IAEA
International Atomic Energy Agency

INDC(NDS)-0598
Distr. G+NM

INDC International Nuclear Data Committee

Missing Levels with Two Superimposed Sequences

J.F. Shriner, Jr.

Tennessee Technological University

and

G.E. Mitchell

North Carolina State University and
Triangle Universities Nuclear Laboratory

October 2011

IAEA Nuclear Data Section, Vienna International Centre, A-1400 Vienna, Austria

Selected INDC documents may be downloaded in electronic form from
http://www-nds.iaea.org/reports-new/indc-reports/
or sent as an e-mail attachment.
Requests for hardcopy or e-mail transmittal should be directed to *services@iaeand.iaea.org*
or to:
Nuclear Data Section
International Atomic Energy Agency
Vienna International Centre
PO Box 100
A-1400 Vienna
Austria

Printed by the IAEA in Austria

October 2011

INDC(NDS)-0598
Distr. G+MN

Missing Levels with Two Superimposed Sequences

J.F. Shriner, Jr.

Tennessee Technological University

and

G.E. Mitchell

North Carolina State University and
Triangle Universities Nuclear Laboratory

ABSTRACT

Neutron, and to a lesser extent proton, resonance data play an important role in the determination of nuclear level densities. Those densities in turn have a variety of applications in both, pure and applied physics. A key factor to consider when using resonance data to determine level densities is the possibility of missing levels. Previously, we have studied how predictions of the Gaussian Orthogonal Ensemble (GOE) version of Random Matrix Theory (RMT) can be used to estimate the fraction of missing levels for a sequence of levels, a group of levels which share the same quantum numbers. Here we look at a related problem, how to estimate the number of missing levels when the data consist of two superimposed (and perhaps not well separated) sequences. We have developed tests based on four different eigenvalue statistics (the nearest-neighbor spacing distribution, the Dyson-Mehta Δ_3 measure of long-range order, the internal energy, and a statistic related to the *Q* statistic originally proposed by Dyson and Mehta). Fortran codes implementing these tests are available at http://www-nds.iaea.org/missing-levels/ (see MF and MF2 codes). The techniques were applied to the known resonances in the n+^{235}U reaction; results from the new analysis were consistent with the single-sequence analysis and produced an average spacing consistent with the value listed in the RIPL-3 database.

October 2011

CONTENTS

Missing Levels with Two Superimposed Sequences

J. F. Shriner, Jr.

Tennessee Technological University

G. E. Mitchell

North Carolina State University and Triangle Universities Nuclear Laboratory

I. INTRODUCTION

The Handbook for Calculations of Nuclear Reaction Data (RIPL2) [1] discusses the importance and usage of average neutron resonance parameters as core input for nuclear reaction codes. The key parameters are neutron strength functions, capture widths, and level spacings. In a previous report [2], we discussed the importance of average level spacings and some of the issues involved in determining them. For those cases where one uses identified energy levels to determine an average spacing, any missing levels or spurious (i.e., misassigned) levels will cause an error in the measured value of the average spacing. Our previous report discussed methods to estimate the number of missing levels among a sequence of levels which share the same quantum numbers. The prototypical case to which these techniques can be applied is a set of $\ell = 0$ resonances of neutrons (or protons) on a target with spin zero. The compound nuclear states all then have angular momentum and parity $J^\pi = 1/2^+$.

However, what is often true in studying neutron resonances is that while the orbital angular momentum ℓ can be determined with reasonable certainty, to determine the angular momentum is much more difficult and/or time-consuming. For the case discussed above, s-wave resonances on spin-zero targets, there is no ambiguity. However, considering either resonances with a different value of ℓ or targets with spin greater than zero leads to more than one possible value of angular momentum for the compound nuclear states, and it is the general issue of estimating the number of missing levels in cases where J is less well determined than is ℓ that we begin to examine here. Specifically we consider the case where two different J values are possible for a given value of ℓ: this occurs for s-wave resonances on targets with spin greater than zero or for resonances with $\ell \neq 0$ on targets with spin zero. If the experimental values of J are well-determined, then one can apply our earlier techniques

to each sequence of levels. (We use the term sequence to indicate a group of levels which share a complete set of quantum numbers). However, if J is not consistently determined with confidence, the techniques we describe here can be applied to the combined data (all levels without separation by J) to estimate the overall number of missing levels.

In section II, we describe the four eigenvalue statistics that we utilize to estimate the fraction of missing levels. The section after that describes the Monte Carlo studies that we performed to characterize how each statistic varies with the missing fraction. Section IV discusses combining the results of the four tests to provide a single overall estimate of the missing fraction, and Section V illustrates the application of our method to the case of s-wave neutron resonances in the n + ^{235}U reaction. An overall summary is given in Section VI. The Appendices describe briefly the Fortran program MF2 that implements this analysis and updates to our earlier Fortran program MF that analyzes single-spectrum data. Throughout this work, we denote the observed fraction of levels by f_o and the missing fraction of levels by f_m. These two variables are very simply related via $f_o = 1 - f_m$, but there are times when it proves convenient to switch back and forth between them.

II. EIGENVALUE STATISTICS

We have chosen four eigenvalue statistics to use in our analysis. Each of these is discussed briefly in this section, both in terms of pure and complete sequences and in terms of the effects of superimposed sequences and of missing levels. We use the notation E_i for energy eigenvalues and take N_{lev} as the total number of levels under consideration.

A. Nearest-neighbor Spacing Distribution

The nearest-neighbor spacing distribution examines the distribution of the energy spacings $S_i \equiv E_{i+1} - E_i$ between adjacent levels. It is both traditional and convenient to utilize the dimensionless variable $x \equiv S/D$ instead of S itself. For pure and complete spectra obeying the GOE distribution, the probability density function is given essentially by the Wigner distribution

$$P(x) = \frac{\pi x}{2} e^{-\pi x^2/4}. \tag{1}$$

For the case where only a fraction f_o of the original levels are observed, the nearest-neighbor spacing distribution becomes [3, 4]

$$P(x) = \sum_{k=0}^{\infty} f_o \, (1 - f_o)^k \, P(k; x), \tag{2}$$

where $P(k; x)$ is the k-th nearest neighbor distribution. Note that $P(0; x)$ is the Wigner distribution and that Eq. (2) reduces to the Wigner distribution for $f_o = 1$.

To consider the two-sequence case, we start with the probability distributions when multiple sequences are present with fractions f_i. The nearest-neighbor spacing distribution when the underlying sequences are pure and complete has been given by Mehta [5], and we follow that notation here:

$$P(x) \;=\; E(x) \left\{ \sum_i f_i^2 \frac{p_i(f_i x)}{E_i(f_i x)} + \left[\sum_i f_i \frac{1 - \psi(f_i x)}{E_i(f_i x)} \right]^2 - \sum_i \left[f_i \frac{1 - \psi(f_i x)}{E_i(f_i x)} \right]^2 \right\}. \tag{3}$$

For underlying GOE spectra for each sequence, the various terms are

$$\begin{aligned} E(x) \;&=\; \prod_i E_i(f_i x), \\ p(x) \;&=\; \frac{\pi}{2} x e^{-\pi x^2/4}, \\ 1 - \psi(x) \;&=\; e^{-\pi x^2/4}, \\ E_i(x) \;&=\; 1 - \frac{2}{\sqrt{\pi}} \int_0^{\frac{\sqrt{\pi} x}{2}} e^{-y^2} \, dy. \end{aligned} \tag{4}$$

It is important to realize that the variable x here already refers to the combined sequence:

$$\begin{aligned} x \;&\equiv\; S/D \\ \frac{1}{D} \;&=\; \sum_i \frac{1}{D_i}, \end{aligned} \tag{5}$$

where D is the average spacing of the combined spectra and D_i is the average spacing of the i-th sequence. Once again the formula reduces to the Wigner distribution in the case of a pure and complete spectrum (e.g., $f_1 = 1$, all other $f_i = 0$).

