the step size is just $\ln(\tau_{\text{max}}/\tau_{\text{min}})/(M-1)$, which can be set equal to w_m for all m in the simplest possible quadrature scheme.

The *ab initio* selection of a set of fixed-τ points is particularly inappropriate for discrete distributions where the line spacing is initially unknown and almost always irregular. But, as will be shown later, even for a continuous distribution the choice of fixed points is far inferior to the alternate of taking such points as free parameters of a weighted, nonlinear least-squares inversion procedure. The main difference in the present inversion methods from earlier ones using linear or nonlinear least-squares fitting is the present provision that all the $\{d_m, \tau_m\}$ or $\{c_m, \tau_m\}$ parameters can be free variables of the fitting.

Allowing the τ_m [or $\ln(\tau_m/\tau_0)$] parameters to be free introduces two complications. Fitting is generally carried out with M initially small, then increased to larger and larger values. The successive τ_m values found in this way are not usually in monotonically increasing order, but ordering is needed for an integration routine appropriate when the spacing between $\ln(\tau_m/\tau_0)$ values is not constant. Thus, sorting on τ_m is required before the use of such a routine, but it is of course unnecessary for the discrete-distribution situation.

The actual procedure for the dielectric case is to fit the model predictions

$$\epsilon_{\text{mod}}(\omega_n) = \epsilon_\infty + (\epsilon_0 - \epsilon_\infty)I_{\text{mod}}(\omega_n) \tag{7}$$

to the experimental data, $\epsilon_{\text{meas}}(\omega_n)$, for all n values. Such fitting allows one to obtain least-squares estimates of ϵ_∞, ϵ_0, and the $2M$ distribution parameters, and their estimated standard deviations as well. In the fitting, any of these parameters can be taken free or fixed. Since dielectric frequency response data are usually complex, complex nonlinear least-squares (CNLS) fitting of both real and imaginary parts simultaneously is generally employed in order to use all the available data, but real- or imaginary-part fitting is also possible.

All the present inversions and direct fits were carried out with the readily available LEVM computer program described in Appendix A. The nonlinear least-squares fitting procedure itself involves some regularization effects[4] which help contribute to good inversion estimates. The various fitting choices available in specific CNLS fitting program used herein (LEVM), allow several different approaches to be used. To define the relevant DRT inversion possibilities, let "*D*" stand for discrete, "*C*" for continuous, "*F*" for fixed, and "*V*" for variable (free) τ_m. Then the following combinations are of particular interest: discrete-distribution inversion method involving free, variable τ_m's (DV), continuous-function inversion method involving fixed τ_m's (CF), and continuous-function inversion method involving free, variable τ_m's (CV).

The linear least-squares problem associated with use of Eq. (6) for inversion with constant spacing involves an $N \times M$ "design" matrix, and, when $N = M$, direct inversion

of the $M \times M$ matrix may be employed to obtain c_m values.[6(b)] This approach has been called the matrix-quadrature method of numerical inversion,[5] and was applied to the present problem long ago by Uhlmann and Hakim.[23] But, it is generally unsatisfactory, and its generalization when $M < N$ (for example, the use of linear least squares fitting) is also quite inadequate and usually leads to meaningless oscillatory estimates[5,24] of c_m. Note that when $N = M$, either matrix inversion or, equivalently, linear least squares, should lead to an exact solution, "one which 'models' the discrete data function to within numerical round-off,"[5] but ill conditioning intervenes when, as usual, the condition number of the matrix is much greater than unity.[5] Thus it is important that the number of degrees of freedom in the present inversion procedures always be substantial (see Appendix A).

II. INVERSION RESULTS

A. General considerations

A cardinal point about inversion solutions which yield estimates of continuous distributions, one associated with their usually ill-posed and ill-conditioned character, is that no matter how well a frequency-response model, such as Eq. (7), fits a given set of data, the estimate of the DRT associated with the data is not guaranteed to approximate the true DRT well. In fact, the accuracy of such an estimate does not necessarily improve beyond a certain point as the S_F of the frequency-response fit is decreased by increasing M, the number of free c_m parameters in the fitting model. Here S_F is a measure of goodness of fit at the frequency-response level (see Appendix A). This accuracy limitation does not apply for DV estimation using data arising from an inherently discrete DRT since the data themselves determine the appropriate number of d_m points.

B. Estimation of distributions using synthetic data

Morgan and Lesmes[17] (ML) recently inverted both synthetic and experimental data using a variant of the present CF approach (see Appendix B), and some of their same data sets will be reanalyzed herein. For synthetic data, I consider only their choices of single Debye-function relaxation response and of Cole–Cole (CC) response involving the parameter $\beta = 0.6$. The more general empirical response function suggested by Havriliak and Negami[25] may be written as

$$I_{HN}(\omega) = [1 + (i\omega\tau_{HN})^\beta]^{-\gamma}, \quad (8)$$

where τ_{HN} is not necessarily equal to τ_0, and β and γ fall in the range $[0,1]$. When $\gamma = 1$, this equation reduces to the Cole–Cole form;[26] for $\beta = 1$ it describes Davidson–Cole response;[27] and for $\gamma = \beta = 1$, it degenerates to Debye response. I follow ML here by using the values $\epsilon_\infty = 5$ and $(\epsilon_0 - \epsilon_\infty) = 20$ in generating synthetic data. The synthetic CC data used covered the range from $10^{-3} \leq \omega_n\tau_0 \leq 10^3$, a logarithmic bandwidth of 6, with 49 points equally spaced in log frequency. The $F(y)$ DRT associated with Cole–Cole behavior may be written[2] as

$$F_{CC}(y) \equiv (2\pi)^{-1} \sin(\psi)/[\cosh(\beta y) + \cos(\psi)], \quad (9)$$

where $\psi \equiv \pi\beta$. Since this expression is normalized, its integral over all y is unity.

