Alternate methods for judging the significance of Hub Test results

Richard Shurtleff *

Abstract

Observations of astronomical objects include vectors that are transverse, perpendicular to the direction from the objects to us. The Hub Test judges how well transverse vectors on the sky align by calculating their mutual alignment angle to each point on the sky. Then the most focussed and least focussed values determine the collective alignment of these transverse vectors. This article deals with statistics. We describe a Library of random distribution data that can estimate the significance of the values obtained by the Hub Test correlation metrics. For a given experimental sample, the Library can be used via interpolation or, somewhat less accurately, via substitution in formulas that fit the Library data. The Library in this second version of the article has 294 sets of probability distributions with the number of sources ranging from 9 to 900, 14 values total, and the nominal sample radii ranging from 0.14° to 64°, 21 values total. An Appendix presents the underlying detailed calculations in a self-contained Mathematica notebook.

Keywords: Alignment; Hub Test; Transverse Vectors; Polarization; Jets; Significance

*Department of Sciences, Wentworth Institute of Technology, 550 Huntington Avenue, Boston, MA, USA, 02115, orcid.org/0000-0001-5920-759X, e-mail addresses: shurtleffr@wit.edu, momentummatrix@yahoo.com

0. Preface

The pdf version of this notebook is available online from the viXra archive.
To find the ready-to-run notebook follow the link in Ref. 1. The notebooks in this series were created using Wolfram Mathematica, Version Number: $13.0 \times .1 \times .0$, Ref. 2.

Note(s):

(1) Some numerical quantities in the pdf version may differ from the live ready-to-run version in Ref. 1 because the ready-to-run version may have been altered after the pdf was produced. The ready-to-run version and the pdf version may be updated out-of-synch.

(2) In Version 2, we expanded the Library to 294 records from Version 1's 90 records. Also, we changed the variable list. Here, the basic variables are the number of sources $N = $ nSrc and the root-mean-square radius ρRMS of a sample's region. Most functions are written in terms of $w = N^{-1/2}$, the inverse square root of N, instead of N. The nominal radius ρNominal is convenient for labeling purposes, but ρRMS is better for comparison with samples found in the wild.

(3) For Version 2, we abandon the inverse radius $\tau = 1/\rho$, reverting to the radius ρ. The radius ρ is bounded by 0 and 90°; its inverse runs off to infinity.

(4) We often hide cells. For example, the Library is based on a large numerical data table. The table is included in the Mathematica notebook, but hidden a few cells before Table A1, because it would take a lot of space to display. The table can be recovered from the links in Ref. 1.

(5) Cells that contain only "print" statements are hidden because they hold no more information than what they print. Again, if one wishes to see hidden cells, follow a link in Ref. 1.

CONTENTS

1. Introduction

Given a set of transverse vectors on the sky, one may ask if their directions are correlated. Candidates include the polarization direction of electromagnetic radiation and the direction of asymmetries such as jets. The Hub Test of alignment extends the transverse directions, making Great Circle geodesics on the Celestial Sphere. The transverse directions are perfectly aligned if they intersect at some point H on the sphere. The directions are well-aligned when they converge in a small area near some point H_{min}. The Hub Test can find correlations for samples with hubs H_{min} that are near the sources as well as the distant Hubs that other alignment tests would uncover.

The basic quantity that measures the convergence of the Great Circles is a function of position on the Celestial Sphere, denoted $\bar{\eta}(H)$. The function $\bar{\eta}(H)$ is an acute angle defined at all points of the Celestial Sphere, except at the sources. The smallest alignment angle $\bar{\eta}_{min}{}^{obs}$ and the largest avoidance angle $\bar{\eta}_{max}{}^{obs}$ provide two independent measures of the correlation of the set of transverse vectors. Given a sample, its alignment correlations are determined by $\bar{\eta}_{min}{}^{obs}$ and $\bar{\eta}_{max}{}^{obs}$.

The significance of the smallest alignment angle $\bar{\eta}_{min}{}^{obs}$ means the likelihood that random transverse directions would deliver an equal or smaller value $\bar{\eta}_{min}{}^{random}$ than $\bar{\eta}_{min}{}^{obs}$, *i.e.* better $\bar{\eta}_{min}{}^{random} <= \bar{\eta}_{min}{}^{obs}$ than the observed value. The significance is sometimes called the "*p*-value".

By its definition, the most direct process we consider is what we call the 'direct' method, "Direct Method A". For the given sample, replace the observed transverse vectors with randomly directed vectors and duplicate the calculations to find the smallest alignment angles $\bar{\eta}_{min}{}^{random}$. Repeat many times to get a distribution of values that can be fit by a formula that approximates the probability distribution of $\bar{\eta}_{min}{}^{random}$ from the random runs of data. Assuming that the fitting formula is accurate along the 'tail' of the distribution, we can estimate the significance by integrating the probability distribution to find the likelihood that random results $\bar{\eta}_{min}{}^{random}$ have values below the smallest observed alignment angle $\bar{\eta}_{min}{}^{obs}$. Direct Method A is the most reliable, but least convenient, of the three methods considered in this article.

To avoid the inconvenience of Direct Method A, we present two alternatives. We have constructed random data for simulated samples that we call a "Library". The probability distributions for random run results $\bar{\eta}_{min}{}^{random}$ are calculated by applying Direct Method A to the simulated samples. The process, discussed in Sec. 4. gives us a Library, a set of probability distributions. The distributions can be reconstructed by two-parameter fitting functions, two-parameters just like Gaussians. We record the location of the distribution's peak, $\eta_0{}^{min}$, and its half-width σ^{min}.

We identify the simulated sample by the number of sources N and the root-mean-square radius ρRMS of the sample region. For each case of $(N, \rho$RMS$)$ we have the two parameters $\eta_0{}^{min}$ and σ^{min} that describe the probability distribution of the $\bar{\eta}_{min}{}^{random}$. And two more, $\eta_0{}^{max}$ and σ^{max}, for the avoidance distribution of the $\bar{\eta}_{max}{}^{random}$.

Thus, the Library consists of the relevant properties of the simulated samples, N and ρRMS, and the four parameters, $\eta_0{}^{min}$, σ^{min}, $\eta_0{}^{max}$ and σ^{max}, needed to replicate the probability distributions of the two metrics of correlation $\bar{\eta}_{min}{}^{random}$ and $\bar{\eta}_{max}{}^{random}$. The Library also includes the standard errors for these parameters, so that estimates of the accuracy of the process can be made. Some vestigial, irrelevant quantities are also collected in the record for each case of $(N, \rho$RMS$)$.

Once the Library is uploaded, an interpolation determines the probability distribution required for an observed sample whose properties N and ρRMS are known. This method of estimating significance is termed "Interpolation Method B". Method B is presented in Sec. 5 just after the Library is built.

Formula Method C finds functions that fit the Library data, the peak $\eta_0{}^{min}$ and half-width σ^{min}, as well as the avoidance distribution parameters for the other measure of correlation, $\bar{\eta}_{max}{}^{obs}$. The functions $\eta_0{}^{min}$ and σ^{min} depend on an observed sample's number of sources N and its root-mean-square radius ρRMS. Once values of the two parameters $\eta_0{}^{min}$ and σ^{min} are determined for an observed sample, the probability distribution can be constructed and integrated to find the chance that random runs have better, smaller values of $\bar{\eta}_{min}{}^{random}$ than the observed sample's $\bar{\eta}_{min}{}^{obs}$. The formulas needed for Formula Method C are discussed in Sec. 6.

We have previously studied three observed samples whose significances were estimated with Direct Method A, the best, but least convenient method. The papers are accessible, already appearing in the on-line non-refereed literature, viXra and ResearchGate. Two samples have 27 and 13 QSO radio sources and one sample has 99 stars from our Galaxy with polarized visible starlight; see Refs. 3,4,5.

In Sec. 7 of this article, for comparison, the significances of the three samples are determined with Interpolation Method B and Formula Method C. Version 1 of the Library, Interpolation Method B and/or Formula Method C, has been applied to two catalogs, one of polarized starlight and one of radio QSOs, Refs. 6,7. These are examples of how the Hub Test can assist in identifying interesting sample for detailed study when confronted with enormous datasets.

Concluding remarks appear in Sec. 8.

2. The Hub Test in brief

The Hub Test, Ref. 8, is based on the alignment of transverse directions with directions toward other points on the Celestial Sphere. In Fig. 1, the "alignment angle" η is the acute angle η between two great circles at S, $0° \leq \eta \leq 90°$. The alignment angle η measures how well the polarization direction $\hat{v}_\psi$ matches the direction $\hat{v}_H$ toward the point H. Perfect alignment occurs when $\eta = 0°$ and the two great circles overlap. Perpendicular great circles, $\eta = 90°$, indicates maximum "avoidance" of the polarization direction $\hat{v}_\psi$ with the point H on the sphere. The halfway value, $\eta = 45°$, favors neither alignment nor avoidance.
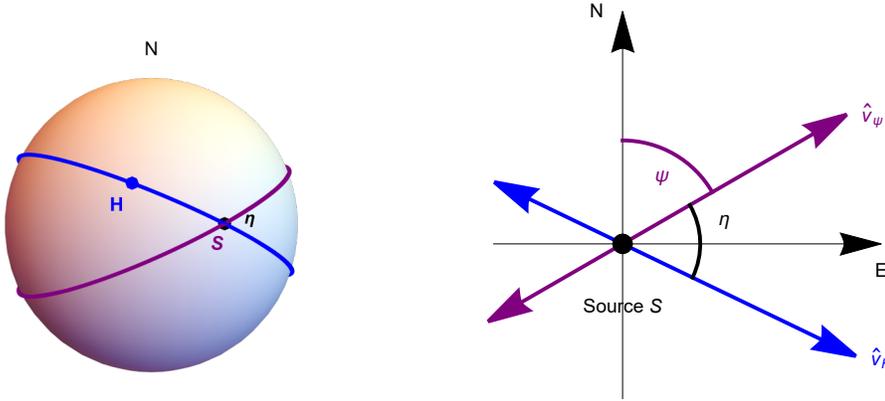

Figure 1: The Celestial sphere is pictured on the left and on the right is the plane tangent to the sphere at the source S. The linear polarization direction $\hat{v}_\psi$ lies in the tangent plane and determines the purple great circle on the sphere. A point H on the sphere together with the point S determine a second great circle, the blue circle drawn on the sphere. Clearly, H and S must be distinct in order to determine a great circle. The angle η measures the alignment of the polarization direction ψ with the point H.

With N sources S_i, $i = 1, ..., N$, there are N alignment angles η_{iH} at each point H . One can calculate an average alignment angle $\bar{\eta}$ at H,

$$\bar{\eta}(H) = \frac{1}{N} \sum_{i=1}^{N} \eta_{iH} , \qquad (1)$$

where

$$\cos(\eta_{iH}) = | \hat{v}_\psi . \hat{v}_H | , \qquad (2)$$

$$\cos(\eta_{iH}) = \left| \frac{\cos\delta_S \cos\psi \sin\delta_H + \cos\delta_H \left[\sin(\alpha_H - \alpha_S) \sin\psi - \cos(\alpha_H - \alpha_S) \cos\psi \sin\delta_S \right]}{\sqrt{1 - \left(\cos(\alpha_H - \alpha_S) \cos\delta_H \cos\delta_S + \sin\delta_H \sin\delta_S \right)^2}} \right| , \qquad (3)$$

where α and δ are the longitude and latitude of the ith source S and the point H. Each angle η_{iH} is taken to be the acute angle solving (2) or (3). Then the average alignment angle $\bar{\eta}(H)$ at the point H must also be acute.

The alignment angle $\bar{\eta}(H)$ is a function of position H on the sphere. See Figs. 2 and 3. It is symmetric across diameters, $\bar{\eta}(H) = \bar{\eta}(-H)$, because H and $-H$ lie on the same great circle through S. The points opposite S and H are located at $(\alpha,\delta,\psi) \rightarrow (\alpha \pm \pi, -\delta, \pi - \psi)$. Substituting into the formula on the right in Eq. (3), and simplifying, shows that each angle η_{iH} is invariant when S moves to $-S$ or when H moves to $-H$. It follows that the sum in Eq. (1) and the alignment angle function $\bar{\eta}(H)$ is invariant across diameters as well.