To include missing levels in the two-sequence analysis, we utilize two papers by Bohigas and Pato [4, 6] that discuss the general issue of missing levels. In the first of these, they discuss specifically the behavior of the $E(k, s)$ functions that give the probability of finding k levels ($k = 0, 1, 2, \ldots$) inside an interval of length s. The corresponding function $e(k, s)$

when a fraction f_o of the levels are observed is

$$e(k,s) = \sum_{j=k}^{\infty} \frac{j!}{k!(j-k)!} f_o^k (1-f_o)^{j-k} E\left(j, \frac{s}{f_o}\right). \tag{6}$$

This in turn allows us to determine the probability function when levels are missing in terms of the pure and complete probabilities and the quantity f_o. In our work here, we will consider only the case of two superimposed spectra with fractions α and $1-\alpha$. For $\alpha = 0$ – the result for a single sequence – we obtain:

$$P(x) = \sum_{k=0}^{\infty} (1-f_o)^k P\left(k, \frac{x}{f_o}\right), \tag{7}$$

which is simply a different way of stating Eq. (2). The more general result can be written

$$P(\alpha, x) = \sum_{k=0}^{\infty} (1-f_o)^k \, P\left(k, \alpha, \frac{x}{f_o}\right). \tag{8}$$

where $P\left(k, \alpha, \frac{x}{f_o}\right)$ is the k-th nearest-neighbor distribution with the appropriate value of α for two superimposed sequences.

Note that values of $f_o > 1$ are not only unphysical but they would negate the use of Eq. (7) as a probability function since the function takes on negative values in this case for some values of x. Thus our mathematical analysis will be restricted to $f_o \leq 1$.

In closing, we point out that in the limit that the number of superimposed sequences approaches ∞, the NNSD density function becomes that of a Poisson distribution:

$$P(x) = e^{-x}. \tag{9}$$

We shall see in our discussion of other statistics that the Poisson limit is relevant to understanding the behavior when levels are missed.

B. Δ_3

Early works by Dyson and Mehta [7–11] introduced several eigenvalue statistics, and the so-called Δ_3 is the best known and most widely used of those. Δ_3 involves a comparison between a best-fit straight line and the number plot (also known as the "staircase function") $N(E)$ for the spectrum eigenvalues:

$$\Delta_3 \equiv \min_{A,B} \frac{1}{E_{max} - E_{min}} \int_{E_{min}}^{E_{max}} [N(E) - AE - B]^2 dE. \tag{10}$$

For GOE statistics, the expected value must be evaluated numerically, but for large N approaches

$$\Delta_3(N) \approx \frac{1}{\pi^2}\left[\ln N - 0.0687\right]. \tag{11}$$

The effect on the average value of Δ_3 due to missing levels has been given by Bohigas and Pato [4]. In terms of the observed fraction f_o, this can be expressed

$$\Delta_3(f_o, N_{lev}) = (1 - f_o)\,\Delta_{3,Poisson}\,(N_{lev}) + f_o^2 \Delta_{3,GOE}\left(\frac{N_{lev}}{fo}\right). \tag{12}$$

The term $\Delta_{3,Poisson}$ is the value of Δ_3 for a Poisson spectrum and is given by $N_{lev}/15$.

The average behavior for two pure and complete superimposed sequences was originally discussed by Dyson and Mehta [10]:

$$\Delta_3(N_{lev}, \alpha) = \Delta_{3,GOE}\,(\alpha N_{lev}) + \Delta_{3,GOE}\,([1-\alpha]\,N_{lev})\,. \tag{13}$$

Therefore, to deal with both effects simultaneously we combine the two results:

$$\Delta_3(N_{lev}, f_o, \alpha) = (1 - f_o)\,\frac{N_{lev}}{15} + f_0^2\left[\Delta_{3,GOE}\left(\frac{\alpha N_{lev}}{f_o}\right) + \Delta_{3,GOE}\left(\frac{[1-\alpha]\,N_{lev}}{f_o}\right)\right]. \tag{14}$$

An important note is that although this function is not defined for $f_o < 0$, it is defined (and continuous) for $f_o > 1$. We allow $f_o > 1$ in the Δ_3 analysis to provide a consistency check, even though such values have no physical meaning.

C. Thermodynamic energy U

Another statistic introduced in Dyson's original papers [7, 9] is the internal energy of a spectrum. Dyson explored this statistic largely within the Circular Orthogonal Ensemble (COE), which confines the eigenvalues to lying on a circle. Applying the calculation to a set of experimental eigenvalues required the development of a potential energy to provide the confinement as well as a scaling of the eigenvalues [12]. The computation is most simply performed when the energy range in question is centered on $E = 0$ and has an average spacing of 1 in appropriate units. We begin with the eigenvalues $\{E_i\}$, and apply two transformations which normalize the average spacing and center the eigenvalues, respectively:

$$\mathcal{L} \equiv N/2, \tag{15}$$

$$\varepsilon_i \equiv \frac{2\mathcal{L} - 1}{E_N - E_1} E_i, \tag{16}$$

$$\epsilon_i \equiv \varepsilon_i - \left(\varepsilon_1 + \mathcal{L} - \frac{1}{2}\right). \tag{17}$$

The necessary potential energy function is

$$V(E) \equiv (\mathcal{L} - E)\left[\frac{1}{2} + \ln\left(\frac{\mathcal{L} - E}{2\mathcal{L}}\right)\right] + (\mathcal{L} + E)\left[\frac{1}{2} + \ln\left(\frac{\mathcal{L} + E}{2\mathcal{L}}\right)\right]. \tag{18}$$

We also must define a so-called "picket fence" spectrum to serve as a "ground state"; this is a spectrum which has equally spaced levels located at energies

$$\xi_i \equiv -\mathcal{L} + i - \frac{1}{2}. \tag{19}$$

The statistic U, which is the thermal energy per level, can then be calculated:

$$U = -\frac{1}{N}\left[\sum_{j>i} \ln\left|\frac{\epsilon_j - \epsilon_i}{j - i}\right| + \sum_i \left[V(\epsilon_i) - V(\xi_i)\right]\right]. \tag{20}$$

The expected value of U for a large GOE spectrum is $U_0 \equiv 1 - \frac{1}{2}\gamma - \frac{1}{2}\ln 2 \approx 0.365$. A Poisson spectrum has

$$U_{Poisson} = \frac{1}{N_{lev}}\left[\sum_{i=1}^{N_{lev}} \left[(N_{lev} - i)\ln i - V\left(\xi_i\right)\right] - \frac{1}{4}N_{lev}\left(N_{lev} - 1\right)\left(2\ln N_{lev} - 3\right)\right]. \tag{21}$$

Dyson himself [9] gave the expected behavior when two GOE spectra with frequencies α and $1 - \alpha$ are superimposed:

$$U = U_0 - \frac{1}{2}\left[\alpha\ln\alpha + (1 - \alpha)\ln\left(1 - \alpha\right)\right]. \tag{22}$$

We are unaware of any theoretical discussion that specifically addresses the effects of missing levels on U.

D. $\hat{Q}$

In the same paper in which Dyson and Mehta defined Δ_3, they set out to design a statistic that was similar to U but had a smaller variance [10]. They labeled this variable Q and defined it in such a way that it is a measure of short-range order in the spectrum (both Δ_3 and U are long-range measures, whereas the nearest-neighbor spacing distribution is another short-range measure). Because the expectation value of Q is proportional to N_{lev}, we have chosen to focus instead on what we call $\hat{Q} \equiv Q/N_{lev}$. The definition and computational details have been given elsewhere [2] and will not be repeated here, except to say that we choose the value $M = 4$ in all our calculations.