1. Discrete distribution: a single spectral line

Although discrete-distribution response with only a few lines present is probably rarer than a response associated with the presence of a continuous DRT, Garg and Smyth[28] found that their data on normal primary alcohols could be well described by the sum of three Debye functions. Morgan and Lesmes[17] applied their real-part CF method to single Debye response. The results of their fitting with $M = 19$ were not entirely satisfactory, since they yielded a central $c_m = c_{10}$ estimate of about 1.3 instead of unity, and showed c_{10} dependence on M. Such behavior arises from the use of CF area normalization rather than discrete-distribution normalization.

In contrast, DV inversion yielded essentially exact estimates when applied to the same synthetic data, namely the input values $\epsilon_\infty = 5$, $\epsilon_0 = 25$, $d_1 = 1$, and $\tau = 1$ s. Further, very few iterations were required for convergence, and only this single value of d_m was nonzero, so the appropriate $M = 1$ value was verified.

CNLS fit results of the same data using the LEVM CF inversion method were instructive when carried out with DV rather than CF normalization. The calculations involved double-precision arithmetic with stringent stopping/convergence criteria and required a large number of iterations for full convergence. With $M = 11$, for example, they led to essentially exact estimates for ϵ_∞ and ϵ_0 as before, to an S_F value of about 10^{-9}, to a value of the central c_6 of 1 ± 10^{-10}, and to positive estimates of all the other $10c_m$'s of less than 10^{-10}. However, as in the ML calculations,[17] it was necessary to ensure that one of the fixed input τ_m values coincided with the τ value of the single line present, an important weakness of the CF method applied to data involving a discrete distribution.

Incidentally, real-part CF fitting yielded essentially the same results as complex fitting, as also did imaginary-part fitting, although the latter does not allow an estimate of ϵ_∞ to be obtained. To illustrate the results of a poor choice of the fixed τ_m values with CF fitting, exact data such as those above were produced for $M = 11$ with τ_p, the value of τ for the single line present, placed halfway between τ_6 and τ_7. Fitting then yielded $S_F \simeq 0.12$, $(\epsilon_0 - \epsilon_\infty) \simeq 18.0$, $\epsilon_\infty \simeq 4.98$, and only two significant c_m values: $c_6 \simeq 0.486$ and $c_7 \simeq 0.514$.

2. Discrete distribution: multiple lines

Over the years there has been considerable effort expended to find ways of resolving the sum of two or more simple Debye responses with relaxation frequencies relatively close together,[19,20,29–32] but no satisfactory approach has been demonstrated. For example, when Fourier deconvolution methods are used,[5,19–22] there is a problem of selecting an appropriate window for high-frequency filtering, and when such methods are applied to Debye spectra they yield limited resolution with spectral responses of finite width and usually with oscillating tails as well. Further, Fourier trans-

formation of discrete data leads to somewhat different results than does continuous-function transformation, as analyzed elsewhere.[33]

In order to demonstrate the great resolution power of the DV, the inversion of data arising from two very closely spaced lines is considered here. In the first case, the following values were used in Eq. (A1) to generate the data: $C_1 = 1$ F, $R_1 = 1$ Ω, $C_2 = 0.01$ F, and $R_2 = 90$ Ω. Thus, the peak height of one dispersion is 100 times smaller than the other, and the ratio of their time constants is 1 to 0.9, very close spacing. DV inversion of the original exact frequency-response yielded essentially the exact input values, including the time constants $\tau_1 = 1$ s and $\tau_2 = 0.9$ s. When the data were rounded to four significant figures, however, the fit yielded $S_F \simeq 2 \times 10^{-4}$, and estimates of 1.003, 0.997, 0.0069, and 129, for the above four circuit elements, respectively. The corresponding τ_m values were thus about 1 and 0.89 s, showing that τ_2 was better estimated than its individual components. Considerably less-accurate estimates were obtained with only three digits remaining in the data.

For a second test, C_1 and R_1 had the same values as above, but C_2 was taken as 0.3 F and R_2 as 2 Ω. Thus, for this case the peak height of the second dispersion was about a third of that of the first, and the ratio of τ's was 1 to 0.6. With three (and two) significant figures, the fit estimates for S_F, C_1, R_1, C_2, and R_2 were, respectively, 0.0018 (0.014), 1.001|0.011 (0.967|0.090), 0.998|0.014 (1.047|0.116), 0.299|0.037 (0.335|0.258), and 2.005|0.029 (1.824|0.202). Here a number like |0.037 is the relative standard deviation estimate of the associated parameter as determined from the fit (see Appendix A). Again, the τ_1 and τ_2 estimates are better than their individual components, and most of the relative standard deviation estimates are appreciably greater than the actual errors of estimation. These results show that although the accuracy is strongly degraded on going from three to two digits in the data, the estimates are still quite significant.

The present inversions indicate that DV fits of noisy, discrete-distribution data can yield much greater resolution than has been demonstrated before. Furthermore, discrete-distribution inversion is not constrained by the limitations of an ill-posed problem which usually apply for continuous-function DRT estimation.

3. Continuous Cole–Cole distribution

Figure 1 shows the results of two complex-fit, proportional-weighting CV inversions of exact CC data with $\beta = 0.6$. Morgan and Lesmes[17] treated this same situation with their CF method but presented their results as low-resolution linear plots of c_m [called $y(\tau)$ by them] vs the ordinary logarithmic variable, $s_m \equiv \log_{10}(\tau_m/\tau_0)$. This variable, rather than y_m, is also employed here for plotting. Better comparison between theory and predictions, particularly in the tails of the distribution, is possible when a logarithmic rather than a linear c_m scale is used, as in Fig. 1.