For random polarization directions, the average $\bar{\eta}(H)$ should be near 45°, since each alignment angle η_{iH} is acute, $0° \leq \eta_{iH} \leq 90°$, and random polarization directions should not favor large values or small values of η_{iH}, and, therefore, average to about 45°. When

the sum in Eq. (1) on the right is a sum over random angles η_{iH} scattered above and below 45°. By Random Walk theory, the sum of $(\eta_{iH} - 45°)$ goes like $N^{1/2}$, so $\bar{\eta}(H)$ should differ from 45° by an amount proportional to $N^{-1/2}$. Therefore as N grows larger for randomly directed samples, the average $\bar{\eta}(H)$ should approach nearer to 45°. This is confirmed by the Library, see Figs. 8, A1, A3.

Points H where the average alignment angle $\bar{\eta}(H)$ is smaller than 45°, the great circles tend to converge and where the angle $\bar{\eta}(H)$ is larger than 45°, the great circles can be said to diverge. The extremes of the function $\bar{\eta}(H)$ measure extreme convergence and extreme divergence of the great circles determined by the polarization directions.

In this article and notebook, we often use "min" to label the smallest alignment angle $\bar{\eta}_{min}$, the minimum value of the function $\bar{\eta}(H)$, Eq. (1). We have $\bar{\eta}_{min} = \bar{\eta}(\pm H_{min}) \leq \bar{\eta}(H)$, for all H. The associated points on the Celestial Sphere are the "alignment hubs" H_{min} and $-H_{min}$. Thus "min" is associated with convergence of the polarization directions. For divergence, the "avoidance hubs" H_{max} and $-H_{max}$ locate places where the polarization directions most avoid, as indicated by the largest alignment angle $\bar{\eta}_{max}$, the maximum value of the function $\bar{\eta}(H)$. We have $\bar{\eta}_{max} = \bar{\eta}(\pm H_{max}) \geq \bar{\eta}(H)$, for all H. Thus, we very often label an avoidance related quantity with "max".
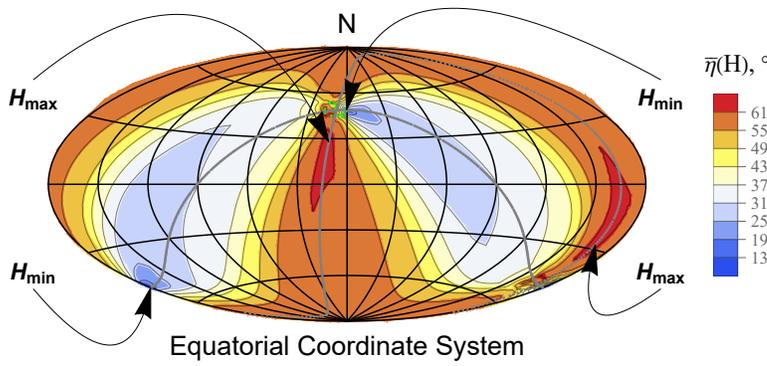

Figure 2: For the sample of 13 QSOs in Ref. 4, the alignment angle function $\bar{\eta}(H)$ mapped on the Celestial Sphere (Aitoff plot, centered on $(\alpha,\delta) = (180°,0)$, East to the right). The QSOs are shaded green ▪. The largest avoidance angle, $\bar{\eta}_{min} = 62.7°$, is located at the hubs H_{max} and $-H_{max}$. The smallest alignment angle, $\bar{\eta}_{min} = 10.86°$, is located at the hubs H_{min} and $-H_{min}$, where the polarization directions converge best. One of the two alignment hubs H_{min} is located very close to the QSOs.
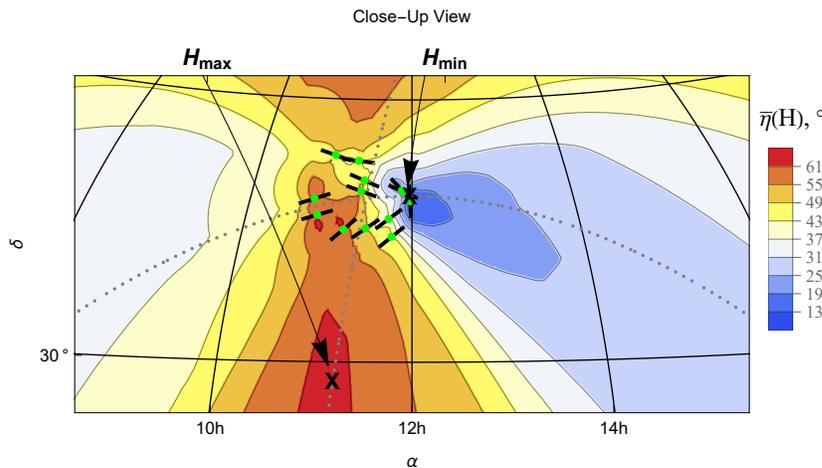

Figure 3: The region near the QSOs in Fig. 2. The QSOs are located at the green dots. The short black lines through the QSOs

indicate the polarization directions. Measuring polarization directions ψ clockwise from North toward East, one sees that the angles ψ range from more than $\psi = 90°$ for the northern-most QSOs to 45° or so for the southerly QSOs. Thus, the QSOs display parallax: all are in the general direction of the alignment hub H_{min}, but their directions depend on where they are located.

3. Significance, Direct Method A.

Once a measure of alignment is made, it is important to judge its significance. The significance of the smallest alignment angle $\overline{\eta}_{min}$ is defined as the likelihood that randomly directed polarization vectors would produce a smaller value of $\overline{\eta}_{min}$. By this definition, one way to determine significance is to repeat the process of making Great Circles from random polarization directions, calculating the alignment function $\overline{\eta}(H)$, and finding $\overline{\eta}_{min}$, all for randomly directed vectors. One such process makes a "random run".

The most reliable method of determining significance that we consider is called "Direct Method A". One generates many random runs with randomly directed transverse vectors assigned to the sources, replacing the observed polarization directions. A histogram of the random-based results for $\overline{\eta}_{min}$ is then approximated by a suitable fitting function. Aside from a scale factor that normalizes the distribution, the fitting function of the histogram is the probability distribution of the random results $\overline{\eta}_{min}$. Having found a function that approximates the probability distribution, one calculates the likelihood that random runs return a smaller value than the observed $\overline{\eta}_{min}{}^{obs}$ and that is the significance of $\overline{\eta}_{min}{}^{obs}$.

One of the samples we discuss here has 27 QSOs and has been treated elsewhere, Ref. 3. The significances of $\overline{\eta}_{min}$ and $\overline{\eta}_{max}$ for the 27 QSO sample were found by generating 10,000 random runs, Direct Method A. The histograms are displayed in Fig. 4. The histogram for the smallest alignment angle $\overline{\eta}_{min}$ on the left in Fig. 4 and the histogram on the right for the largest avoidance angle $\overline{\eta}_{max}$ are fit by functions proportional to the probability distributions.

There is much in common between convergence and divergence, $\overline{\eta}_{min}$ and $\overline{\eta}_{max}$, both are extremes of the alignment angle function $\overline{\eta}(H)$. For brevity and because mutual alignment is of central interest in the literature, convergence is emphasized in this part, Part I the Article. Both are treated equally in Part II the Appendix.
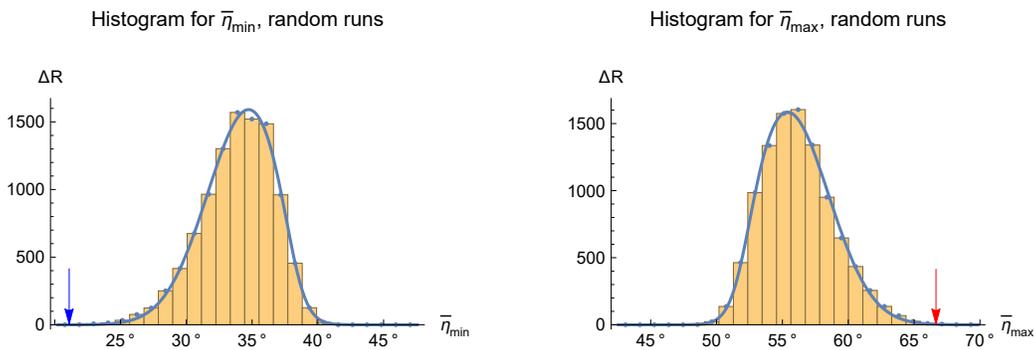

Figure 4: Direct Method A applied to a sample of 27 QSOs, Ref. 3. Histograms for both the smallest alignment angle $\overline{\eta}_{min}$ and the largest avoidance angle $\overline{\eta}_{max}$ are displayed. Adding the heights of the bars in a histogram yields the total number of random runs, here $\Sigma \, \Delta R = R = 10{,}000$ runs. The height ΔR of each bar is the number of runs with results in the indicated bin, so the fraction $\Delta R/R$ is the probability that a random run value of $\overline{\eta}_{max}$ lands in the bin. Once normalized, the solid curve is the probability distribution obtained by fitting the fractions $\Delta R/R$ by Eq. (6). The blue and red arrows locate the values determined with the observed polarization directions. Very few random runs do better than the observed data, so the observed polarization directions are very significantly correlated, both for convergence and divergence.

4. Creating a Library of Random Run Results

One way to avoid generating and analyzing thousands of randomly directed samples is to generate and analyze thousands of artificially created simulated samples beforehand and collect parameters sufficient to estimate their probability distributions. With a sufficiently wide range of samples, the parameters for a particular observed sample can be found by interpolation. The collection of

parameters can be called a "Library". The process is called "Interpolation Method B".

To begin with, we must create a set of samples to analyze. We arrange the sources in square arrays, so that they are spread out uniformly over the region. This forces the number of sources N to be a perfect square. Currently, the Library has fourteen values of N ranging from $N = 9$ to $N = 900$, in detail:

$$N = 9, 16, 25, 36, 49, 64, 81, 100, 121, 225, 256, 324, 625, 900 \ . \qquad (4)$$

It is convenient to pretend that the square samples can be approximated as circular regions with a "radius". Thus, to this end, we build the square arrays using a nominal radius, "ρNominal". The size of each square is determined by averaging two square arrays, a large square with the nominal circle inscribed and a smaller square that has vertices on the nominal circle. The side of the square array for the simulated sample is the average of the inscribed and superscribed squares. Thus we choose values for the radii ρNominal and construct square arrays of samples. As of this writing,

$$\rho\text{Nominal} = \tfrac{1}{7}°, \tfrac{1}{6}°, \tfrac{1}{5}°, \tfrac{1}{4}°, \tfrac{1}{3}°, \tfrac{1}{2}°, 1°, \tfrac{3}{2}°, 2°, 3°, 4°, 6°, 8°, 10°, 15°, 16°, 24°, 32°, 42°, 52°, 64°, \quad (5)$$

twenty one values in total.

For calculations and comparing different samples, a suitable "radius" measure may be the root-mean-square radius ρRMS. The RMS radius is a well-recognized, often employed, measure of the effective size of a region containing sources. For large regions, there can be a problem due to the symmetry of η in Fig. 1 across a diameter. By Eq. (3), a source S located at (α,δ,ψ), with longitude α, latitude δ, polarization direction ψ, has the same alignment angle η_{iH} with point H as the diametrically oppose source $-S$ at $(\alpha + 180°, -\delta, \pi-\psi)$. When S is beyond 90° from the sample's center, the virtual source at $-S$ is closer to the sample center than 90° and the root-mean-square radius ρRMS calculated with S should be instead calculated with $-S$. Therefore, in constructing simulated samples for the Library, we should keep all sources in the square arrays within 90° of the sample's center point on the Celestial Sphere. Well, we should, but we don't. Some of the most far flung sources for ρNominal $= 64°$ are a couple of degrees beyond 90°. We ignore this discrepancy.

For observed data, the use of the Library is complicated by this problem of calculating ρRMS for an observed sample that extends more than 90° from the sample's center on the Celestial Sphere. For example, one might want to apply the Library methods to a whole sky sample. In such cases the sources should be gathered into one hemisphere by creating equivalent, though virtual, sources. If an actual source S is more than 90° from the sample's center, replace the longitude, latitude and polarization directions (α,δ,ψ) for a virtual equivalent source by $(\alpha \pm \pi, -\delta, \pi-\psi)$, the corresponding values on the opposite point $-S$ on the Celestial Sphere.

This is not necessary with Direct Method A, because Direct Method A does not use the Library and the root-mean-square radius ρRMS is not needed with Direct Method A. But, Interpolation Method B and Formula Method C need a value of ρRMS to compare with the Library data. Moving sources to a single hemisphere may be necessary to obtain good results with Methods B and C.