For a large pure and complete GOE spectra, the expected value of $\hat{Q}$ is

$$\hat{Q}_0 \approx U_0 - \left(\pi^2 M\right)^{-1} \approx 0.340. \tag{23}$$

For a superposition of two pure and complete GOE sequences, $\hat{Q}$ behaves similarly to U:

$$\hat{Q} = \hat{Q}_0 - \frac{1}{2}\left[\alpha \ln \alpha + (1-\alpha)\ln(1-\alpha)\right]. \tag{24}$$

We are unaware of an analytic form for $\hat{Q}$ for Poisson spectra but have determined that its average can be described empirically over the range $N_{lev} = 50 - 1250$ by the function

$$\hat{Q}_{Poisson} \approx 1.31 + 0.051 \ln N_{lev}. \tag{25}$$

As with U, we are unaware of any discussion of how $\hat{Q}$ would be affected by missing levels.

E. Linear Correlation Coefficient

In our previous work on estimating missing levels for a single sequence [2], we had utilized a fifth statistic, the linear correlation coefficient between adjacent spacings. We knew from that previous work that this statistic was often the least sensitive of the five. However, as we examined that statistic in more detail, we found too many cases where the value of this particular statistic fell outside the range of expected values. For these reasons, we have not included this statistic in this analysis (and we have removed it from our single-sequences analysis as well).

III. MODELING THE STATISTICS

As stated earlier, our interest here is specifically the case of two superimposed spectra with fractions α and $1-\alpha$. We assume these fractions to have relative magnitudes given by the statistical factor $2J_i + 1$. Thus for two spin states J_1 and J_2 with $J_1 < J_2$, we take $\alpha = \frac{2J_1+1}{(2J_1+1)+(2J_2+1)}$. It is certainly true that for a given set of levels, alpha may differ from this value. However, it seems likely that such information would be known only if the spin assignments are relatively robust. In such cases, analyzing the two sequences separately is almost certainly superior. The cases in which the methods described here will be most useful are those in which there are two known sequences but a significant fraction of the

spin assignments are uncertain. In those cases, it is unlikely that alpha will be known with confidence.

To test the method, Monte Carlo simulations were performed for each of the four statistics. Two GOE sequences of the same size are generated, a value of α is specified, and the appropriate numbers of levels from each of the GOE sequences are chosen so that when combined they will yield a single spectrum with both the desired number of levels and the desired value of α. The two sequences are separately unfolded with the method described in Ref. [2], and then the smaller sequence is rescaled so that it spans the same energy range as the larger sequence. The two sequences are now merged to provide a spectrum to be analyzed. 2500 such spectra are generated for each value of N_{lev}, f_m, and α. Results are determined for (1) $N_{lev} = 50$, 100, 200, 500, and 1000, (2) f_m ranging from 0 to 0.5 in steps of 0.1, and (3) α ranging from 0 to 0.5 in steps of 0.1.

For the three statistics Δ_3, U, and $\hat{Q}$, we utilize the spectral averaging technique described in Ref. [13]; for a spectrum of size N_{lev}, the statistics are evaluated over subsets of size $N_{lev}/2$. The values of these statistics for the various subsets are then averaged to provide the spectral value. As we discuss each of these three statistics below, one will observe that the fitting functions therefore depend on $N_{lev}/2$ even though the spectrum has twice that many levels.

A. NNSD

Eq. (3) provides the $k = 0$ term on the right in Eq. (8). However, analytic expressions are not available for the $k > 0$ terms. We have determined these numerically using Monte Carlo calculations and interpolation. To determine the values to use in interpolation, spacings are analyzed for 2500 spectra of size 2400. An average behavior for $P(k, \alpha, x)$ was determined over a grid of k, α, x values: (1) k varies from 1 to 10 in steps of 1; (2) α varies from 0 to 0.5 in steps of 0.05; (3) x varies from 0 to 16 in steps of 0.05. These values were stored and are used to interpolate for intermediate values of α and x. In practice, we required the function $P(k, \alpha, x)$ to be zero for any value of x that varies from $k + 1$ (the expected mean of the k-th neighbor distribution) by more than 4. This was a very conservative estimate and could probably have been smaller without noticeably affecting our results.

The analysis determines f_o via a maximum likelihood analysis (similar to the single-sequence approach described in Ref. [3]). For each case, 2500 spectra were generated, a

value of f_o was extracted for each spectrum, and the overall behavior of the ensemble was characterized by the median value of f_o along with the 16-th and 84-th percentiles. (Because our ultimate interest is in analyzing a single spectrum and these distributions are noticeably asymmetric, we have chosen the median value as the one that is most suitable to describe our distribution.) We then compare the extracted values of f_m to the known values.

We can draw several conclusions from our analysis. (1) In general, the medians reproduce the true values quite well. (2) There appears to be a tendency for the $\alpha = 0.1$ results to be a bit lower than expected, especially when compared to results for $\alpha = 0$ and $\alpha = 0.2$. The differences are still within normal variations, but the trend occurs frequently enough to warrant suspicion. However, we note that for our intended purposes (either $\ell \neq 0$ waves with spin-zero targets or s-waves with non-zero-spin targets), only values $0.25 \leq \alpha < 0.5$ will be under consideration. (3) The uncertainties get larger as N_{lev} gets smaller. This is not surprising, of course, but the size of the uncertainties suggests that this method *by itself* is not reliable for $N_{lev} \lessapprox 100$. (4) The uncertainties grow significantly as α increases.

As we did with the pure/complete sequence case, we can characterize our extracted value $\tilde{f}$ as a linear function of f_o (independent of N and α):

$$\tilde{f} = a_o + a_1 f_o. \tag{26}$$

The fit parameters are determined to be $a_0 = 0.0460 \pm 0.021$ and $a_1 = 0.938 \pm 0.025$. The behavior is illustrated in Fig. 1.

B. Δ_3

Δ_3 distributions are notably asymmetric, and Eq. (14) describes average behavior. Since we wish to utilize the median rather than the average, we utilize a pair of empirical coefficients b_0 and b_1. (We also note that, as described earlier, we are using intervals of size $N_{lev}/2$; therefore, the right hand side of this equation has $N_{lev}/2$ anytime the number of levels appears.) We fit our simulations with the following equation:

$$\Delta_3(N_{lev}, f_o, \alpha) = b_0 \left(1 - f_o\right) \frac{N_{lev}}{30} + b_1 f_o^2 \left[\Delta_{3,GOE}\left(\frac{\alpha N_{lev}}{2f_o}\right) + \Delta_{3,GOE}\left(\frac{\left[1-\alpha\right] N_{lev}}{2f_o}\right)\right]. \tag{27}$$

The fit parameters are determined to be $b_0 = 0.921 \pm 0.037$ and $b_1 = 0.982 \pm 0.032$. The behavior is illustrated in Fig. 2. We point out that Δ_3, by the nature of Eq. (14) and the