Note that when an $M = 19$ inversion point, denoted by an open circle in Fig. 1, evenly surrounds its exact value, denoted by an asterisk, the inversion estimate is very good indeed, as it is for the present $M = 19$ inversion estimate. Furthermore, comparison of the $M = 13$ and $M = 19$ results
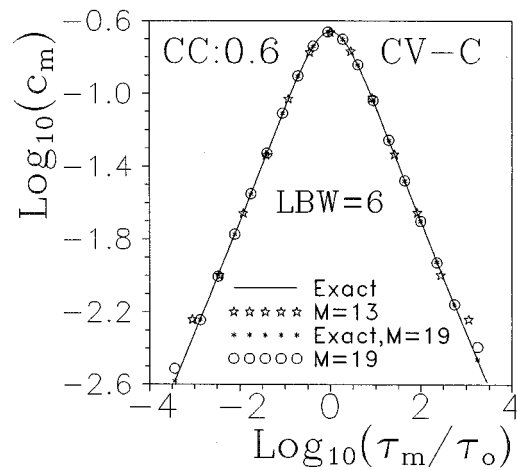

FIG. 1. Log–log plots comparing the exact Cole–Cole (CC) continuous distribution function of Eq. (9) (solid line) with CV inversion results obtained using the complex nonlinear least-squares fitting program, LEVM, with $M = 13$ and 19 c_m points. Here the combination "−C" denotes complex data fitting, and τ_0 is a scaling relaxation time taken as 1 s here and elsewhere. The points denoted by asterisks are exact distribution values plotted at the same values of $s_m \equiv \log_{10}(\tau_m/\tau_0)$ as those found from the $M = 19$ inversion. Here LBW is the logarithmic bandwidth, $\log_{10}(\omega_{max}/\omega_{min})$.

unambiguously demonstrates that the data analyzed involve a continuous, not a discrete DRT. The s_m values of any discrete lines present in the distribution would remain nearly independent of the value of M or of the span of the frequency-response data, but we see that they do not do so here. The same variable-τ_m behavior was found with DV inversion of these data, so either approach can be used to identify distribution type. Such a capability is absent from all previous approaches which use fixed points.

Although log–log plots are useful to show behavior over a wide range, errors are better presented with a linear scale, as in Fig. 2. Here complex-fit CF and CV inversion errors are compared. Note the approximately 10-times greater errors of the CF inversion as compared to the CV ones. The strong
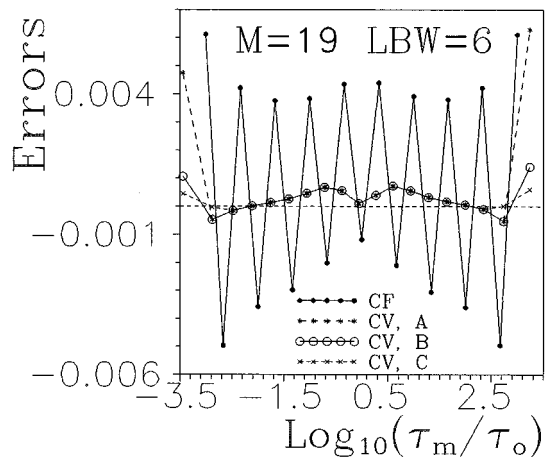

FIG. 2. CF and CV inversion errors, $[c_m - F_{CC}(s_m)]$, for the CC data of Fig. 1 vs $\log_{10}(\tau_m/\tau_0)$ for $M = 19$. The A, B, and C CV inversion results were obtained with different end-point corrections, as discussed in the text. The B curve corresponds to the $M = 19$ results of Fig. 1.

TABLE I. Inversion results for water at 20 °C; $\tau_0 = 1$ s. Here quantities written as $A|B$ indicate an estimate, A, and its estimated relative standard deviation, B.

Bottreau et al. (Ref. 32) DV		Morgan and Lesmes (Ref. 17) CF		This study DV			
$\tau_m|\tau_0$	d_m	$\tau_p|\tau_0$	c_p	$\tau_m|\tau_0$	d_m	$\tau_m|\tau_0$	d_m
$4.63\times10^{-14}|0.002$	0.036			$7.17\times10^{-14}|0.19$	$0.037|0.09$	$3.37\times10^{-14}|0.45$	$0.026|0.27$
		3.09×10^{-13}	0.014			$2.93\times10^{-13}|0.70$	$0.019|0.39$
$8.92\times10^{-12}|0.017$	0.914	9.55×10^{-12}	3.04	$9.39\times10^{-12}|0.005$	$0.963|0.003$	$9.44\times10^{-12}|0.006$	$0.955|0.003$
$2.86\times10^{-11}|0.09$	0.051						
ϵ_0 80.4		80.9		80.08		80.15	
ϵ_∞ 1.78		3.30		$2.26|0.10$		$1.94|0.16$	
S_F				0.332		0.320	

error oscillations of the CF curve are characteristic of such inversions of data involving continuous distributions.

All three of the CV inversions of Fig. 2 used a type of extended trapezoidal quadrature with variable spacing[34] but with different end-point corrections. Curve A involved no such corrections and therefore shows appreciable errors in the two end points which arise from the limited range of the data. Because the numerical integration involves a finite range of s_m, area outside this range is ignored. Curve B partly corrects for these omissions by changing the factor of 0.5 used in the integration only for the end-points to unity. Finally, curve C, the best of the three, corrects the last two points on each end by factors appearing in an extended quadrature formula[6(c)] of order $1/M^3$. Since there is no evident theoretical rationale for this formula, the curve B approach is taken standard in the LEVM CV procedure.