For the Library, we calculate the RMS radius of the square arrays that we construct for the Library. The RMS radii ρRMS are then used in calculations. However, we organize the samples by the nominal radius ρNominal. Fig. 5 displays one of the samples constructed for the Library. It has a nominal radius of 16°, *i.e.* ρNominal $= 16°$, but the RMS radius is smaller, ρRMS $= 12.5°$.
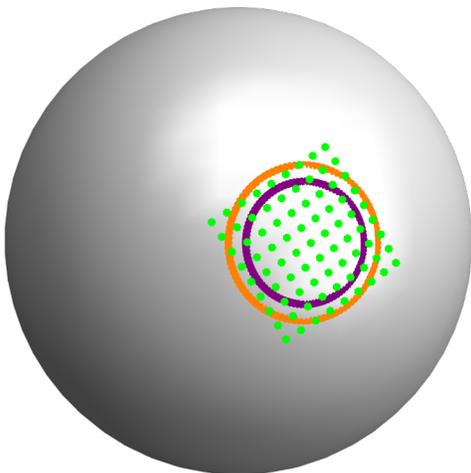
Figure 5: A random run sample with $N = 81$ sources, here shaded green, and a nominal radius ρNominal $= 16°$, the orange circle, and a root-mean-square radius ρRMS $= 12.5°$, the purple circle. The sample is rotated to avoid coincidences with coordinate axes. The length of the side of the square is set so that the square approximates a circle with the nominal radius. However, for calculations and comparison with any observed samples, the root-mean-square radius is considered to be more appropriate.

In total, there are 14 values for N and 21 values of ρNominal, so there are 294 combinations (N, ρNominal), each with its own, unique ρRMS.

Once the samples are created, the sources are given randomly directed transverse vectors. The function $\bar{\eta}(H)$, Eq. (1), is calculated and the smallest alignment angle $\bar{\eta}_{min}$ and the largest avoidance angle $\bar{\eta}_{max}$ are determined. At the time of writing, there are 294 samples and each is processed with 10,000 or more random runs. Some of the histograms for the square array of $N = 25$ sources are displayed in Fig. 6 and 7.
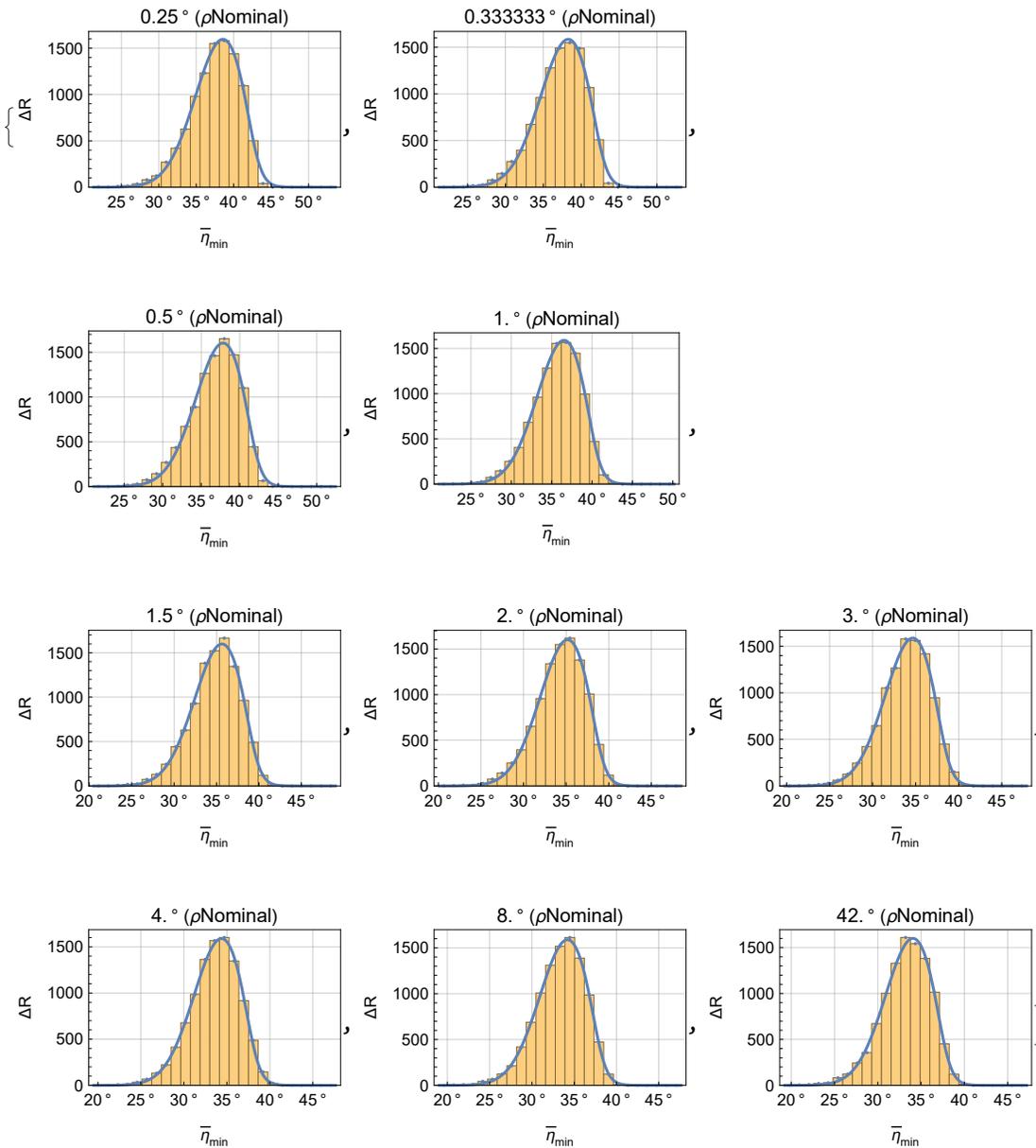
Figure 6: Thumbnails of ten histograms for the values of $\bar{\eta}_{min}$ from random runs with $N = 25$ and various nominal radii. The plots are labelled with the nominal radius ρNominal. Functions, Eq. (6), that fit the distributions are plotted as blue lines. Note that all the

distributions tilt as if to avoid $\bar{\eta}_{\min} = 45°$. They are distorted from a symmetric Gaussian shape.

A glance at Fig. 6 above reveals that the distributions for $\bar{\eta}_{\min}$ are steeper on the high side toward $\eta \to 45°$. This behavior is repeated for all such distributions and appears to be an inherent property of the statistics.
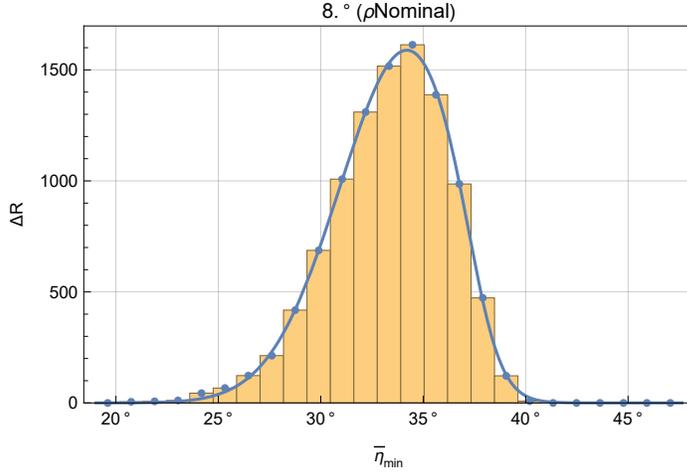
Figure 7: Enlarged view of the histogram and fit for $N = 25$ and ρNominal $= 8°$, the distribution of the $R = 10,000$ values of the smallest alignment angle $\bar{\eta}_{\min}$ from the random runs. As noted above with Fig. 4, the histogram approximates the shape of the probability distribution, aside from a normalizing scale factor. These distributions are fit by a Gaussian multiplied by a step function. The step function vanishes to the right of the peak and is unity on the more important left side where the smallest alignment angles $\bar{\eta}_{\min}$ for well-aligned samples are found. Thus the tail of the curve for small $\bar{\eta}_{\min}$ drops off like a Gaussian on the important low side of the distribution.

A histogram of the smallest alignment angles $\bar{\eta}_{\min}$ from R $= 10,000$ runs is enlarged from Fig. 6 and displayed in Fig. 7. The height of a bar in Fig. 7 is the number ΔR of the random runs that gave a value of $\bar{\eta}_{\min}$ in the interval $\delta\eta$, the width of a bin. Thus, the quantity $\Delta R/R$ is the likelihood that random runs give a value of $\bar{\eta}_{\min}$ in the bin. The histogram has the shape of the probability distribution $P_{\min}(\eta)$ for $\bar{\eta}_{\min}$. The probability distribution is normalized, $1 = \int P_{\min} d\eta$. Since the sum $\Sigma \Delta R = R$, it follows that $1 = \Sigma \frac{\Delta R}{R \delta \eta} \delta\eta$ and that $P_{\min}(\eta) \approx \frac{\Delta R}{R \delta \eta}$. Therefore, by fitting the histogram, we obtain the probability distribution $P_{\min}(\eta)$, within a numerical factor.

There is a complication. Look closely at the distribution in Fig. 7. The right side, the side toward $\bar{\eta}_{\min} \to 45°$, has a steeper slope than the left side, the side toward $\bar{\eta}_{\min} \to 0$. Thus, the low $\bar{\eta}_{\min}$ side is favored and probability is pushed from the right side to the left side. A simple, symmetrical Gaussian would not fit the data well. The fitting curve shown in blue in Fig. 7 combines a Gaussian with a unit step-function, that is unity to the left of the peak, and zero to the right. We have

$$P_{\min}(\eta) = \left(\frac{\text{norm}}{\sigma \, (2\pi)^{1/2}} \right) \left(1 + e^{4 \frac{(\eta - \eta 0 - \sigma)}{\sigma}} \right)^{-1} e^{-\frac{1}{2} \left(\frac{\eta - \eta 0}{\sigma} \right)^2}, \qquad (6)$$

where "norm" is a scale factor that normalizes the probability density. One finds that norm ≈ 1.22.

There are two free variables that can be chosen to fit a random run histogram like the one shown in Fig. 7. We call η_0 the 'peak' value and σ is the 'half-width'. The presence of the step function $\left(1 + e^{4 \frac{(\eta - \eta 0 - \sigma)}{\sigma}} \right)^{-1}$ moves the peak and half-width a little from their Gaussian values without the step function. It is remarkable that the fitting function has two parameters just like a Gaussian, namely the peak η_0 and the half-width σ. The step function is introduced without an associated parameter.

Given the form in Eq. (6), the problem of determining the significance of an observed $\bar{\eta}_{\min}$ reduces to estimating the values of two parameters, η_0 and σ, since those two values determine the probability distribution $P_{\min}(\eta)$. Once determined, the probability

distribution can be integrated to obtain the significance of an observed $\bar{\bar{\eta}}_{\min}$.

   For small N one confronts a common issue with distribution formulas. Probability distributions like that in Eq. (6) return non-zero probabilities for unphysical values of the variables. The acute angle $\bar{\bar{\eta}}_{\min}$ can not be negative even though the probability distribution in Eq. (6) assigns a probability to negative values of η. The problem is unimportant when there are a sufficient number of sources because low values of $\bar{\bar{\eta}}_{\min}$ are then unlikely. With more sources, the likelihood of alignment decreases, as is evident from the $\bar{\bar{\eta}}_{\min}$ plot in Fig. 8 top. Thus, having valid statistics requires setting a minimum number of sources; we take 7. The number of sources N must be at least 7, $N \geq 7$ for the significance of negative $\bar{\bar{\eta}}_{\min}$ to be sufficiently small.