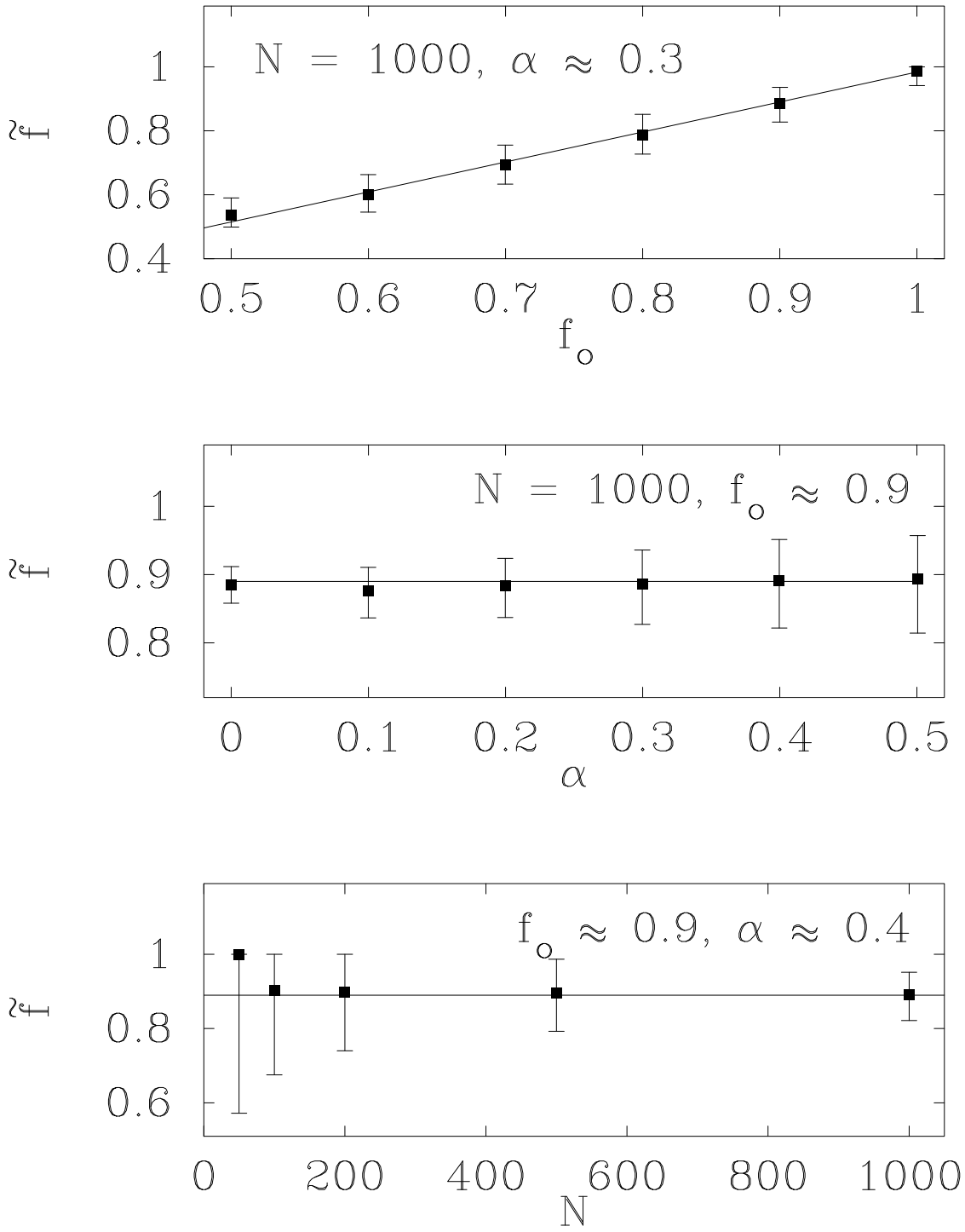

FIG. 1: Median values of $\tilde{f}$ as a function of N (lower part of the figure), α (center part), and f_o (upper part). In each case, the other two parameters are held constant at the values shown. The solid lines represent the best-fit function of $\tilde{f}$ given in Eqn. (26).

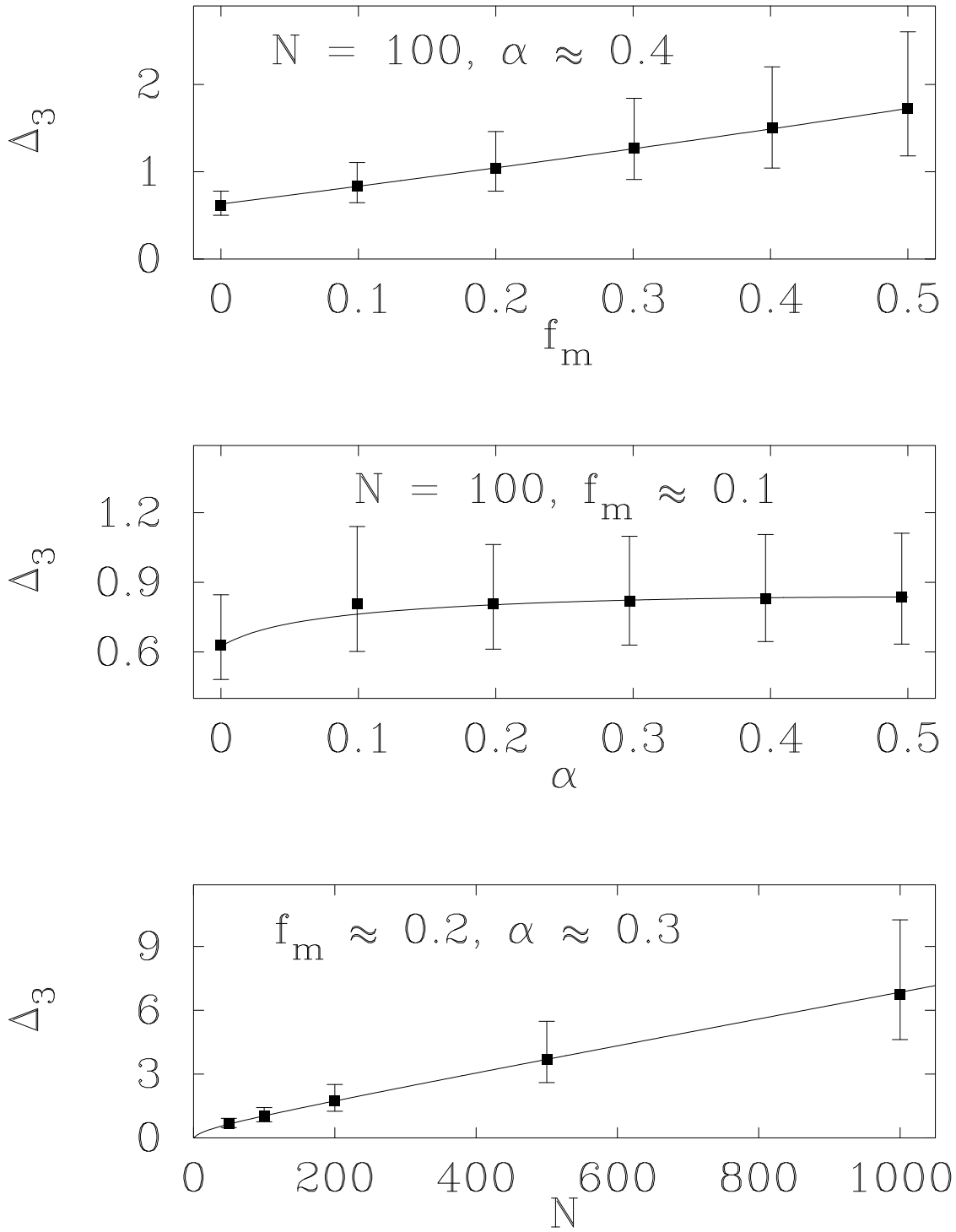

FIG. 2: Median values of Δ_3 as a function of N (lower part of the figure), α (center part), and f_o (upper part). In each case, the other two parameters are held constant at the values shown. The solid lines represent the best-fit function of Δ_3 given in Eqn. (27).

fact that $\Delta_{3,Poisson}$ is so much larger than $\Delta_{3,GOE}$, is much more sensitive to the presence of missing levels than it is to the particular value of α. That is a very attractive feature for our purposes, since it means that our assumptions about the value of α are unlikely to be critical. Two recent paper by Mulhall [14, 15] also examined the problem of using Δ_3 to estimate the fraction of missing levels. He utilized the results of Monte Carlo simulations in lieu of a functional dependence as given in Eq. (27) and then empirically parameterized the probability function to allow a maximum-likelihood approach. He also treated Δ_3 for each spectrum as a function of different numbers of levels rather than just utilizing a single value of Δ_3 for $N_{lev}/2$ levels as we have. The most recent of these papers does compare the maximum-likelihood method with the use of Eq. (27), but it applies Eq. (27) using a least-squares evaluation over different values of the averaging interval. Because of these differences, a direct comparison of the two approaches is difficult. Mulhall suggests that uncertainties in f_o are smaller with the maximum-likelihood approach, but it seems to us that the values of Δ_3 for different lengths of the averaging interval are not statistically independent. That would make it much more difficult to obtain an appropriate assessment of the uncertainty with that method.