Incidentally, although spline-fitting and quadrature[6(d)] could be used in place of the present quadrature method, this approach was found to yield somewhat larger errors than did the curve B method when the spline fitting involved natural boundary conditions, and the use of more appropriate estimated boundary conditions did not seem to justify the resulting complications. Although the curve B errors are remarkably small, their existence, even for exact input data, is evidence of ill conditioning which must be expected for the inversion of discrete data associated with a continuous distribution.

With $M=19$, the fit of the original synthetic frequency-response data was very good, with an S_F value of about 10^{-5}. This quantity has been found to decrease for CV fitting as $\exp(-aQ)$, where a is a constant which depends on the form of the convolution kernel; $Q \equiv M/(\text{LBW})$ (LBW is the logarithmic bandwidth of the original exact data). For such data, S_F continues to decrease as M is increased until finally limited by the number of decimal digits carried in the data and calculations. There is thus always a limit on the resolution the present CV method can deliver. This limit is usually unimportant for exact data but can become crucial for experimental data. Then the minimum S_F, and so the maximum useful M value, is limited by the error in the data. This limitation is demonstrated for the data analyzed in Sec. II C 3.

C. Estimation and identification of distributions from experimental data

1. Inversion of water data

Morgan and Lesmes[17] used their real-part CF method to invert 20 °C water data compiled by Boutreau et al.[32] from a variety of sources. These 27 data values extended over the wide range from about 6×10^8 Hz to over 10^{13} Hz and are quite irregular, having appreciable lacunae and some large outliers. To allow direct comparison with earlier results, these same data were used for the present inversions, in spite of their irregularities. Before trying to invert these data, they were fitted, using LEVM, to the CC form of Eq. (8). It was found that an adequate fit could not be obtained by using the 24 lowest-frequency original data values, those employed by ML. Therefore, generalized cross validatory quintic spline smoothing[35] was carried out on the full 27 points listed in Ref. 32. Of these, the last three points were extrapolations from infrared data. This procedure automatically provides statistically appropriate smoothing when carried out on logarithmically transformed data if the original data have predominantly proportional errors.[10] The smoothed data are presented in Fig. 3 but still show appreciable irregularity.

An adequate $N=24$ CC fit was then found for the smoothed data. It yielded $\tau_{CC}/\tau_0 \approx 9.33\times10^{-12}|0.006$, $\beta \approx 0.984|0.004$, $(\epsilon_0-\epsilon_\infty) \approx 76.34|0.004$, $\epsilon_\infty \approx 4.28|0.051$, and $S_F = 0.356$. Most of these results agree closely with those
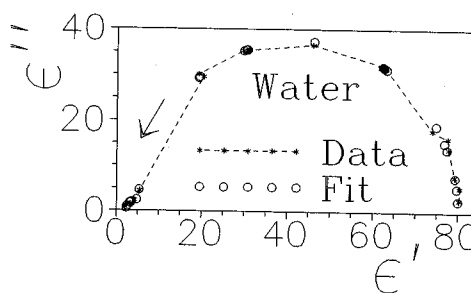
FIG. 3. Complex-plane plot of the smoothed 20 °C water dielectric data obtained by many workers, and of the $M=2$ DV inversion fitting points for these data. The arrow shows the direction of increasing frequency. The dominant discrete relaxation time obtained from thie inversion of these data was 9.39 ps. See Table I for further results.

quoted by Hasted,[36] who concluded that a CC continuous distribution fit the water data somewhat better than did a single-line discrete distribution.

Results of different analyses are summarized in Table I. The $M=2$ DV inversion points are also plotted in Fig. 3. The most plausible estimates shown in Table I seem to be the two DV ones on the right, calculated with $N=27$. Of these, one should choose that with two isolated spectral lines, since the parameter uncertainties are smaller than those with $M=3$. Fitting without the three extrapolated points yielded only a single line with a τ very close to the larger τ one, showing that the second small-τ line in Table I was associated with these points.

When three DV lines are extracted with $N=27$, it is evident that two of them are then not well defined, especially the one with $\tau_m/\tau_0 \simeq 2.93 \times 10^{-13}$. When four lines are estimated, it is found that the relative standard deviations of five of the nine free parameters are of the order of or greater than unity, meaningless estimates. As shown in Table I, the central line of the three-line discrete-function-approach results of the present study is clearly close to the first of the two peaks of the continuous distribution estimated by ML using their CF approach, one which used -1 degrees of freedom rather than an appreciable positive value, as always employed with LEVM fits. Here c_p denotes the value of c_m at a peak of a continuous distribution. Since the central DV line is itself not well defined, little credence should be given to either it or the ML estimate. Difficulties to be expected with underdetermined fits are discussed in Appendix B.

Comparison of the results of Bottreau *et al.*[32] with the present ones is also instructive. These authors used a discrete-function inversion approach limited to $M \leq 4$, one which is less accurate and much less general than the present independently developed one. As one might expect, their results and the present DV ones are reasonably similar, but the present DV method yields more resolution and accuracy and is far easier to apply. When it was used to estimate four spectral lines, one with $\tau_m/\tau_0 \simeq 1.2 \times 10^{-11}$ appeared, roughly corresponding to one of the lines listed at the left of Table I. The relative standard deviation associated with this value was about 0.36 and the relative standard deviation estimate for the corresponding d_m was about 2, not a well-defined line.

Because the DV approach yields adequate results with just two spectral lines, it is not sensible to apply the CF method to these data, which are clearly better and more appropriately defined by a few isolated, discrete lines than by a continuous DRT. Even though the data are of poor quality and consistency, the combination of least-squares-fit smoothing and DV inversion yields reasonable inversion parameters, but more-accurate data, extending over a very wide frequency range for several different temperatures, are desirable for such an important material as water.