   To find the significance, or p-value, of the smallest alignment angle $\bar{\bar{\eta}}_{\min}$, one finds the likelihood of smaller random run values by integrating $P_{\min}(\eta)$ from below,

$$p(\bar{\bar{\eta}}_{\min}) = \int_{-\infty}^{\bar{\bar{\eta}}_{\min}} P_{\min}(\eta) \, d\eta \qquad . \qquad\qquad (7)$$

It is unfortunate, but conventional, to denote the significance with the letter "p." By Eqs. (6) and (7) the significance p of the correlated behavior indicated by the smallest alignment angle $\bar{\bar{\eta}}_{\min}$ rests on the two probability distribution parameters $\eta_0^{\min}$ and $\sigma^{\min}$.
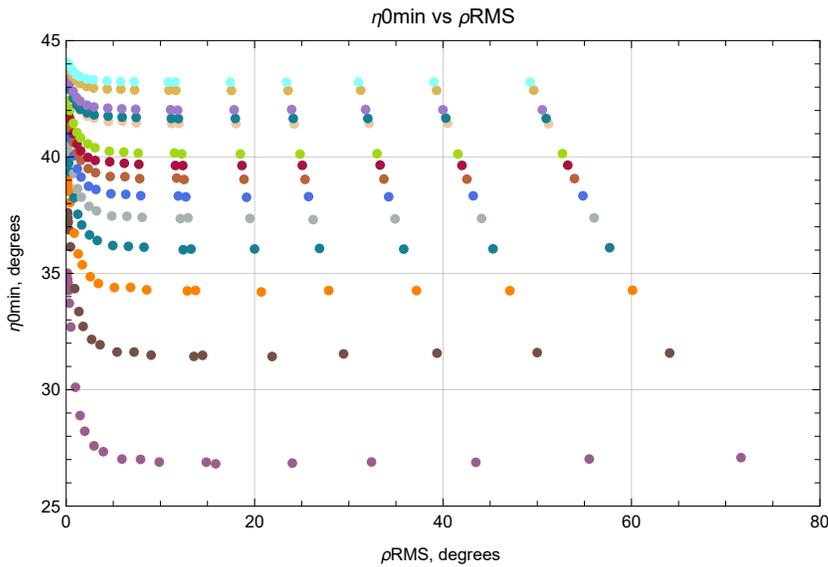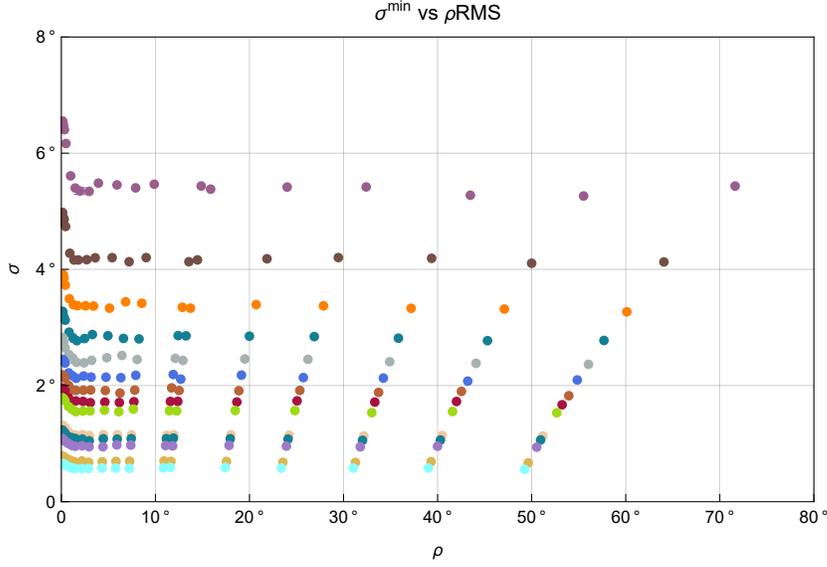
$$\{N, \text{color}\} = \{\{9, \blacksquare\}, \{16, \blacksquare\}, \{25, \blacksquare\}, \{36, \blacksquare\}, \{49, \blacksquare\}, \{64, \blacksquare\}, \{81, \blacksquare\},$$
$$\{100, \blacksquare\}, \{121, \blacksquare\}, \{225, \blacksquare\}, \{256, \blacksquare\}, \{324, \blacksquare\}, \{625, \blacksquare\}, \{900, \blacksquare\}\}$$

Figure 8: The Library data for the two parameters of probability distributions for the smallest alignment angle $\bar{\eta}_{\min}$ of $\bar{\eta}(H)$. The Library consists of the values of η_0 and σ in Eq. (6) for the random run distributions of $\bar{\eta}_{\min}^{\text{random}}$, such as Fig. 7. *Top.* The parameter $\eta_0^{\min}$ locates the peak of a random run distribution. For a given number of sources N, the parameter $\eta_0^{\min}$ curves upward for ρRMS less than about 4°, ρRMS $\lesssim 4$ °. The standard errors of the $\eta_0^{\min}$ are smaller than the dots and are not displayed. *Bottom.* Similar behavior is found for the half-width $\sigma^{\min}$, with constant values at large ρRMS and curving upward below about 4°.

One motivation for emphasizing w, with $w = N^{-1/2}$, is the use of square arrays of sources, as illustrated in Fig. 5. The quantity N is associated with an area. By taking the square root, *i.e.* $N^{1/2}$, one obtains a 'linear' quantity. Then, both $N^{1/2}$ and ρRMS are associated with angular arc lengths. A reason for inverting $N^{1/2}$ to make $w = N^{-1/2}$ is the fact that N is unbounded $N \longrightarrow \infty$, while w is bounded, $0 \le w \le 1/3$, since we must have N greater than 9 for the statistics to be sufficiently accurate and for N to be a perfect square. Yet another reason follows by noting that the sum in (1) is a sum over random values for random runs. A sum over random values evokes the mathematics of random walks and random walks involves $N^{1/2}$, where N is analogous to the number of steps in a Random Walk.

The root-mean-square radius ρRMS is already bounded, $0 \le \rho$RMS $\le 70°$, since a radius of 90° makes a complete hemisphere and then diametric symmetry covers the sphere. On occasion, we will divide ρRMS by 90° to make a variable rRMS that varies from 0 to 1, $0 \le$ rRMS ≤ 1. Power series approximations may be better contained when the variable itself is bounded.

The variables that characterize a sample are

$$N \quad \text{or} \quad w \quad \text{and} \quad \rho\text{RMS or rRMS}, \qquad\qquad (8)$$

where

$$w := N^{-1/2} \quad \text{and} \quad \text{rRMS} := \rho\text{RMS}/90° \quad . \qquad\qquad (9)$$

The two quantities w and rRMS have ranges from zero to some value less than one, $w \le 9^{-1/2} = 1/3$ and rRMS $\lesssim 70°/90° \approx$ 0.8, making them suitable for the power series in the formulas of Formula Method C. First we simply interpolate the Library to find the parameters η_0 and σ of the probability distributions.

5. Interpolation Method B

Interpolation Method B is more convenient than Direct Method A. Given experimentally observed data, an interpolation is a quick way to get values of the peak η_0 and the half-width σ in the probability distribution Eq. (6). For an observed sample whose

number of sources N^{obs} and RMS radius ρRMS^{obs} are in the range of the Library's resources, an interpolation of the Library data produces values of η_0 and σ.

The Library data displayed in Fig. 8 above and Figs. A1 - A4 in Part II the Appendix. With Eq. (6), these values determine the probability distribution $P_{min}(\eta)$ for the observed sample. By integrating the distribution $P_{min}(\eta)$ over all η less that the observed sample's $\bar{\eta}_{min}{}^{obs}$, one finds the significance or p-value, $p(\bar{\eta}_{min})$, of the smallest alignment angle $\bar{\eta}_{min}{}^{obs}$ calculated for the observed sample, as in Eq. (7).

Method B may not give reliable parameters η_0 and σ when the number of sources N^{obs} and/or the RMS radius ρRMS^{obs} is/are outside the range of Library data. In such cases, asking Mathematica for an interpolation produces an extrapolation instead. While extrapolation may be less reliable than interpolation, it does yield values of η_0 and σ allowing one to determine a probability distribution $P_{min}(\eta)$.

It might be better to consider Formula Method C whenever the observed sample's N and ρRMS leads to an extrapolation of the Library data to find η_0 and σ. And Formula Method C may be found to be more convenient than setting up the interpolations needed with Method B.

6. Fitting the Library, Formula Method C

"Formula Method C", like Interpolation Method B, is based on the Library data in Fig. 8. Thus, both Methods B and C avoid the time-consuming random-runs needed with Direct Method A.

Formula Method C finds functions of $w = N^{-1/2}$ and ρRMS that fit the two parameters η_0 and σ listed in Library data table. Then, with these two functions, one substitutes the N and ρRMS of the sample into the formulas and that gives values for the peak η_0 and half-width σ. The two parameters are all that is needed by Eq. (6) to construct the probability function P_{min} and then an integration yields the significance of an observed smallest alignment angle $\bar{\eta}_{min}{}^{obs}$.

Therefore, we need to find functions $\eta_0{}^{min}(w, \rho RMS)$ and $\sigma^{min}(w, \rho RMS)$ that fit the Library data in Fig. 8. Looking closely at the left graph in Fig. 8, one sees that $\eta_0{}^{min}$ is relatively flat for large ρRMS and curves upward sharply at small ρRMS. This behavior is reminiscent of the hyperbolic tangent, $\tanh(u) = \sinh(u) / \cosh(u)$, displayed in Fig. 9.
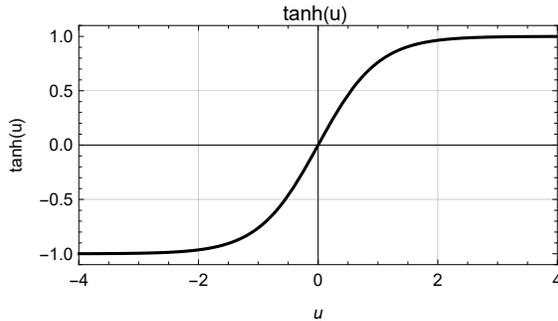
Figure 9: The hyperbolic function tanh(u). The $\eta_0{}^{min}$ and σ^{min} data in Fig. 8 may be fit by inverting, rescaling and translating as necessary the tanh(u) function displayed here. One can fit the Library data for $\eta_0{}^{min}$ to about 0.2% and σ^{min} to about 0.8%, see Fig. 10.

By contorting the tanh(u) with an application of polynomials and varying their coefficients to get a least squares fit to the Library data, we get the following functions. See Part II the Appendix, Sec. A4.

$\eta_0{}^{min}(w, \rho RMS) =$

$$44.9986 + w\,\big(26.9406 + 3.08965\,w - \tag{10}$$
$$(-25.9516 - 0.00244035\,\rho RMS)\,\mathrm{Tanh}\big[0.718833\,\rho RMS + 0.25621\,w\,\rho RMS - 0.00771145\,\rho RMS^2\big]\big)$$

$\sigma^{min}(w, \rho RMS) = 0.25\,w\,(74.5043 - 5.83586\,w + (4.07678 + 10.7564\,w)\,\mathrm{Tanh}[1.33841 - 2.31346\,\rho RMS])\,,$ $\qquad$ (11)

where w $= N^{-1/2}$ and the units for $\eta_0{}^{\min}$ and $\sigma^{\min}$ are degrees.

The fit to the Library values of $\eta_0{}^{\min}$ in the Left plot of Fig. 10 is better than the fit to the $\sigma^{\min}$ on the Right. In Part II the Appendix, it is found that the percent difference of the fit of $\eta_0{}^{\min}(w, \rho\text{RMS})$ in Eq. (10) to the Library data averages 0.13%, where by 'average' we mean an arithmetic mean, and the % diff. is always better than 1.1%. For $\sigma^{\min}(w, \tau)$, one finds that the functions fit the Library to 4.4% or better, with an average of 0.93%. Thus, the 8-parameter fit in Eq. (10) fits 294 Library data points of $\eta_0{}^{\min}$ to a few tenths of a percent, while the 6-parameter function in Eq. (11) fits the 294 values of $\sigma^{\min}$ to a percent or so. See Fig. 10 and the results preceding Figs. A6 and A11 in Part II the Appendix.
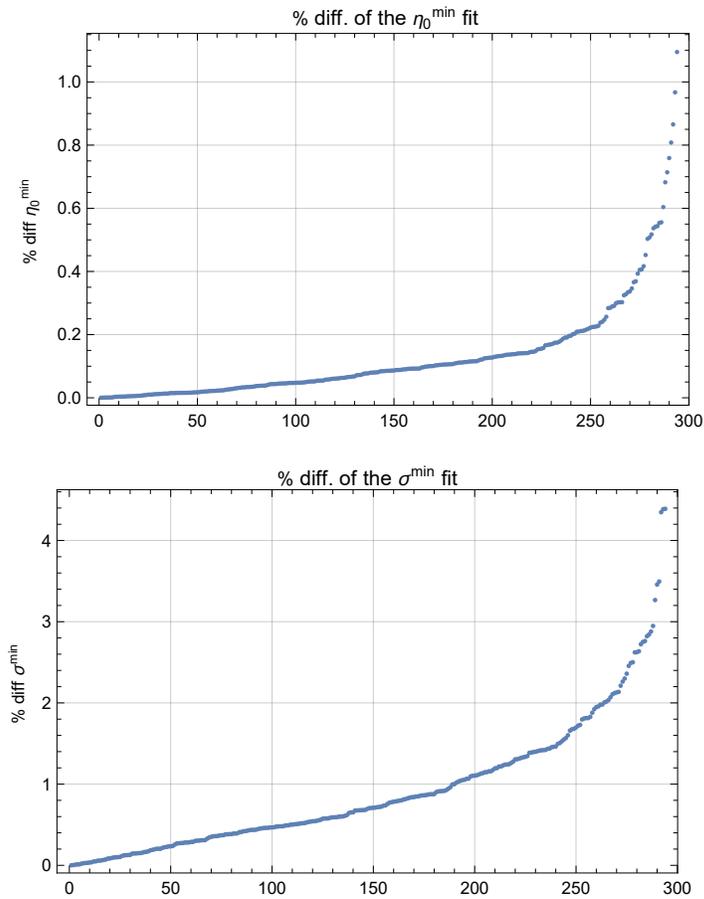
Figure 10. The percent difference between the Library data and the fitting functions for $\eta_0{}^{\min}$ and for $\sigma^{\min}$.
For $\eta_0{}^{\min}$, the mean % diff is 0.13% and, for $\sigma^{\min}$, the mean % diff is 0.93%. Their standard deviations are 0.16% and 0.81%, respectively.