C. U

To fit the statistic U, we choose to try a function similar to that used for Δ_3:

$$\begin{aligned} U\left(N_{lev}, f_o, \alpha\right) &= c_0\left(1-f_o\right) U_{Poisson}\left(\frac{N_{lev}}{2}\right) + c_1 f_o^2 U_{GOE}\left(\frac{N_{lev}}{2 f_o}\right) \\ &= c_0\left(1-f_o\right) U_{Poisson}\left(\frac{N_{lev}}{2}\right) + c_1 f_o^2\left[U_0 - \frac{1}{2}\left\{\alpha \ln \alpha + (1-\alpha)\ln(1-\alpha)\right\}\right] \end{aligned} \tag{28}$$

This proves to be a much simpler expression from which to extract f_o than the corresponding one for Δ_3 in Eq. (27) because U_{GOE}, at least in the limit of large N_{lev}, does not depend on N_{lev}. The best-fit parameters are $c_0 = 0.957 \pm 0.020$ and $c_1 = 1.004 \pm 0.014$. This function does a good job of characterizing the data, as is illustrated in Fig. 3.

Further examination of Eq. (28) reveals that for fixed N_{lev} and α, U is a quadratic function of f_o that monotonically decreases over the range $f_o \epsilon [0,1]$. U reaches a minimum value of

$$U_{min} = c_0 U_{Poisson}\left(\frac{N_{lev}}{2}\right)\left[1 - \frac{c_0 U_{Poisson}\left(N_{lev}/2\right)}{4 c_1 U_{GOE}}\right]. \tag{29}$$

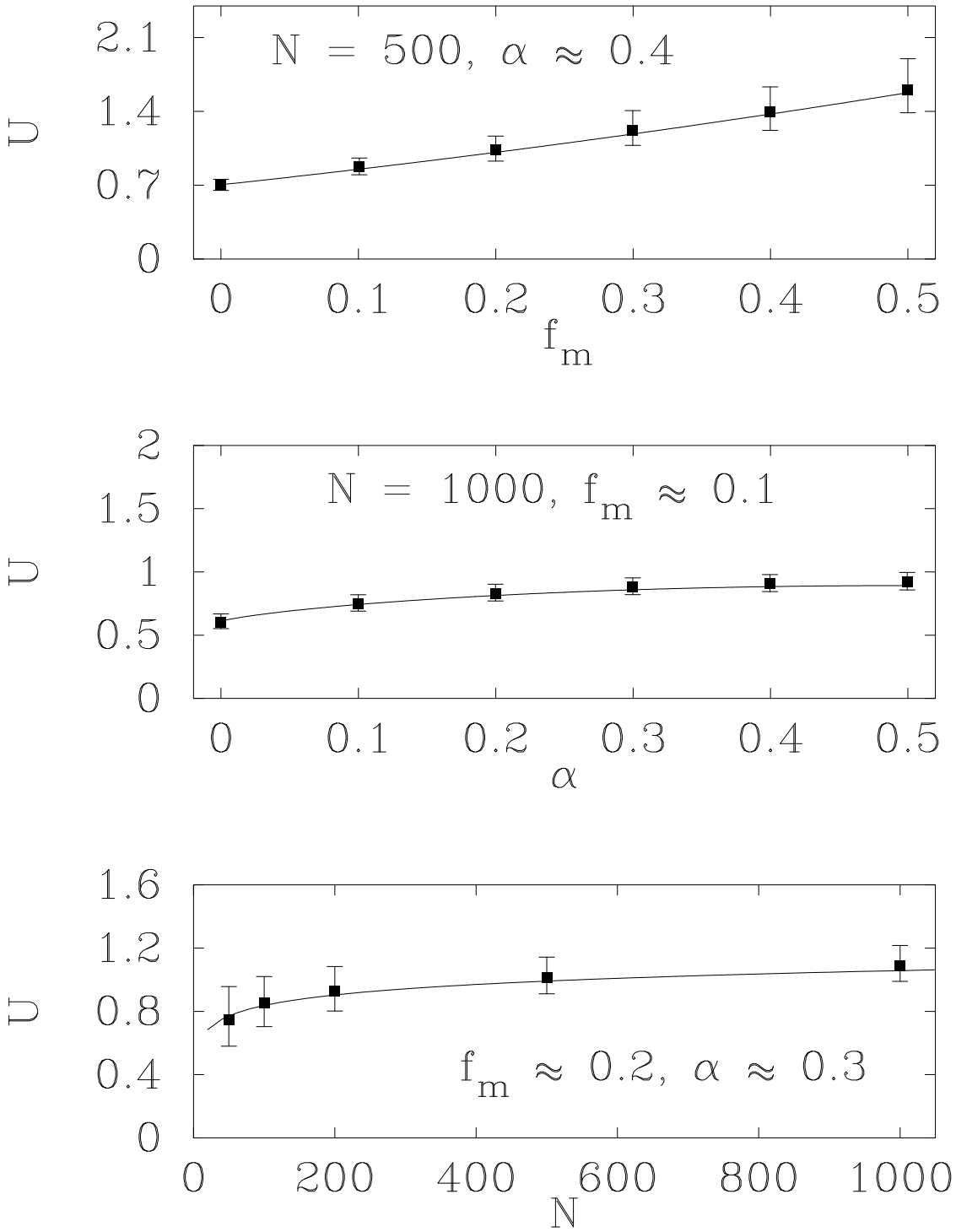

FIG. 3: Median values of U as a function of N (lower part of the figure), α (center part), and f_m (upper part). In each case, the other two parameters are held constant at the values shown. The solid lines represent the best-fit function of U given in Eqn. (28).

For values $N_{lev} \geq 50$, all values of U_{min} are negative, and the minimum value occurs at a value $f_{crit} > 1$. Thus the smaller of the two solutions to Eq. (28) is the appropriate one for our purposes.

D. $\hat{Q}$

The statistic $\hat{Q}$ behaves similarly to U in many ways, and we try a fitting function of a similar form:

$$\begin{aligned}\hat{Q}\left(N_{lev}, f_o, \alpha\right) &= d_0\left(1-f_o\right)\hat{Q}_{Poisson}\left(\frac{N_{lev}}{2}\right) + d_1 f_o^2 \hat{Q}_{GOE}\left(\frac{N_{lev}}{2f_o}\right) \\ &= d_0\left(1-f_o\right)\left[1.31 + 0.051\ln\left(\frac{N_{lev}}{2}\right)\right] + d_1 f_o^2\left[\hat{Q}_0 - \frac{1}{2}\left\{\alpha\ln\alpha + (1-\alpha)\ln(1-\alpha)\right\}\right] \quad (30)\end{aligned}$$

We find that this function provides a reasonable description of our Monte Carlo results. The best-fit parameters are $d_0 = .968 \pm 0.013$ and $d_1 = 0.980 \pm 0.012$. This behavior is illustrated in Fig. 4.