2. Inversion of n-pentanol alcohol data

Morgan and Lesmes sent me the $N=30$ real-part data they used[17] to estimate a continuous DRT for n-pentanol alcohol.[37] These 25 °C data, shown in Fig. 4, were produced, according to ML, by digitizing the graphical $\epsilon'(\omega)$ curve for
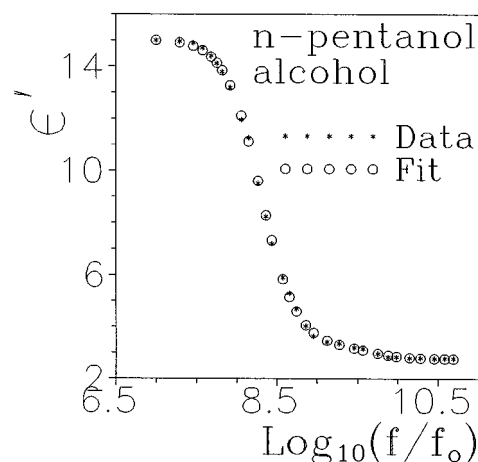

FIG. 4. The 25 °C n-pentanol alcohol ϵ' frequency-response data, and the $M=2$ DV inversion fitting points for these data. Here $f_0=1$ Hz. The dominant discrete relaxation time obtained from the inversion of these data was 0.802 ns. See Table II for further results.

this material appearing in Ref. 20; thus they may be expected to contain errors arising from this process as well as some of their original experimental errors.

LEVM has been used to fit these frequency-response data to a wide variety of continuous-distribution response functions. The best fit was found for the Davidson–Cole response function[27] form of Eq. (8). Fitting results with a smoothed and interpolated $N=74$ set of data were $\tau_{CD} \simeq 9.54 \times 10^{-10} | 0.022$ s, $\gamma \simeq 0.862 | 0.011$, $\epsilon_\infty \simeq 2.787 | 0.006$, $(\epsilon_0 - \epsilon_\infty) \simeq 12.374 | 0.007$, and $S_F \simeq 0.0207$. This fit is significantly worse than the DV one in Table II.

Table II compares DV inversion results to two continuous-function inversions obtained by earlier workers. Those of Ref. 20 were obtained by a Fourier-transform deconvolution method involving imaginary part data only, one which leads to broadening of any delta-function spectral lines present. Their τ values quoted in Table II are those defining the peaks of their continuous distribution estimate, omitting a small response for which τ_0/τ_m is well above the measured $\omega_{max}\tau_0$. Their peak coefficients have been normalized[17] such that the amplitude of the major peak was set to 1.000.

The ML CF results[17] were obtained from real-part fitting with $N=M=30$, again yielding -1 for the available number of degrees of freedom. Thus, one might expect very poor estimates. We see that their ϵ_0 estimate is, in fact, appreciably larger than the lowest-frequency limiting data value of 15 and the DV estimates of ϵ_0 presented in Table II. More significantly, their CF inversion led to a smooth, apparently continuous, distribution curve with two peaks, very different results, except for the location of the peaks, than the predictions of the DV inversions discussed below.

The DV results listed in Table II used the original 30 data values obtained from ML, and only two significant spectral lines could be obtained from the inversion of these data. When I attempted to extract three lines, the relative standard deviation estimates of one of the τ_m/τ_0's and two of the three d_m's were of the order of unity or greater. Had the

TABLE II. Inversion results for n-pentanol alcohol at 25 °C; $\tau_0 = 1$ s.

Salefran and Dutoit (Ref. 20)		Morgan and Lesmes (Ref. 17) CF		This study DV	
τ_p/τ_0	c_p	τ_p/τ_0	c_p	τ_m/τ_0	d_m
3.6×10^{-11}	0.056	3.43×10^{-11}	0.026	3.48×10^{-11}|0.16	0.04|0.08
7.6×10^{-10}	1.000	8.55×10^{-10}	2.5	8.02×10^{-10}|0.014	0.96|0.007
ϵ_0		15.4		15.03	
ϵ_∞		2.71		2.76|0.007	
S_F				0.0143	

$\epsilon(\omega)$ data actually been associated with a continuous distribution of some width, such as those estimated from the n-pentanol data by Salefran and Dutoit[20] and by ML,[17] it should have been possible to obtain many significant nonzero d_m or c_m estimates. Since this was not the case, it seems virtually certain that the true distribution is discrete, and thus the present DV procedure should yield good estimates for the DRT coefficients.

Although the present results clearly show that the present data are associated with a discrete DRT rather than with a continuous one, it is of interest to compare real-part CF results with the corresponding results of ML. The present CF inversion process with $M = 10$ was found to yield only a few nonzero c_m results which matched the discrete lines of Table II as closely as possible within the limitations set by equally spaced τ_m/τ_0 values. All other c_m estimates were effectively driven to zero, further indication that for these data a discrete DRT is appropriate and that a continuous one is not. By contrast, with a negative value for the available degrees of freedom, none of ML's c_m estimates were less than about 3×10^{-3}.

3. Inversion of glycerol data

Since only discrete DRT's were found for the two liquids considered earlier, it seemed worthwhile to analyze one with a continuous distribution. Professor Sydney Nagel kindly sent me dielectric response data on salol and glycerol,[37] materials of considerable current interest because of their usefulness in allowing the verification of an important new scaling relationship for supercooled liquids.[38] The salol data exhibited much more irregularity in their high-frequency tail regions than did the glycerol, so the latter have been used for the inversion results shown here. Within the limitations of the data, a similar response was found for the salol data sets investigated.

First, it was found that all the data sets examined could be quite well fitted directly with continuous-distribution response functions represented by the following three circuit elements all in parallel at the admittance level: a capacitance accounting for ϵ_∞, a fractional exponential response function[39] dominating the low-frequency response, and a Havriliak–Negami response function[25] important at the high frequency end of the data. For glycerol at 230 K, the fractional exponent was found to be about 0.72, and the Havriliak–Negami function degenerated to a Davidson–Cole response with $\gamma \approx 0.30$. The S_F of the proportional weighting LEVM complex data fit was about 0.022.