Graphing the functions in Eqs. (10) and (11), we have
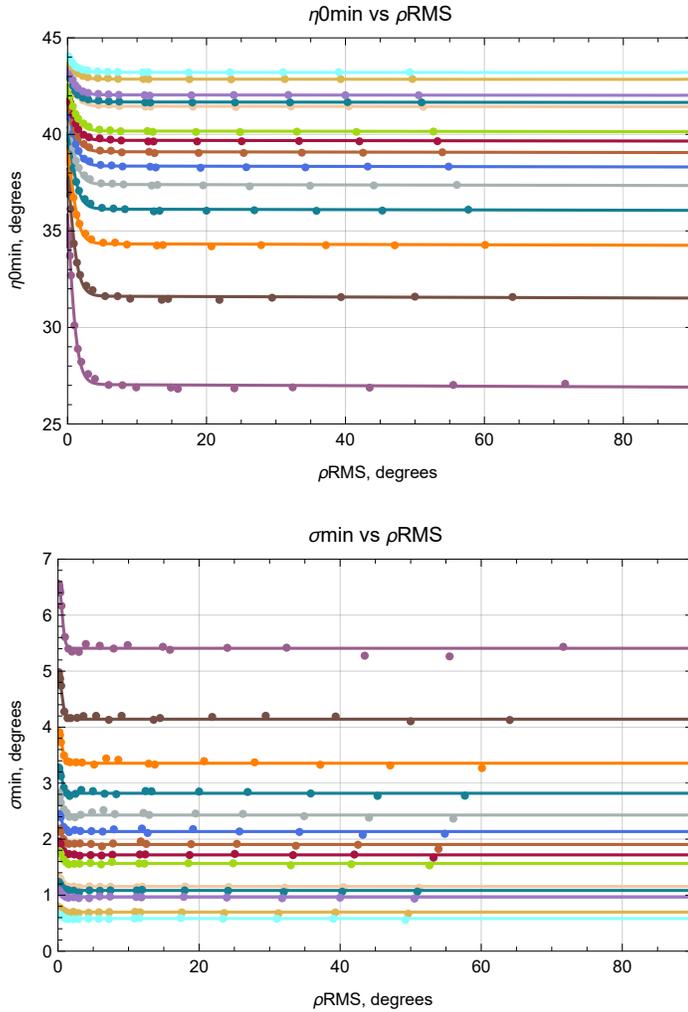
N:{{9, ▨}, {16, ▨}, {25, ▨}, {36, ▨}, {49, ▨}, {64, ▨}, {81, ▨},
{100, ▨}, {121, ▨}, {225, ▨}, {256, ▨}, {324, ▨}, {625, ▨}, {900, ▨}}

Figure 11: Fitting functions of $\eta_0{}^{\min}$ (*Top*) and $\sigma^{\min}$(*Bottom*), the two parameters for the probability distribution in Eq. (6) for the smallest alignment angle $\bar{\eta}_{\min}$. The curves are flat at large angles, and, for all angles, the curves flatten out for large *N*. The angle 45° is an upper bound for $\eta_0{}^{\min}$ that is approached by large randomly directed samples. Large numbers of randomly directed sources are exceedingly unlikely to align by chance.

With Formula Method C, we now have three methods to determine the parameters $\eta_0{}^{\min}$ and $\sigma^{\min}$ for the probability distribution Eq. (6) needed to determine significance or *p*-value in Eq. (7). Let us see how the three methods work in practice with three samples.

7. Applying the Three Methods to Three Sets of Experimental Data

In previous work, Refs. 3,4,5, we looked at two samples of radio QSOs and one sample of polarized light from Milky Way Galaxy stars. The QSO samples had 27 and 13 radio sources each, while there were 99 stars in the polarized light sample. The data analyzed in the QSO studies are from a catalog JVAS1450, Ref. 9. The JVAS1450 catalog adds object-specific data to a subset of an

earlier catalog JVAS/CLASS 8.4-GHz, Ref. 10. The observations of the 99 Stars sample are contained in two catalogs, Heiles 2000 and Berdyugin 2014, Refs. 11,12.

In the previous work, all three samples had their significance determined by applying Direct Method A. For each sample, 10,000 random runs were generated with random directions replacing the polarization directions. We used the observed sites of the sources. The histogram distributions for the 10,000 smallest alignment angle $\bar{\eta}_{\min}^{\text{random}}$ were fit with functions of the form in Eq. (6). Each distribution gave values for the two adjustable parameters for the fitting function, the value of $\bar{\eta}_{\min}^{\text{random}}$ at the peak, $\eta_0^{\min}$, and the half-width $\sigma^{\min}$ where the distribution is down by a factor of $e^{-1/2}$ from the peak value. The terminology is accurate for Gaussian distributions; the presence of a step-function in Eq. (6) shifts the values only slightly.

Sample	N	ρRMS, deg	$\bar{\eta}_{\min}$, deg
27 QSOs	27	6.82492	21.094
13 QSOs	13	4.72812	10.865
99 Stars	99	6.83803	7.007

Sample	$\eta_0^{\min}$, deg	$\sigma^{\min}$, deg
27 QSOs	34.923 ±0.029	3.272 ±0.034
13 QSOs	30.26 ±0.17	4.64 ±0.20
99 Stars	39.947 ±0.018	1.737 ±0.021

Table 1: Three previously studied samples of observed polarizations, Refs. 3,4,5. *Left*: The samples' properties important for the current article are the number of sources N and the root-mean-square radius of the sample ρRMS. The alignment correlation is judged on the calculated value of $\bar{\eta}_{\min}$, the smallest alignment angle of the function $\bar{\eta}(H)$, Eq. (1), for any point H.   *Right*: The significance of the sample's $\bar{\eta}_{\min}$ requires statistics. The distribution of $\bar{\eta}_{\min}$ in random runs depends on two parameters, the alignment angle at the peak $\eta_0^{\min}$ and the half-width $\sigma^{\min}$. From the numbers in the table on the right and Eq. (6), one can determine the probability distribution and significance of any value of $\bar{\eta}_{\min}$ for each sample.
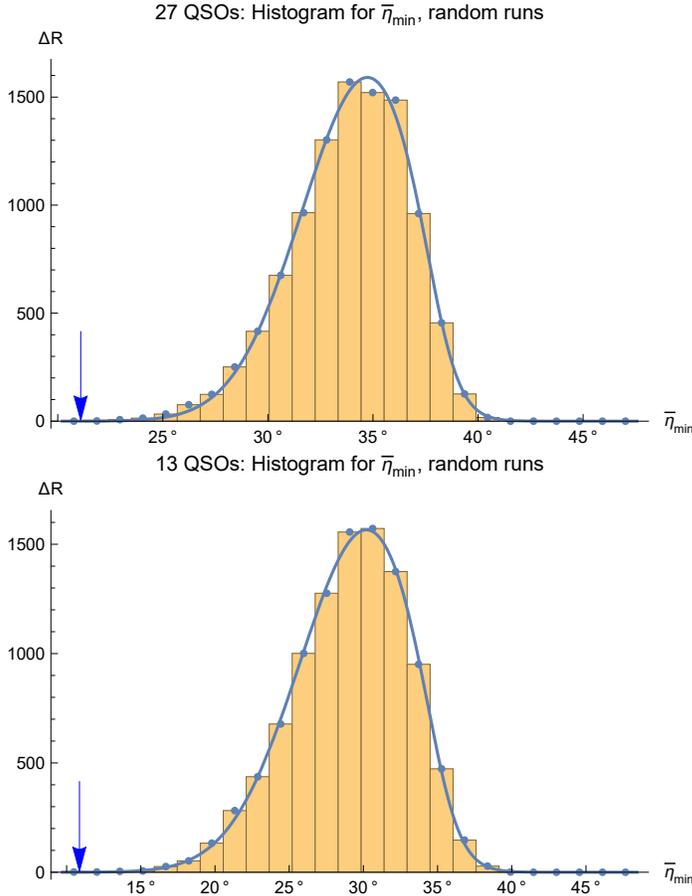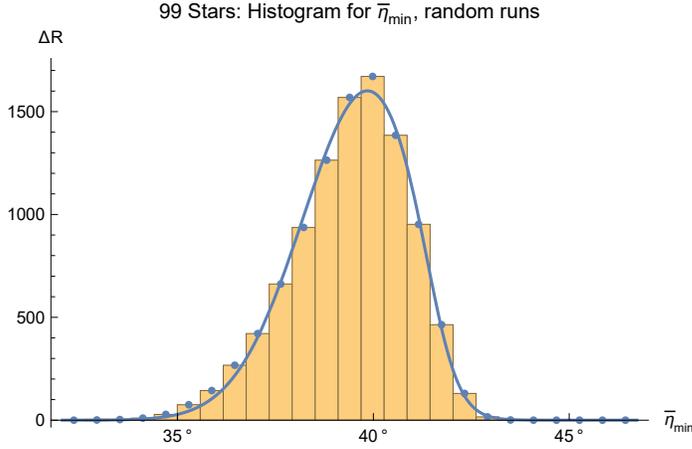
Figure 12: Histograms and fits for random runs found with Direct Method A for the three samples, 27 QSOs, 13 QSOs, and 99 Stars. The Blue arrows mark the observed smallest alignment angles $\overline{\eta}_{\min}$. The value of $\overline{\eta}_{\min}$ for the 99 Stars is off-scale at $\overline{\eta}_{\min} = 7.01°$. The fraction of the random run results to the left of the Blue arrows defines the significance of the observed smallest alignment angle $\overline{\eta}_{\min}$. The significance of the alignments for the 27 QSOs and the 13 QSOs is apparent, while $\overline{\eta}_{\min} = 7.01°$ for the 99 Stars is so far down the $\overline{\eta}_{\min}$-axis that the likelihood that random data would be better aligned is infinitesimal.

As discussed earlier, since Direct Method A uses the actual location of the sources and has the correct radius ρRMS with the correct number N of sources, Method A should be more reliable than Interpolation Method B or Formula Method C. Therefore, we consider Methods B and C inferior compared to a superior Method A.

All three methods rely on the same general probability distribution formula, Eq. (6). The methods differ in how the parameters for the location of the peak, $\eta_0{}^{\min}$, and the half-width, $\sigma^{\min}$, are calculated. Table 2 compares the values of $\eta_0{}^{\min}$ and $\sigma^{\min}$ from methods A, B and C.

(a) peak $\eta_0{}^{\text{min}}$:

Sample	A (degrees)	B (degrees)	C (degrees)
27 QSOs	34.923 ± 0.029	34.797 ± 0.015	34.74 ± 0.10
13 QSOs	30.26 ± 0.17	30.250 ± 0.031	30.14 ± 0.16
99 Stars	39.947 ± 0.018	39.678 ± 0.012	39.67 ± 0.05

(b) half-width σ^{min} :

Sample	A (degrees)	B (degrees)	C (degrees)
27 QSOs	3.272 ± 0.034	3.306 ± 0.018	3.23 ± 0.04
13 QSOs	4.64 ± 0.20	4.61 ± 0.04	4.56 ± 0.06
99 Stars	1.737 ± 0.021	1.729 ± 0.015	1.728 ± 0.015

Table 2: Probability distribution parameters $\eta_0{}^{\min}$ and $\sigma^{\min}$ in Eq. (6). The ± values are the standard errors reported in fitting the data with the Mathematica command NonlinearModelFit. (a) The values of $\eta_0{}^{\min}$ from Methods B and C are within 0.3% of the values by Method A, but not within the ± standard errors. (b) In contrast, since the ranges overlap, the ± standard errors for the half-widths $\sigma^{\min}$ appear adequate to describe the uncertainties .