As was also true with U, Eq. (30) reveals that for fixed N_{lev} and α, $\hat{Q}$ is a quadratic function of f_o that monotonically decreases over the range $f_o \epsilon [0, 1]$. The function in Eq. (30) has a minimum value of

$$\hat{Q}_{min} = d_0 \hat{Q}_{Poisson}\left[1 - \frac{d_0 \hat{Q}_{Poisson}}{4 d_1 \hat{Q}_{GOE}}\right]. \quad (31)$$

The minimum value of $\hat{Q}$ is negative for $N_{lev} \geq 50$, and the corresponding value of f is $f_{crit} > 1$. So again, it is the smaller of the two solutions to Eq. (30) that is relevant for our purposes.

IV. COMBINING THE STATISTICS

For any spectrum, we end up with four different values of f_o, since one is obtained from analysis of each of the four statistics. We want to combine those four values to produce a single value of f_o that represents the set of eigenvalues in question. However, the standard assumptions of statistical error propagation may not be appropriate: in some cases, the uncertainties on a value may be asymmetric or limited by the fact that a particular distribution is limited to certain values of f_o. For example, in applying the maximum likelihood method

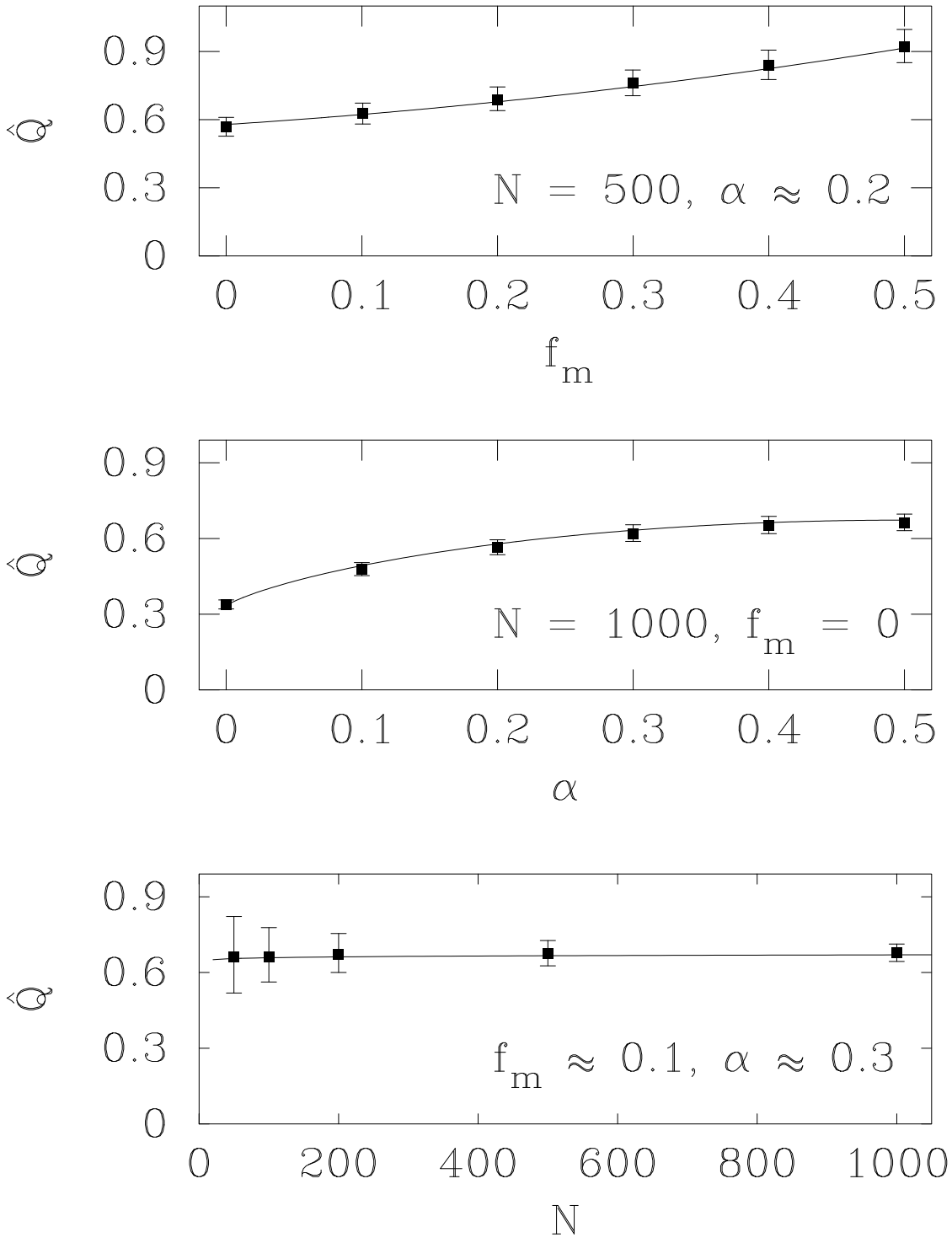

FIG. 4: Median values of $\hat{Q}$ as a function of N (lower part of the figure), α (center part), and f_m (upper part). In each case, the other two parameters are held constant at the values shown. The solid lines represent the best-fit function of $\hat{Q}$ given in Eqn. (30).

to the nearest-neighbor spacing distribution, one can estimate uncertainties in each direction by determining where the logarithm of the likelihood function has decrease by 0.5, but the probability function is not defined for $f_o > 1$. Thus, one has the possibility of obtaining a result like $f_o = 1.00^{+0.00}_{-0.12}$. The upper uncertainty of zero is problematic in applying standard error propagation methods.

Instead we apply a Monte Carlo technique; from the distribution for each statistic, a random number representing that distribution is generated. Any random numbers corresponding to $f_o > 1$ are replaced with $f_o = 1$ due to the physical requirement that $f_o \epsilon [0, 1]$. The four randomly generated values of f_o are then averaged to provide one representative value of f_o for the spectrum. The procedure is repeated many times, and the average of the various spectral values is what we cite as our best estimate of f_o. Uncertainties on this value are determined from values corresponding to the 16th and 84th percentiles of the distribution of these averages.

The observed value of the average spacing, D_{obs} is evaluated via

$$D_{obs} = \frac{E_{Nlev} - E_1}{N_{lev} - 1} \tag{32}$$

with a relative uncertainty of $\sqrt{0.27/N_{lev}}$. The values of f_o and D_{obs} are then multiplied to yield the best estimate of the average spacing D for this series of levels.