Figure 5 shows some DRT estimates for these $N = 41$ glycerol data. First, it is evident from the differences in the $M = 9$ and 10 τ_m values that the response involves a continuous distribution rather than a discrete one. For $M = 10$, the S_F of the fit was about 0.0094, and the estimated relative standard deviations of the c_m's were all below 0.1 and mostly below 0.06. When a fit was carried out with $M = 11$, the value of S_F was only slightly reduced, and the relative standard deviations of six of the 11 c_m's were then above 0.1, with one about 0.34. The leveling of S_F, the increase in the estimated uncertainties of the c_m, and a concomitant increase in those of the τ_m's all indicate that the limit of CV resolution determined by the present data error level has been reached with $M \approx 10$ (see the discussion at the end of Sec. II B 3).

Figure 5 also shows fractional exponential DRT response. It was produced by first fitting the original $\epsilon''(\omega)$ data with a capacitance and fractional-response function in parallel, but using unity weighting to emphasize the peak region of the response. This part of the response was well fitted by this procedure and the resulting fit parameters were used to generate exact, complex-response synthetic data covering the same frequency region with 87 points equally spaced in $\log_{10}(\omega/\omega_0)$. These data were then used to obtain the two CV inversion responses identified by WW in the figure.

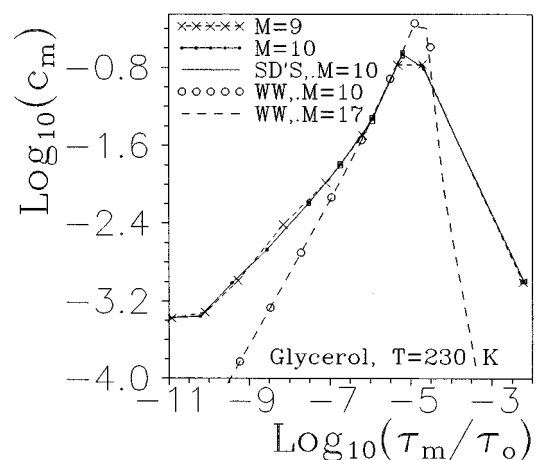Figure 5 shows that there is a constant-slope region with

FIG. 5. CV inversion results for glycerol at 230 K for different values of M. Calculation of the fractional-exponential DRT responses, marked WW for William–Watts, is described in the text. The small, solid line, uncertainty rectangles surrounding each $M = 10$ $\{c_m, \tau_m\}$ point were formed by adding and subtracting one estimated standard deviation of each quantity from the point, then transforming to log–log space and connecting the results with straight lines.

a log–log slope of about 0.72 just to the left of the peak, a slope which is shared by the two glycerol inversion responses and by the fractional-exponential ones, as one might expect from the aforementioned discussion. The presence of a high-frequency tail is evident, however, in the glycerol response below $s_m \approx -7$. Except for the end point, which is likely to be inaccurate because of the finite span of the data, another nearly straight-line region is apparent, but it has a log–log slope of nearer 0.45 than 0.3. To verify the reasonableness of this result, the direct-fit results mentioned earlier involving three response elements were used to generate new $N=87$ exact data and then this data set was fitted by proportional weighting CV inversion. The resulting $M=10$ curve was found to lie almost exactly over the $M=10$ glycerol curve of the figure from the left side of the peak to the minimum s_m value, except for the end point. These results thus show that both the direct and the inverse fitting results are mutually consistent over most of the range of the data. Furthermore, they demonstrate the usefulness of CV inversion of experimental data of the present type in bringing out details of a complicated, distributed DRT.

III. CONCLUSIONS AND FUTURE DIRECTIONS

It seems likely that the identification by ML and others of the DRT's of water and n-pentanol alcohol as continuous rather than discrete arose from the lack of a general inversion method that allowed the τ_m's to be free rather than initially fixed. The present DV and CV methods do so and can thus remove the chance of misidentification. Moreover, ML may not have used sufficiently stringent convergence criteria for their least-squares fits, and in their inversions of data for these liquids they applied their usual least-squares procedure even when their number of degrees of freedom was -1.

Although it has been demonstrated herein and elsewhere[14] that DV inversion is not an ill-posed problem and has tremendous resolving power, the maximum resolution of CV inversion is limited by errors in the data. Preliminary work indicates that resolution can be somewhat improved for experimental data inversion by first smoothing the data by the method used herein for the water data, but the gain achieved is usually relatively small. An alternative is to modify the present nonlinear-least-squares CV algorithm to include Tikhonov regularization. Up to the present, such regularization has generally involved fixed independent variable (here τ) values. In order to avoid the need for a subjective choice of the regularization parameter, in the combined approach it might be taken as a free parameter of the fitting or possibly determined by the L curve method of Hansen.[40,41] If nonlinear least-squares inversion, including regularization and free τ_m parameters, proves successful, much larger useful values of M than now practical should be possible, even for data with appreciable errors.

ACKNOWLEDGMENTS

they used in their work and for providing some helpful comments on an earlier version of the present work. I also appreciate the useful suggestions of two reviewers.

APPENDIX A: THE LEVM DISCRETE FUNCTION INVERSION METHOD

A powerful instantiation of the DV discrete function inversion approach has been included for a decade as one of a great many choices in the very general LEVM complex nonlinear least-squares impedance spectroscopy data fitting program.[9,42–45] This program uses the modified Levenberg–Marquardt algorithm of Moré,[46] which is robust, fast converging, and involves implicitly scaled variables. LEVM has been used for direct impedance-spectroscopy data fitting, for estimating DRT's, and as a powerful substitute for Kronig–Kramers transformation.[12,13] Nearly all the present fits with LEVM used proportional or function proportional weighting, and fitting to full complex data or to their real or imaginary parts is straightforward. Details of the many weighting possibilities available in LEVM are given elsewhere.[9,45,47] LEVM fitting yields a value of the standard deviation of the weighted fit residuals, S_F (which involves the number of degrees of freedom available), estimates of the free parameter values, estimates of the standard deviations of the parameter estimates, a parameter correlation matrix, and an F test of the overall fit.