Out[•]=
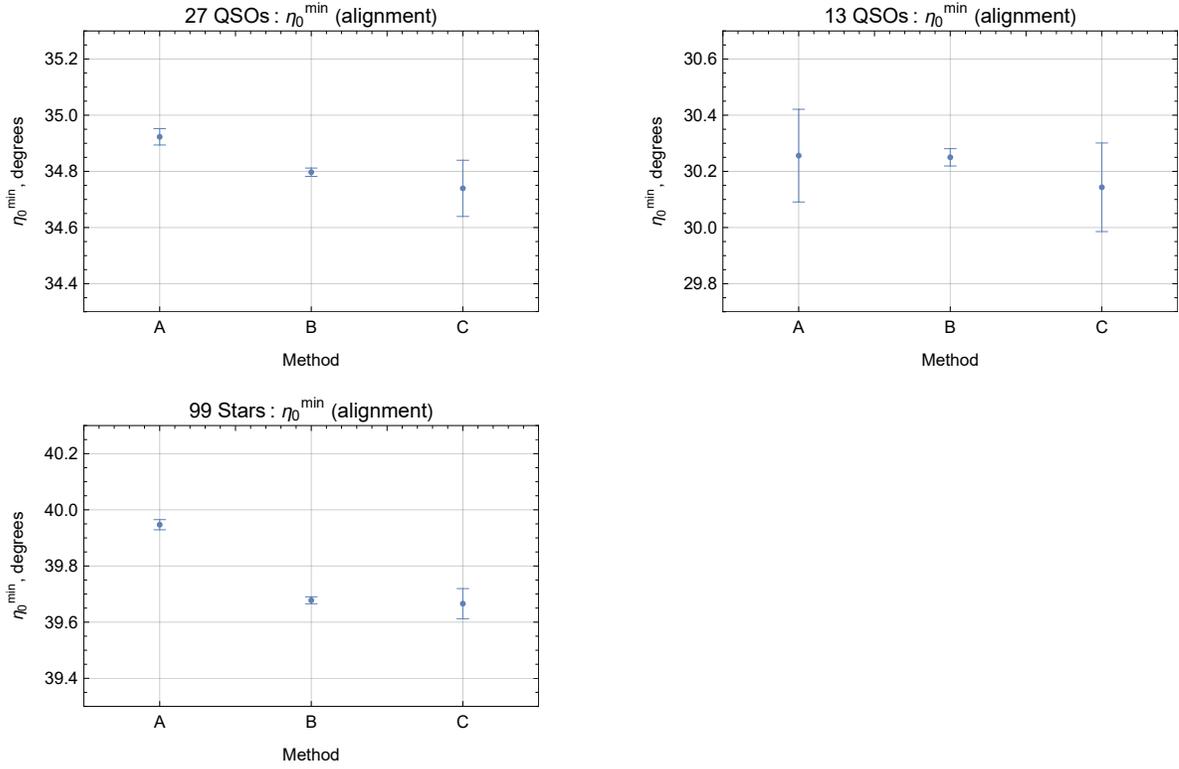
Figure 13: Compare parameter $\eta_0{}^{\min}$ for the 3 methods. Method B has small standard errors, so it is precise, but it is not accurate for the 27 QSOs and 99 Stars when compared to A.

Out[•]=
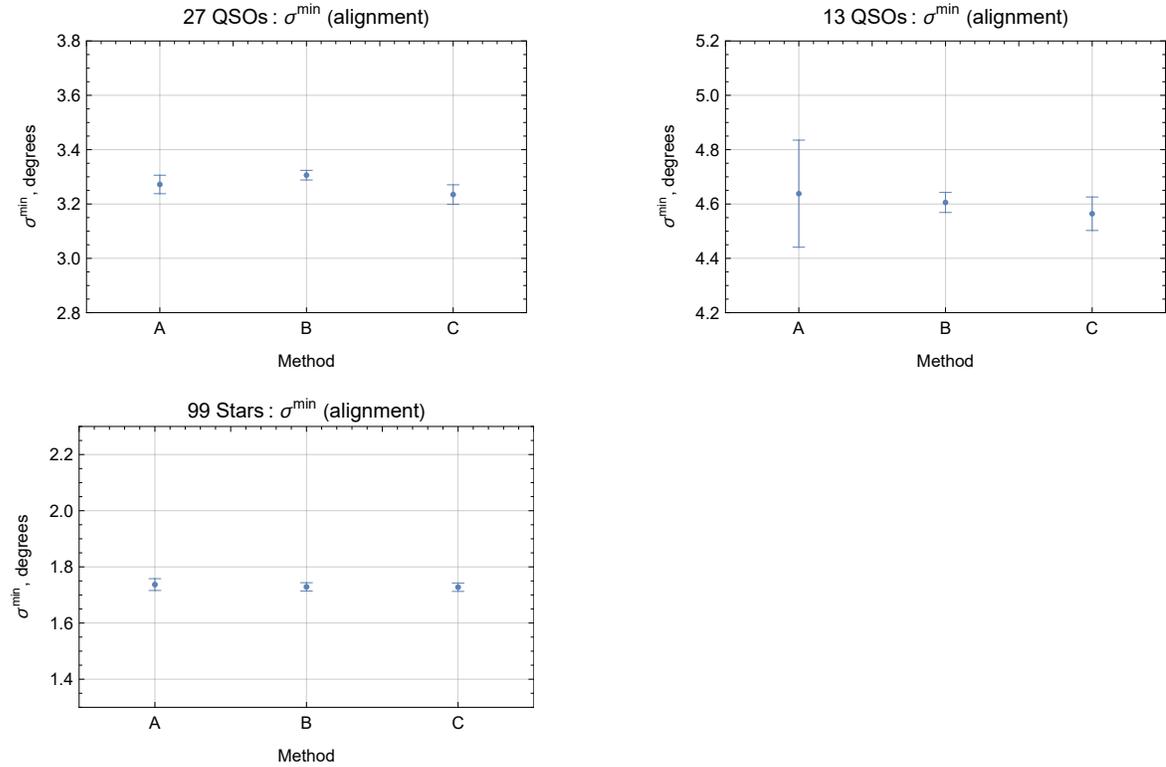
Figure 14: Compare parameter $\sigma^{\min}$ for the 3 methods with 3 samples. Methods B and C give values that overlap with A. The range

of σ^{min} is the same, $\Delta\sigma = 0.5°$, in all three graphs, so the three samples can be compared with each other.

   Finally, we display the significances of the alignments of the three samples with B and C and compare those with the significances found by A. The calculations can be found in Part II the Appendix. Even though the distribution parameters η_0 do not always agree within the error bars, the significances, p-values, for all three methods A,B,C, do agree within the tolerances. One effect to notice in Fig. 15 is the asymmetry of the plus and minus uncertainties and that is most evident for the 99 Stars.

Out[ ]=
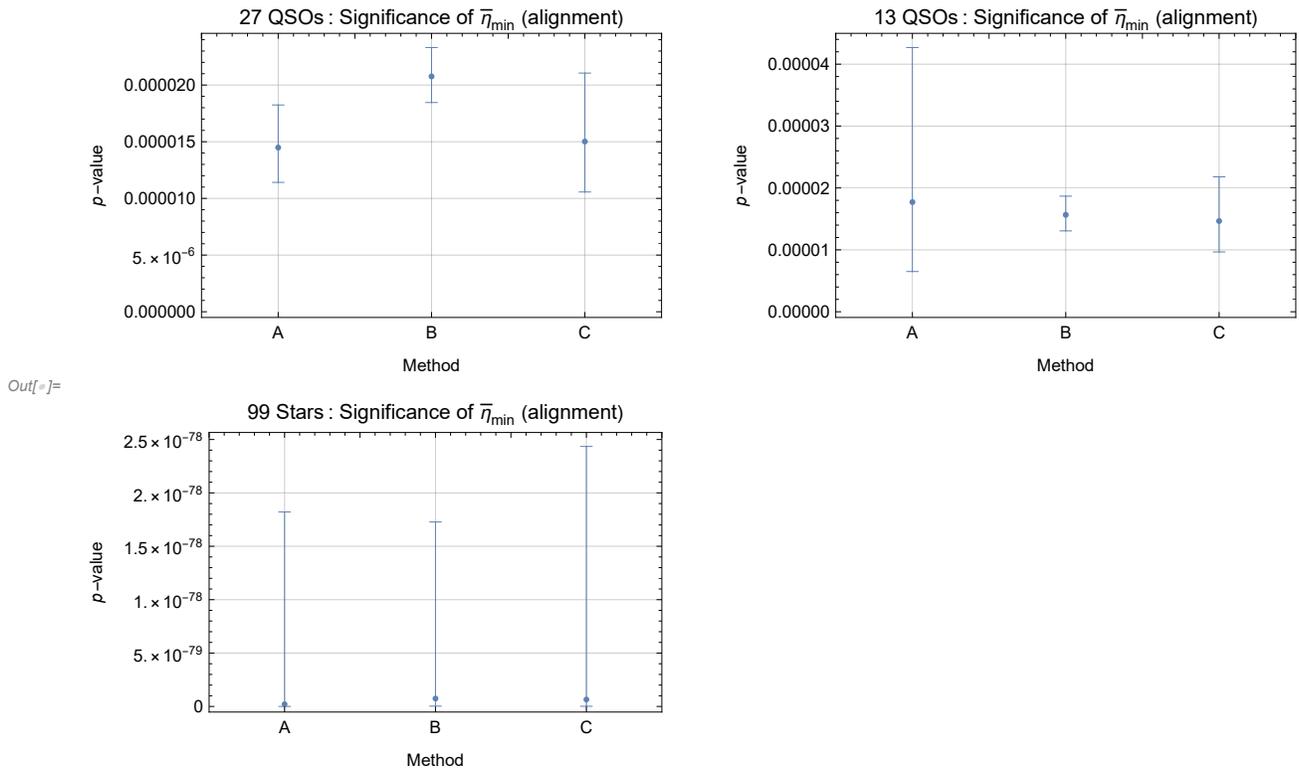
Figure 15: The significance of the smallest alignment angle $\overline{\eta}_{\texttt{min}}$ for three samples as determined by Methods A, B, and C. The error bars are due to the standard errors of the parameters η_0^{min} and σ^{min} determined by the Mathematica software in the process of fitting the histograms in Fig. 12. For all three samples, the significances determined by the Interpolation Method B and the Formula Method C agree with the Direct Method A, within the error bars drawn here.
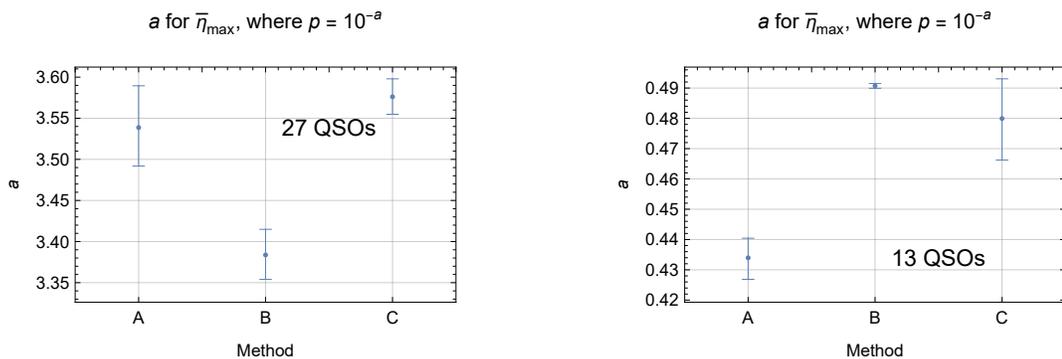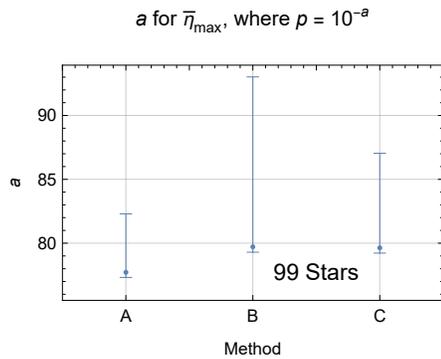
Figure 16: The significance exponent a for the smallest alignment angle $\overline{\eta}_{min}$ for three samples as determined by Methods A, B, and C. The error bars overlap for the sample of 99 Stars. However, for the first two samples, the error bars of B and C do not overlap with those of A. It should be expected that Interpolation of Library data and fitting Library data with functions would not produce results as good as those with Direct Method A. But, upon noting the fine scale of the vertical axes, it is clear that Methods B and C may be used to approximate the statistics quite well.