V. APPLICATIONS

As discussed in the Introduction, the purpose of this study is to provide an estimate of the missing level fraction in cases where known levels belong to two different J^π values and those values are not well determined. However, to test the method, one would like a case (or cases) where those J^π values are well determined so that one can compare this approach to what would be obtained by analyzing the two sequences separately. One such case is provided by the s-wave resonances in the n + ^{235}U reaction [16], which have J^π values of 3^- and 4^-. The levels with $E_n \leq 500$ eV appear to be suitable for this analysis; above this energy there is a significant change in the observed level density as illustrated in Fig. 5. The levels assigned $J = 3$ and $J = 4$ are analyzed separately with the program MF described in our earlier report [2]. Then all $\ell = 0$ levels are grouped together and evaluated with the new program MF2 designed to implement the analyses described in this report. Our results

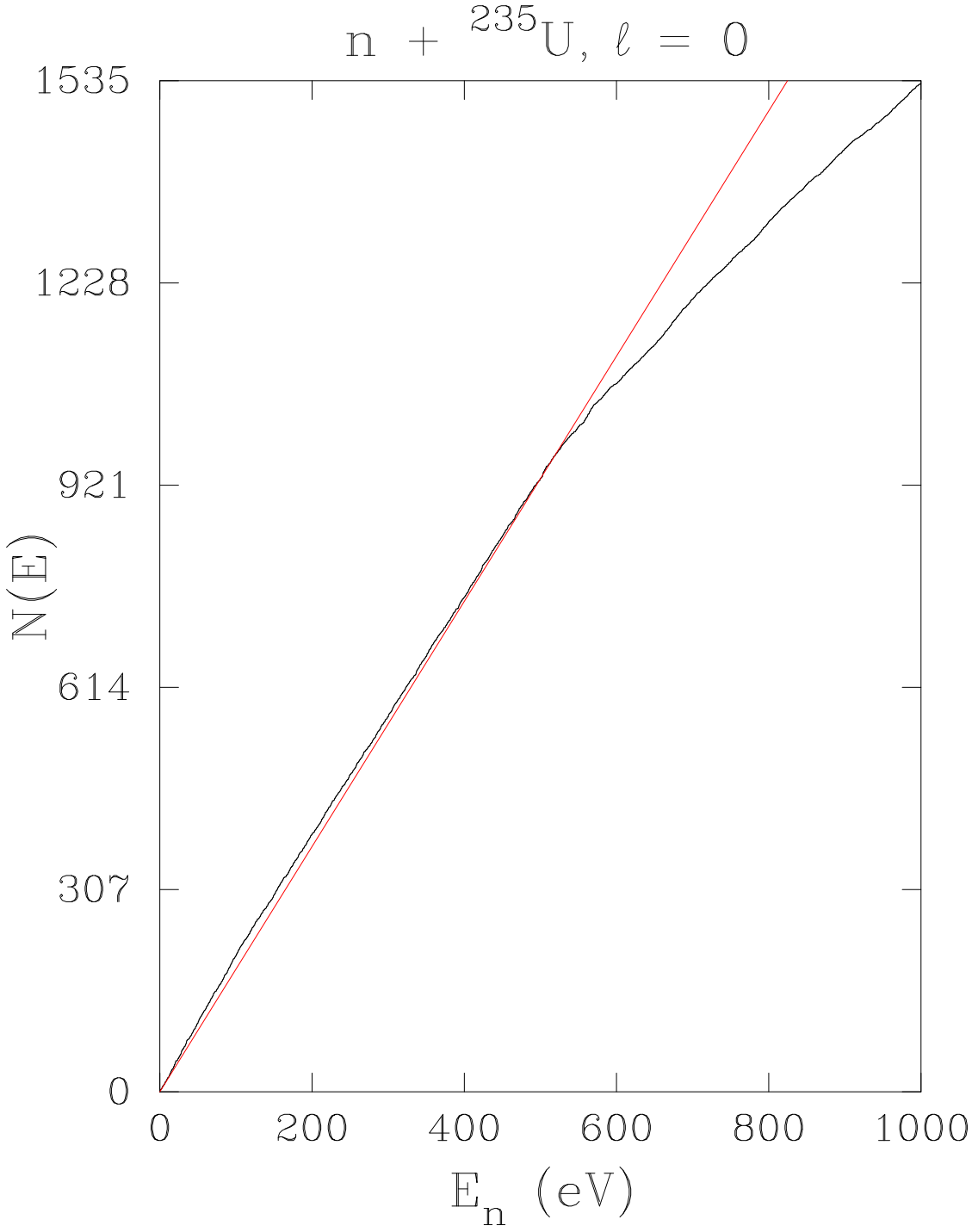

FIG. 5: The cumulative number of neutron levels as a function of neutron energy for the n + ^{235}U reaction. The red line shows the expected behavior if the average spacing over the entire energy range was equal to its value in the range $E_n = 0 - 500$ eV.

are summarized in Table I.

TABLE I: Analysis of n + ^{235}U s-wave resonances with $E_n = 0 - 500$ eV. There are 373 levels with $J = 3$ and 558 levels with $J = 4$.

	$J = 3$		$J = 4$		Combined	
	Value	f_o	Value	f_o	Value	f_o
NNSD	N/A	$0.87^{+0.05}_{-0.06}$	N/A	$0.96^{+0.03}_{-0.04}$	N/A	$1.00^{+0.00}_{-0.04}$
Δ_3	$1.7^{+1.5}_{-1.4}$	$0.88^{+0.14}_{-0.13}$	$1.1^{+0.5}_{-0.6}$	$0.97^{+0.03}_{-0.04}$	$4.2^{+3.6}_{-3.7}$	0.87 ± 0.14
U	$0.63^{+0.15}_{-0.16}$	$0.85^{+0.09}_{-0.08}$	$0.53^{+0.03}_{-0.04}$	0.92 ± 0.02	$0.90^{+0.16}_{-0.17}$	0.89 ± 0.09
$\hat{Q}$	0.44 ± 0.01	$0.87^{+0.01}_{-0.02}$	$0.44^{+0.02}_{-0.01}$	$0.87^{+0.02}_{-0.03}$	0.65 ± 0.01	1.09 ± 0.14
Combined	N/A	$0.86^{+0.04}_{-0.03}$	N/A	$0.91^{+0.02}_{-0.01}$	N/A	0.93 ± 0.04
	$D_{0,J=3} = 1.16^{+0.05}_{-0.06}$ eV		$D_{0,J=4} = 0.82 \pm 0.02$ eV			
	$D_0 = 0.48 \pm 0.01$ eV				$D_0 = 0.50 \pm 0.02$ eV	

In general, the values of f_o for a given analysis are internally consistent. Even more importantly, the value of the average $\ell = 0$ spacing D_0 obtained by analyzing the two sequences separately is in excellent agreement with the value obtained with our new methods, albeit with a slightly larger uncertainty. The derived value is also in agreement with the RIPL-3 recommended value of 0.45 +/- 0.03 eV [17].

VI. SUMMARY

Since values of the level density are crucial for pure and applied nuclear science, the motivation for determining the fraction of missing levels in neutron resonances is clear. The standard methods all assume the validity of the Gaussian Orthogonal Ensemble (GOE) version of Random Matrix Theory (RMT) and, in particular, that the Porter-Thomas distribution describes the width distribution. A difficulty is that non-statistical effects can

distort the width distribution and lead to incorrect values for the fraction of missing levels. In RMT the eigenvalues and eigenvectors are independent of one another, and the spacing distribution is unaffected by non-statistical effects. Therefore we developed methods of determining the fraction of missing levels from the spacings rather than the widths.

We previously described the application of properties of the spacing distribution to determining the missing level fraction for spin zero targets [2]. The Nearest Neighbor Spacing Distribution (NNSD) is the most widely used statistic to evaluate spectral statistics and the most complicated to apply to the missing level problem. Several other statistics were also used (the Δ_3 statistic, the thermodynamic energy, etc.). All except the NNSD determine a specific value of the statistic for a spectrum and therefore are somewhat easier to apply in practice. In Ref. [2] we discussed all of the tests as well as a computer program for their application.