The actual discrete function approach circuit used in LEVM to represent dielectric response response involves a capacitance C_∞ in parallel with M parallel branches, each consisting of a capacitance, C_m, in series with a resistance, R_m. When the admittance of this structure is divided by $i\omega$, one obtains the complex capacitance

$$C(\omega_n) = C_\infty + \Lambda \sum_{m=1}^{M} C_m / (1 + i\omega_n C_m R_m), \qquad (A1)$$

with $\Lambda \equiv 1$. The capacitances appearing in Eq. (A1) can be converted to dielectric constant form by dividing each one by C_V, the capacitance of the empty measurement cell. Weighted CNLS fitting of this circuit yields estimates of the $2M+1$ free parameters. Note that the individual relaxation times are defined as just $\tau_m \equiv C_m R_m$, and the program includes a choice which leads to estimates of C_m (and thus d_m) and τ_m directly if desired. The number of degrees of freedom for discrete or continuous function fitting is $(2N-2M-1)$ for complex data fitting, $(N-2M-1)$ for real-part fitting, and $(N-2M)$ for imaginary part data.

After the iterative fitting process has converged, the program calculates the quantity $\Delta C \equiv (C_0 - C_\infty)$ (equal to the sum of the C_m's), the τ_m's, and the dimensionless distribution strengths, $d_m \equiv C_m/(C_0 - C_\infty)$. The results may then be expressed in just the form of Eq. (1) with $I(\omega)$ given by the I_G of Eq. (5).

When LEVM is used for discrete function inversion, it not only returns estimates of the standard deviations of the free distribution parameters but also those of the relative standard deviations of the parameters, where for a parameter estimate B, the relative standard deviation is just the standard deviation divided by $|B|$. Unfortunately, such estimates have usu-

ally been unavailable and/or unreported in the past. Incidentally, a relative standard deviation estimate as large as 0.4 or 0.5 provides only a small degree of confidence in the related parameter estimate, and one of the order of unity or greater suggests that the related parameter estimate is meaningless unless it itself is zero.

Although estimated relative standard deviations can be calculated for the d_m coefficients and for $(C_0 - C_\infty)$ when those of the C_m's and their correlation matrix are known, the calculation is only approximate. It involves not only a quotient of uncertain quantities, but also possibly very uncertain correlation coefficient values. These values are obtained by a linearization procedure from the off-diagonal elements of the variance-covariance matrix, which is calculated after convergence of the fit. A Monte Carlo study[34] has shown that in calculations of the present type, the off-diagonal elements are generally much less accurate than the diagonal elements, the ones which lead to the parameter relative standard deviation estimates. Thus, this approach is not used in LEVM.

Luckily, it turns out that the *relative* standard deviation estimates of the distribution coefficients are closely the same as those of their associated parameters, and so they need not be separately estimated. This result has been established by taking the Λ of Eq. (A1) fixed at the value of $(C_0 - C_\infty)$ obtained from a prior fit with Λ fixed at unity, and repeating the fitting. Then, the resulting dimensionless C_m values are the direct distribution coefficients themselves and their relative standard deviation values are found to be the same as those with $\Lambda = 1$.

It should be clear that only minor changes in the LEVM discrete function approach program are needed to convert it to a CV or CF continuous-function approach. Because of the parameter optimization element in the present inversion methods, a variable-τ, continuous-function solution differs from a corresponding discrete function one by more than a scale factor. For all of the different kinds of inversions discussed herein, LEVM incorporates the physically required constraint that distribution-strength coefficients cannot be negative.[24]

APPENDIX B: THE MORGAN–LESMES CONTINUOUS FUNCTION INVERSION METHOD

Morgan and Lesmes[17] used nonlinear regression for their real-part CF inversions. They employed a modified Levenberg–Marquardt algorithm which, upon convergence, should yield results very close to those obtained from LEVM CF real-part fitting using function-proportional weighting, a weighting which is determined by taking the uncertainty of each data point equal to the model estimate for that point.[42,43,47] One significant difference between the present inversion approaches and that of ML is that they used what they termed a "data rms error," rms$_D$, which involved $N-1$ rather than the standard degrees-of-freedom choice mentioned in Appendix A for S_F. Furthermore, rms$_D$ should not generally be defined as an *error* measure.[17] There are unknown errors present in experimental data, and thus such quantities as rms$_D$ and S_F are (different) measures of only the sum of the weighted, squared differences between the data values and the model predictions. Only when the model

is known to describe error-free data exactly (information never available for experimental data) do these quantities involve the true weighted errors, as opposed to the weighted residuals.

Some of the ML inversions involved underdetermined systems of equations, those where $M > N$. The problem of obtaining solutions under these conditions is known in econometrics as the undersized sample problem.[6(e),48] These authors suggest that one should not expect to obtain a unique solution; they mention such techniques as singular value decomposition and the repeated use of the generalized matrix inverse; but they do not suggest that it is reasonable to use the same least-squares approach appropriate for an overdetermined system for an undetermined one, as ML have done.

In the ML work,[17] no notational distinction is made between the DRT's $G(\tau)$ and $F(y)$, the right side of Eq. (10) needs to be inverted; the undefined quantity "N" is the number of discrete frequency values, S_m in Eqs. (15) and (16) should be s_m, and the term $\sin(\pi\tau)$ in Eq. (14) should be replaced by $\sin(\pi\beta)$.