8. Concluding Remarks

The article presents two methods, Interpolation Method B and Formula Method C, to calculate the significance of the smallest alignment angle $\overline{\eta}_{min}$ in the Hub Test. Methods B and C, provide quick alternatives to the laborious process of generating and analyzing thousands of random runs required with Direct Method A.

The "Library" is an extensive collection of two parameters, the peak η_0 and the half-width σ of probability distributions. The random samples in the Library have numbers of sources N from 9 to 900 and region radii from 1/7 ° to 64°. Method B, interpolation of the Library data, should deliver estimates of the two parameters, η_0 and σ, that can be used to determine the probability distribution and significance of an observed result $\overline{\eta}_{min}$. With Method C, the fitting formulas to the two parameters, the peak η_0 and the half-width σ, provide excellent coverage of the 294 Library samples.

One application of Interpolation Method B and Formula Method C is mapping the significance of catalogued data, as done in Refs. 6,7. Catalogs of polarized sources or jet directions grow in size, so it may be helpful to have one or two quick ways to gauge the significance of alignment correlations for finite subsets of the data. A possible strategy is to use Methods B or C to locate regions of interest. Once a sample is found to be significantly aligned by Methods B or C, one is advised to apply Direct Method A to be confirm the result.

Whatever the project, the Library of probability distribution parameters may make the Hub Test easier to apply and check.

References

1. R. Shurtleff, the ready-to-run Mathematica version of this notebook is available at the following URL:
   https://www.dropbox.com/s/eyypcaql0b7dt5f/20220414InterpolateAndFormula5.nb?dl=0
   https://www.wolframcloud.com/obj/shurtleffr/Published/20220414InterpolateAndFormula5.nb
2. Wolfram Research, Inc., Mathematica, Version 13.0.1.0, Champaign, IL (2022).
3. Shurtleff, R., "Evaluating the Alignment of the Polarized Radio Waves from 27 QSOs in a Region near the NGP", https://vixra.org/abs/2105.0091 (2021).
4. Shurtleff, R., "Evaluating the Alignment of the Polarized Radio Waves from 13 QSOs in Ursa Major", https://vixra.org/abs/2110.0077 (2021).
5. Shurtleff, R., "Evaluating the Alignment of the Polarized Starlight from 99 Stars in a Region off the Disk of the Milky Way", https://vixra.org/abs/2104.0074 (2021).
6. Shurtleff, R., "Finding Regions of Correlated Polarization Directions of Polarized Starlight from 5830 Sources in the Milky Way Galaxy", https://vixra.org/abs/2202.0128 (2022).
7. Shurtleff, R., "Finding Regions of Correlated Polarization Directions of the 1450 QSO Radio Sources in the JVAS 1450 Catalog", https://vixra.org/abs/2201.0074 (2022).
8. Shurtleff, R., "Indirect polarization alignment with points on the sky, the Hub Test" , https://vixra.org/abs/2011.0026 (2020).
9. See Table 1 QSOs and Sec. 2 in Pelgrims, V. and Hutsemékers, D., Polarization alignments of quasars from the JVAS/CLASS 8.4-GHz surveys, MNRAS, 450, 4161-4173, doi: 10.1093/mnras/stv917, arXiv:astro-ph:1503.03482 (2015).   The JVAS1450 catalog was kindly emailed to me by V. Pelgrims .
10. Jackson, N., Battye, R. A., Browne, I. W. A.,   Joshi, S., Muxlow, T. W. B., and Wilkinson, P. N., A survey of polarization in the JVAS/CLASS flat-spectrum radio source surveys - I. The data and catalogue production, MNRAS, 376, 371-377, doi: 10.1111/j.1365-2966.2007.11442.x , arXiv:astro-ph/0703273 (2007).
11. Heiles, C., An agglomeration of stellar polarization catalogs, Astron. J. 119, 923 (2000).
12. Berdyugin A., Piirola V., Teerikorpi P., Astron. Astrophys. 561, A24 DOI: 10.1051/0004-6361/201322604 (2014).

Part II the Appendix

00. Preface

The problem of "significance" is to determine the likelihood that random polarizations directions would have better alignment or avoidance than the observed polarization directions. We talk about polarization directions, but one can analyze any set of transverse directions on the sky for their alignment.

Note that Part I the Article discusses the measure of alignment, the smallest alignment angle $\bar{\eta}_{min}$ almost exclusively. However, the Hub Test also determines correlations of the transverse directions by finding points on the Celestial Sphere that the great circles avoid. In that case, the relevant quantity is the largest avoidance angle $\bar{\eta}_{max}$ which is the maximum of the alignment function $\bar{\eta}(H)$. Both alignment and avoidance are treated in this part, Part II the Appendix.

A1. Introduction and Probability Distributions, Significance Formulas

Given a sample whose sources have measured transverse directions, say electromagnetic polarization or physical features like jets, then the alignment angle function $\bar{\eta}(H)$, Eq. (1) and Figs. 2,3, can be calculated at all point H on the sphere except at the sources themselves. The directions are best aligned with the points H_{min} and $-H_{min}$ where the smallest alignment angle occurs, $\bar{\eta}_{min}$, i.e. $\bar{\eta}_{min}$ = $\bar{\eta}(H_{min}) \leq \bar{\eta}(H)$. Thus $\bar{\eta}_{min}$ gauges alignment. The largest alignment angle $\bar{\eta}_{max}$ is found at hubs H_{max} and $-H_{max}$ and gauges a different correlation, avoidance.

The probability distributions for alignment and avoidance of samples with randomly oriented polarization directions are fit with pseudo-Gaussian formulas. We choose

$$P_{min}(\eta) = \left(\frac{norm}{\sigma (2\pi)^{1/2}} \right) \left(1 + e^{4 \frac{(\eta - \eta 0 - \sigma)}{\sigma}} \right)^{-1} e^{-\frac{1}{2} \left(\frac{\eta - \eta 0}{\sigma} \right)^2} , \qquad (A1)$$

and

$$P_{max}(\eta) = \left(\frac{norm}{\sigma (2\pi)^{1/2}} \right) \left(1 + e^{-4 \frac{(\eta - \eta 0 + \sigma)}{\sigma}} \right)^{-1} e^{-\frac{1}{2} \left(\frac{\eta - \eta 0}{\sigma} \right)^2} , \qquad (A2)$$

where norm is the normalization constant, norm $= 1.22029$, so that the definite integral over all η of the probability distributions is equal to one. The two distribution formulas Eq. (A1) and (A2) involve two parameters η_0 and σ, one set for alignment, η_0^{min} and σ^{min}, and one set for avoidance, η_0^{max} and σ^{max}. The "min" and "max" are dropped in Eq. (A1) and (A2) for convenience.

To find the significance, or p-value, of the smallest alignment angle $\bar{\eta}_{min}$, one finds the likelihood of smaller random run values by integrating $P_{min}(\eta)$ from below,

$$p(\bar{\eta}_{min}) = \int_{-\infty}^{\bar{\eta}_{min}} P_{min}(\eta) \, d\eta \qquad (A3)$$

For avoidance, the largest avoidance angle $\bar{\eta}_{max}$ is likely not as large as random runs with a significance of

$$p(\bar{\eta}_{max}) = \int_{\bar{\eta}_{max}}^{\infty} P_{max}(\eta) \, d\eta . \qquad (A4)$$

By Eqs. (A1 - A4) the calculation of significances for a sample of correlated behavior indicated by the smallest alignment angle $\bar{\eta}_{min}$ and the largest avoidance angle $\bar{\eta}_{max}$ rests on obtaining values of the four probability distribution parameters $\eta_0^{min}, \sigma^{min}, \eta_0^{max}$, and σ^{max}. In my humble opinion, it would be better to use "s" or "sig" for significance rather than the accepted notation "p".

The Direct Method A finds the four needed parameters $\eta_0^{min}, \sigma^{min}, \eta_0^{max}$, and σ^{max} by generating many, many random runs and analyzing the results. This notebook's usefullness is based on a "Library", a collection of the four parameters from a wide selection of samples that have been created and analyzed elsewhere in other notebooks. For any given sample, Interpolation Method B makes an interpolation of the Library data, while Formula Method C employs a set of four functions that fit the Library data. Either Method yields values of the four needed parameters $\eta_0^{min}, \sigma^{min}, \eta_0^{max}$, and σ^{max} for a given observed sample.

Sec. A2 describes the Library. Interpolation Method B is the topic of Sec. A3 and Sec. A4 has Formula Method C. Formula Method C has formulas which are uncertain. Sec. A5 treats the uncertainties in the formulas of Method C. The three previously studied samples offer the means to compare the three different Methods A, B, C in Sec. A6.

A2 The Library

A2a. Preliminary

Definitions:

homeDirectory   location of this notebook and many associated files. Other directories are used also.
er(α,δ)   radial unit vector in the direction from the origin to the point with longitude and latitude (α,δ)
mean   the arithmetic average of a set of numbers, $\frac{1}{N} \sum_{i=1}^{N} n_i$
stanDev   the standard deviation. Given a set of N numbers n_i with mean value m, the standard deviation is
$\left(\frac{1}{N} \sum_{i=1}^{N} (n_i - m)^2\right)^{1/2}$, the square root of the average of the squares of the differences of the numbers with the mean. Note that we
divide by N to get the average of the deviations squared.
norm   a constant used to normalize the distribution so the integral of probability is 1.
probMIN0, probMAX0   probability distributions for η, dependent on parameters η_0, σ. [alignment (MIN) and avoidance (MAX)]
signiMIN0($\bar{\eta}_{min}$, η_0^{min}, σ^{min})   significance of $\bar{\eta}_{min}$, dependent on probability distribution parameters η_0^{min}, σ^{min}. (alignment)
signiMAX0($\bar{\eta}_{max}$, η_0^{max}, σ^{max})   significance of $\bar{\eta}_{max}$, dependent on probability distribution parameters η_0^{max}, σ^{max}. (alignment)

In[1]:=
```
Print["The date and time that this statement was evaluated: ", Now]
Print["The computer time expended so far is ", TimeUsed[], " seconds."]
```

The date and time that this statement was evaluated: Mon 18 Apr 2022 08:36:46 GMT−4

The computer time expended so far is 1.328 seconds.

In[3]:=
```
homeDirectory =
    "C:\\Users\\shurt\\Dropbox\\HOME_DESKTOP-0MRE5OJ\\SendXXX_CJP_CEJPetc\\SendViXra\\
      20200715AlignmentMethod\\20210606RunRandomTryToUpdate\\20210719Series\\20220108
      NewSeriesRunData\\20220220NewSeriesAlternateMethodsABC";
```

In[4]:=
```
(*The radial unit vector in the direction from the origin to the
 point with longitude and latitude (α,δ), 0 ≤ α ≤ 2π, -π/2 ≤ δ ≤ π/2.*)
er[α_, δ_] := er[α, δ] = {Cos[α] Cos[δ], Sin[α] Cos[δ], Sin[δ]}
```

In[5]:=
```
mean[data_] := (1 / Length[data]) Sum[data[[i4]], {i4, Length[data]}];
(* arithmetic average *)
stanDev[data_] := ((1 / Length[data]) Sum[(data[[i5]] - mean[data])^2, {i5, Length[data]}])^(1/2)
(*standard deviation*)
```

In[7]:= `(* y = ((η - η0)/σ); dy = dη/σ *)`

`(* The normalization factor "norm" is needed to make`
`the integral of the probability distribution equal to unity. *)`

$$\text{norm} = \left(\frac{1}{(2\pi)^{1/2}} \, \text{NIntegrate}\left[\left(1 + e^{4(y-1)}\right)^{-1} e^{-\frac{y^2}{2}}, \{y, -\infty, \infty\} \right] \right)^{-1};$$

`Print["The normalization scale factor is norm = ",`
`  norm , " for both alignment` probMIN0 and avoidance probMAX0.`"];`

The normalization scale factor is norm = 1.22029 for both alignment probMIN0 and avoidance probMAX0.