In the present work we considered how to determine the missing fraction for resonances on targets with spin greater than zero. Specifically, we studied estimating the missing fraction when the data are a superposition of two GOE spectra. In practice the primary use is for s-wave resonances on targets with spin greater than zero, although these same methods can be applied to p-wave resonances on targets with spin zero. We applied four tests that were previously utilized in the one-sequence analysis (for various reasons we decided to omit the linear correlation coefficient for adjacent spacings). In the process we implemented fitting functions for the U and $\hat{Q}$ statistics that had a more physical rationale than the empirical functions we had used in our previous work (and we updated our previous computer code to reflect these new functions as described in Appendix B). We also implemented a Monte Carlo approach to the averaging of the four different results, primarily due to the missing fraction's being limited to values in the interval $0 \leq f_o \leq 1$. We applied the methods to data for the n + ^{235}U reaction, a case where the spin values are believed known well enough to allow us to compare our two-sequence analysis with results from analyzing each set of spin states separately. The two methods provide values of the s-wave average spacing D_0 that are in excellent agreement with each other. The uncertainty on D_0 is a bit larger in the two-sequence analysis, and we expect that such behavior will be true in general.

In general, if the spins of a collection of states with two possible spins are known fairly reliably, then using the methods described in our original report and implemented in the computer code MF is likely the best approach to correct for missing levels. However, if the

spins are unreliable or unknown, then the methods described in this report and implemented in the code MF2 become necessary. Even for cases where the spins are thought reliable, a check with the program MF2 seems advisable; strongly differing results between the two approaches might be evidence of more significant problems with the spin assignments than believed.

VII. ACKNOWLEDGMENTS

This work was supported in part by the US Department of Energy under grant DE-FG02-97ER41042 and the National Nuclear Security Administration under grant DE-FG52-09NA29460.

USE OF THE PROGRAM MF2

Input for the Fortran program MF2 consists of either one file containing a list of the energy levels contained in both sequences or two separate files, each containing a list of the levels in a particular sequence. The program asks the user whether 1 or 2 input files are to be utilized and then asks the user to specify the name(s) of the file(s). Each input file must have the minimum and maximum energies to be included in the analysis on its first line (if using two files, these should be the same in the two cases). Then each successive line should have a level energy as the first value on each line (more than one value per line may be present as long as the energy is the first one). The program also requests the name of an output file to which the results will be written; output is directed both to the specified file and to the terminal.

There are a series of 10 auxiliary files that must accompany the program, named Pxx.dat where xx takes on the numbers 01, 02, ... 10. These are used in the interpolations necessary to evaluate the $P(k, \alpha, x)$ functions used in the NNSD analysis. These files must be placed in an appropriate directory so that the program can access them by name without any additional path information.

UPDATES TO THE PROGRAM MF [2]

As we implemented the ability to analyze two-sequence data, we learned different ways of describing the behavior of various statistics when levels are missing. Our earlier descriptions of U and $\hat{Q}$ for single-sequence data had been purely empirical, but the functions described by Eqs. (28) and (30) have a more physical basis. It seemed appropriate to update program MF to reflect this new information. Thus in version 2 of MF, the fitting function for U is now

$$U(f_{m,lev}) = c_0\,(1 - f_o)\,U_{Poisson}\,(N_{lev}) + c_1 f_o^2 U_{GOE} \tag{33}$$

with parameters $c_0 = 0.893 \pm 0.043$ and $c_1 = 0.994 \pm 0.035$. The best-fit function for $\hat{Q}$ has been changed to

$$Q(f_m, N_{lev}) = d_0\,(1 - f_o)\,\hat{Q}_{Poisson}\,(N_{lev}) + d_1 f_o^2 \hat{Q}_{GOE} \tag{34}$$

with parameters $d_0 = 0.892 \pm 0.025$ and $d_1 = 0.997 \pm 0.032$. Other changes to MF are the removal of the linear correlation coefficient as one of the statistics and the implementation of the Monte Carlo technique described in Sect. IV to combine the results of the different statistics.

---

[1] T. Belgya, O. Bersillon, R. Capote, T. Fukahori, Zhigang Ge, S. Goriely, M. Herman, A.V. Ignatyuk, S. Kailas, A. Koning, V. Plujko, and P. Young. Handbook for calculations of nuclear reaction data, RIPL-2 (Reference Input Parameter Library-2). Technical Report IAEA-TECDOC-1506, International Atomic Energy Agency, Vienna, Austria, 2006. Available online at *http://www-nds.iaea.org/RIPL-2*.

[2] G. E. Mitchell and J. F. Shriner, Jr. Missing level corrections using neutron spacings. Technical Report *INDC(NDS)-0561*, International Atomic Energy Agency, Vienna, 2009. Available online at *http://www-nds.iaea.org/publications/indc/indc-nds-0561.pdf*.

[3] U. Agvaanluvsan, G. E. Mitchell, J. F. Shriner, Jr., and M. Pato. Missing level corrections in nuclear resonances. *Nucl. Instrum. Methods* A498 (2003) 459.

[4] O. Bohigas and M. P. Pato. Missing levels in correlated spectra. *Phys. Lett.* B595 (2004) 171.

[5] M. L. Mehta. *Random Matrices*. Academic, New York, second edition, 1991.

[6] O. Bohigas and M. P. Pato. Randomly incomplete spectra and intermediate statistics. *Phys. Rev.* E74 (2006) 036212.

[7] F. J. Dyson. Statistical theory of the energy levels of complex systems. i. *J. Math. Phys.* 3 (1962) 140.

[8] F. J. Dyson. Statistical theory of the energy levels of complex systems. ii. *J. Math. Phys.* 3 (1962) 157.

[9] F. J. Dyson. Statistical theory of the energy levels of complex systems. iii. *J. Math. Phys.* 3 (1962) 166.

[10] F. J. Dyson and M. L. Mehta. Statistical theory of the energy levels of complex systems. iv. *J. Math. Phys.* 4 (1963) 701.

[11] M. L. Mehta and F. J. Dyson. Statistical theory of the energy levels of complex systems. v. *J. Math. Phys.* 4 (1963) 713.

[12] J. F. Shriner, Jr., M. P. Pato, G. E. Mitchell, and A. P. B. Tufaile. A thermodynamic analysis of energy eigenvalues. *Nucl. Instrum. Methods* A581 (2007) 831.

[13] D. Mulhall, Z. Huard, and V. Zelevinsky. Ergodicity of the Δ_3 statistic and purity of neutron resonance data. *Phys. Rev.* C76 (2007) 064611.

[14] D. Mulhall. Using the δ_3 statistic to test for missed levels in mixed sequence neutron resonance data. *Phys. Rev.* C80 (2009) 034612.

[15] D. Mulhall. Maximum likelihood method to correct for missed levels based on the δ_3 statistic. *Phys. Rev.* C83 (2011) 054321.

[16] L. C. Leal, H. Derrier, N. M. Larson, and R. Q. Wright. R-matrix analysis of ^{235}u neutron transmission and cross-section measurements in the 0- to 2.25-kev energy range. *Nucl. Sci. Eng.* 131 (1999) 230.

[17] R. Capote, M. Herman, P. Oblozinsky, P. G. Young, S. Goriely, T. Belgya, A. V. Ignatyuk, A. J. Koning, S. Hilaire, V. A. Plujko, M. Avrigeanu, O. Bersillon, M. B. Chadwick, T. Fukahori, Z. Ge, Y. Han, S. Kailas, J. Kopecky, V. M. Maslov, G. Reffo, M. Sin, E. Sh. Soukhovitskii, and P. Talou. RIPL - Reference Input Parameter Library for calculation of nuclear reactions and nuclear data evaluations. *Nucl. Data Sheets* 110 (2009) 3107. Available online at *http://www-nds.iaea.org/RIPL-3*.

Nuclear Data Section
International Atomic Energy Agency
Vienna International Centre, P.O. Box 100
A-1400 Vienna
Austria

e-mail: services@iaeand.iaea.org
fax: (43-1) 26007
telephone: (43-1) 2600-21710
Web: http://www-nds.iaea.org