[1] J. R. Macdonald and M. K. Brachman, Rev. Mod. Phys. **28**, 393 (1956).
[2] C. J. F. Böttcher and P. Bordewijk, *Theory of Electric Polarization*, (Elsevier Scientific Publishing Company, Amsterdam, The Netherlands, 1978), Vol. II pp. 47 and 48 and 88–92.
[3] F. R. de Hoog, in *The Application and Numerical Solution of Integral Equations*, edited by R. S. Anderssen, F. R. de Hoog, and M. A. Lukas (Suthoff and Noordhoff, Alphen aan den Rijn, The Netherlands, 1980), pp. 119–124.
[4] J. V. Beck, B. Blackwell, and C. R. St. Clair, Jr., *Inverse Heat Conduction* (Wiley-Interscience, New York, 1985), pp. 108–112 and 134–137.
[5] I. J. D. Craig and J. C. Brown, *Inverse Problems in Astronomy* (Adam Hilger, Bristol, England, 1986), pp. 59 and 60.
[6] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes*, 2nd Ed. (Cambridge University Press, Cambridge, England, 1992); (a) pp. 779–817; (b) p. 666; (c) p. 132; (d) p. 134; (e) pp. 57 and 670–672.
[7] J. R. Macdonald, J. Appl. Phys. **58**, 1955 (1985).
[8] J. R. Macdonald, J. Appl. Phys. **61**, 700 (1987).
[9] J. R. Macdonald, J. Appl. Phys. **65**, 4845 (1989).
[10] J. R. Macdonald, J. Appl. Phys. **75**, 1059 (1994).
[11] J. R. Macdonald, Phys. Rev. B **49**, 9438 (1994).
[12] B. A. Boukamp and J. R. Macdonald, Solid State Ionics **74**, 85 (1994).
[13] J. R. Macdonald, J. Electroanal. Chem. **378**, 17 (1994).
[14] J. R. Macdonald, Comp. Phys. (submitted).
[15] K. W. Wagner, Ann. Phys. **40**, 817 (1913).
[16] J. R. Macdonald and J. C. Wang, Solid State Ionics **60**, 319 (1993).
[17] F. D. Morgan and D. P. Lesmes, J. Chem. Phys. **100**, 671 (1994).
[18] P. C. Hansen, Inv. Probl. **10**, 894 (1994).
[19] J. L. Salefran and A. M. Bottreau, J. Chem. Phys. **67**, 1909 (1977).
[20] J. L. Salefran and Y. Dutuit, J. Chem. Phys. **74**, 3056 (1981).
[21] A. D. Franklin and H. J. de Bruin, Phys. Status Solidi A **75**, 647 (1983).
[22] N. Link, S. Bauer, and B. Ploss, J. Appl. Phys. **69**, 2759 (1991).
[23] D. R. Uhlmann and R. M. Hakim, J. Phys. Chem. Solids **32**, 2652 (1971).
[24] J. R. Cost, J. Appl. Phys. **54**, 2137 (1983).
[25] S. Havriliak and S. Negami, Polmer **8**, 161 (1967).
[26] K. S. Cole and R. H. Cole, J. Chem. Phys. **9**, 341 (1941).
[27] D. W. Davidson and R. H. Cole, J. Chem. Phys. **19**, 1484 (1951).
[28] S. K. Garg and C. P. Smyth, J. Phys. Chem. **69**, 1294 (1965).
[29] D. J. Misell and R. J. Sheppard, J. Phys. D **6**, 379 (1973).
[30] R. J. Sheppard and E. H. Grant, Adv. Mol. Relaxation **6**, 61 (1974).
[31] D. J. Misell and R. J. Sheppard, Adv. Mol. Relaxation **6**, 319 (1975).
[32] A. M. Bottreau, J. M. Moreau, J. M. Laurent, and C. Marzat, J. Chem. Phys. **62**, 360 (1975).
[33] W. J. Thompson and J. R. Macdonald, Proc. Natl. Acad. Sci. USA **90**, 6904 (1993).
[34] J. R. Macdonald, Electrochim. Acta **38**, 1883 (1993).
[35] H. J. Woltring, Adv. Eng. Software **8**, 104 (1986).

[36] J. B. Hasted, *Aqueous Dielectrics* (Chapman & Hall, London, 1973), pp. 40–49.

[37] Authors, private communication.

[38] P. K. Dixon, L. Wu, S. R. Nagel, B. D. Williams, and J. P. Carini, Phys. Rev. Lett. **63**, 1108 (1990).

[39] R. Kohlrausch, Pogg. Ann. Phys. **91**, 179 (1854); G. Williams and D. C. Watts, Trans. Faraday Soc. **66**, 80 (1970).

[40] P. C. Hansen, Inv. Prob. **8**, 849 (1992).

[41] P. C. Hansen, Siam Rev. **34**, 561 (1993).

[42] *Impedance Spectroscopy - Emphasizing Solid Materials and Systems*, edited by J. R. Macdonald (Wiley-Interscience, New York, 1987).

[43] J. R. Macdonald and L. D. Potter, Jr., Solid State Ionics **23**, 61 (1987).

[44] J. R. Macdonald, Solid State Ionics **58**, 97 (1992).

[45] J. R. Macdonald, Electrochim. Acta **35**, 1483 (1990).

[46] J. J. Moré, in *Numerical Analysis*, Vol. 630 of *Lecture Notes in Mathematics* (Springer, New York, 1978), p. 105.

[47] J. R. Macdonald and W. J. Thompson, Commun. Stat. Sim. Comput. **20**, 843 (1991).

[48] H. D. Vinod and A. Ullah, *Recent Advances in Regression Methods* (Marcel-Dekker, New York, 1981).