In[9]:= (∗Eqs. (A1, A3∗)

$$\text{probMIN0}[\eta\_, \eta0\_, \sigma\_] := \left(\frac{\text{norm}}{\sigma (2\pi)^{1/2}} \right) \left(1 + e^{4 \frac{(\eta - \eta0 - \sigma)}{\sigma}} \right)^{-1} e^{-\frac{1}{2}\left(\frac{\eta - \eta0}{\sigma}\right)^2}$$

`signiMIN0[`$\eta\_, \eta0\_, \sigma\_$`] := NIntegrate[`probMIN0$[\eta1, \eta0, \sigma]$`, {`$\eta1$`, -∞, `η`}]`

$$\text{probMAX0}[\eta\_, \eta0\_, \sigma\_] := \left(\frac{\text{norm}}{\sigma (2\pi)^{1/2}} \right) \left(1 + e^{-4 \frac{(\eta - \eta0 + \sigma)}{\sigma}} \right)^{-1} e^{-\frac{1}{2}\left(\frac{\eta - \eta0}{\sigma}\right)^2}$$

`signiMAX0[`$\eta\_, \eta0\_, \sigma\_$`] := NIntegrate[`probMAX0$[\eta1, \eta0, \sigma]$`, {`$\eta1$`, `η`, ∞}]`

The significance signiMIN0$[\eta, \eta0, \sigma]$ is the Integral of probMIN0, i.e. signiMIN0 $= \int_{-\infty}^{\eta} P_{MIN}(\eta) \, d\eta$.

The significance signiMAX0$[\eta, \eta0, \sigma]$ is the Integral of probMAX0, i.e. signiMAX0 $= \int_{\eta}^{\infty} P_{MAX}(\eta) \, d\eta$.

In[13]:= `(* {{nSrc,`ρ`RgnRadius},{x0min,dx0min},{bmin,dbmin},`
`{amin,damin},{x0max,dx0max},{bmax,dbmax},{amax,damax},`
`{`$\sigma\theta$`rHminToCenter,`θ`rHminToCenter},{`$\sigma\theta$`rHmaxToCenter,`θ`rHmaxToCenter}} *)`

A2b. The Library data

Definitions:

fitData0   This table is output from another notebook. Not all of the data is relevant here.
powr[i,j], fitDataReduced1   The number of digits are adjusted so that the standard error has two significant figures,2 sig figs

"fitData0" table

1a. nSrc, number of sources 1b. rgnRadius, nominal radius of region   1c. RMS radius 1d. # Random Runs
2a. x0min: x0 = $\eta0$ align (min)   2b. x0min: err: dx0 - σ for x0 = $\eta0$ align (min)
3a. bmin: b = σ align (min) 3b. dbmin: err: db - σ for b = σ align (min)
4a. amin: a = Amplitude align (min) 4b. damin: err: da - σ for a = Amplitude align (min)
5a. x0max: x0 = $\eta0$ avoid (max) 5b. dx0maxx0max: err: dx0 - σ for x0 = $\eta0$ avoid (max)
6a. bmax: b = σ avoid (max)   6b. dbmax: err: db - σ for b = σ avoid (max)
7a. amax: a = Amplitude avoid (max)7b. damax: err: da - σ for a = Amplitude avoid (max)
8a. $\sigma\theta$rHminToCenter: stanDev[anglerHminToCenter] - σ for θ to H   8b. θrHminToCenter: mean[anglerHminToCenter]   - θ to H
9a. $\sigma\theta$rHmaxToCenter: stanDev[anglerHmaxToCenter] - σ for θ to H   9b. θrHmaxToCenter: mean[anglerHmaxToCenter] - θ to H

fitData   Parameters of the alignment (min) and avoidance (max) random run distributions. Originally in radians, converted to degrees below.

1a. nSrci[i]   Number of sources   1b. ρNomi[i] Nominal radius, deg. 1c. ρRMSi[i] RMS radius, deg.

2a. η0mini[i]   peak alignment distribution   2b. dη0mini[i] standard error

3a. σmini[i]   half-width alignment distr. 3b. dσmini[i] standard error

4a. η0maxi[i]   peak alignment distribution 4b. dη0maxi[i] standard error

5a. σmaxi[i]   half-width alignment distr.   5b. dσmaxi[i] standard error

wi[i]   inverse square root of the number of sources, $w = 1/N^{1/2}$

τRMSi[i]   inverse RMS radius, in deg.$^{-1}$

nSrcList   list of the number of sources in the Library samples

wList   list of the inverse square root of the number of sources for Library samples

ρNomList   list of the nominal radii, deg., of the Library samples

colornSrcList   assigned colors for the numbers of sources nSrc and the corresponding w

colorρList   assigned colors for the nominal radii ρNominal and the corresponding inverses τNominal

nSrcColorKey   Legend for the number of sources nSrc

wNColorKey   Legend for the number of sources nSrc and w $= $ nSrc$^{-1/2}$

ρNomColorKey   Legend for the nominal radii ρNominal

$\rho\tau$NomColorKey   Legend for the nominal radii ρNominal and their inverses, and

 τNominal

idsFORnSrc[iN]   ID #s in fitData, one set for each value of nSrc

idsFORρNom[iρ]   ID #s in fitData, one set for each value of ρNominal

lpRhoη0min[iN]   list plot of Library data, alignment peak parameter $\eta_0^{\min}$ versus RMS radius ρRMS

eta0MinVSRhoData   Combined plot of the lpRhoη0min for all Library data

lpρSigmamin, sigmaMINvsRadiusData Library data plots, except alignment half-width parameter $\sigma^{\min}$ replaces $\eta_0^{\min}$

lp$\rho\eta$0max, eta0VSradiusDataMax   Library data plots, except avoidance peak $\eta_0^{\max}$ replaces $\eta_0^{\min}$

lpρSigmamax, sigmaMAXvsRadiusData   Plots of Library data

fitData:

The following hidden cells upload the fitData file and check the fitData. The last hidden cell contains the fitData. You can view the data by clicking "Open" in "Cell Properties" in the Mathematica menu. On my computer, I see "Cell" four items to the right of File. Click "Cell" first.

In[16]:= **ρNomList = {1 / 7., 1 / 6., 1 / 5., 1 / 4., 1 / 3., 1 / 2., 1., 3 / 2.,**
     2., 3., 4., 6., 8., 10., 15., 16., 24., 32., 42., 52., 64.};(*degrees*)

expectedρNom = ρNomList $\left(\dfrac{2.\,\pi}{360.}\right)$;(*radians*)

Print["The ", Length[ρNomList], " nominal radii are ", ρNomList, ", in degrees."]

The 21 nominal radii are {0.142857, 0.166667, 0.2, 0.25, 0.333333, 0.5,
  1., 1.5, 2., 3., 4., 6., 8., 10., 15., 16., 24., 32., 42., 52., 64.}, in degrees.

fitData:

The fitData table is accessible from the live Mathematica notebook in Ref. 1. It is contained in the last hidden cell.

You can see the fitData table if you "Open" the cell by highlighting it, clicking "Cell" which is four items to the right of "File" above, then "Cell Properties", then "Open".

In[27]:= `(*Identify the items in the fitData table with functions having recognizable names.*)`

`nSrci[i_] := fitData[[i, 1, 1]]`

`(*Convert fitData radians to DEGREES:*)`

ρ`Nomi[i_] := fitData[[i, 1, 2]]` $\left(\dfrac{360.}{2.\ \pi}\right)$ `(*Use the nominal radius in degrees*)`

ρ`RMSi[i_] := fitData[[i, 1, 3]]` $\left(\dfrac{360.}{2.\ \pi}\right)$ `(*Use the RMS radius in degrees*)`

η`0mini[i_] := fitData[[i, 2, 1]]` $\left(\dfrac{360.}{2.\ \pi}\right)$

`d`η`0mini[i_] := fitData[[i, 2, 2]]` $\left(\dfrac{360.}{2.\ \pi}\right)$

σ`mini[i_] := fitData[[i, 3, 1]]` $\left(\dfrac{360.}{2.\ \pi}\right)$

`d`σ`mini[i_] := fitData[[i, 3, 2]]` $\left(\dfrac{360.}{2.\ \pi}\right)$

η`0maxi[i_] := fitData[[i, 5, 1]]` $\left(\dfrac{360.}{2.\ \pi}\right)$

`d`η`0maxi[i_] := fitData[[i, 5, 2]]` $\left(\dfrac{360.}{2.\ \pi}\right)$

σ`maxi[i_] := fitData[[i, 6, 1]]` $\left(\dfrac{360.}{2.\ \pi}\right)$

`d`σ`maxi[i_] := fitData[[i, 6, 2]]` $\left(\dfrac{360.}{2.\ \pi}\right)$

In[38]:= `(*Quantities w and rRMS, calculated from fitData.*)`

`wi[i_] := ` $\dfrac{1}{\text{nSrci[i]}^{1/2}}$ `(*w = 1`$\big/$`N`$^{1/2}$`*)`

`rRMSi[i_] := ` ρ`RMSi[i] / 90.`

In[40]:= `nSrcList = Union[Table[nSrci[i], {i, Length[fitData]}]];`

ρ`NomList = Union[Table[`ρ`Nomi[i], {i, Length[fitData]}]];`

`wList = Table`$\left[\dfrac{1}{\text{nSrcList[[i]]}^{1/2.}}\right.$`, {i, Length[nSrcList]}`$\Big]$`;`

The number of sources for samples in the Library are nSrc =
{9, 16, 25, 36, 49, 64, 81, 100, 121, 225, 256, 324, 625, 900}.

Their inverse square roots are w = {0.333333, 0.25, 0.2, 0.166667, 0.142857,
0.125, 0.111111, 0.1, 0.0909091, 0.0666667, 0.0625, 0.0555556, 0.04, 0.0333333} .

The list of nominal radii is ρ = {0.142857, 0.166667, 0.2, 0.25, 0.333333, 0.5,
1., 1.5, 2., 3., 4., 6., 8., 10., 15., 16., 24., 32., 42., 52., 64.} , in degrees.

There are 21 nominal radii.

```
In[430]:= (*A color for each number N*)
     tryTheseColorsnSrc = {75, 85, 15, 45, 1, 7, 65, 114, 12, 14, 45, 25, 59, 39};
     Table[ColorData["Atoms", "ColorList"]⟦i⟧, {i, tryTheseColorsnSrc }];
```

```
In[50]:= (*Table[{i,ColorData["Atoms","ColorList"]⟦i⟧},
      {i,Length[ColorData["Atoms","ColorList"]]  }]*)
```

```
In[432]:= (*A color for each nominal radius *)tryTheseColorsρNom =
        {88, 102, 62, 5, 86, 8, 28, 42, 82, 92, 15, 4, 63, 40, 61, 108, 45, 80, 90, 20, 76};
     Table[ColorData["Atoms", "ColorList"]⟦i⟧, {i, tryTheseColorsρNom }];
```

```
In[54]:= (*Set up the color codes.*)
     colornSrcList = Table[ColorData["Atoms", "ColorList"]⟦i⟧, {i, tryTheseColorsnSrc }];
     colorρList = Table[ColorData["Atoms", "ColorList"]⟦i⟧, {i, tryTheseColorsρNom }];
     nSrcColorKey = Table[{nSrcList⟦i⟧, colornSrcList⟦i⟧}, {i, Length[nSrcList]}];
     wNColorKey =
       {{nSrcList⟦1⟧, colornSrcList⟦1⟧, "1/3"}, {nSrcList⟦2⟧, colornSrcList⟦2⟧, "1/4"},
        {nSrcList⟦3⟧, colornSrcList⟦3⟧, "1/5"}, {nSrcList⟦4⟧, colornSrcList⟦4⟧, "1/6"},
        {nSrcList⟦5⟧, colornSrcList⟦5⟧, "1/7"}, {nSrcList⟦6⟧, colornSrcList⟦6⟧, "1/8"},
        {nSrcList⟦7⟧, colornSrcList⟦7⟧, "1/9"}, {nSrcList⟦8⟧, colornSrcList⟦8⟧, "1/10"},
        {nSrcList⟦9⟧, colornSrcList⟦9⟧, "1/11"} , {nSrcList⟦10⟧, colornSrcList⟦10⟧, "1/15"},
        {nSrcList⟦11⟧, colornSrcList⟦11⟧, "1/16"}, {nSrcList⟦12⟧, colornSrcList⟦12⟧, "1/18"},
        {nSrcList⟦13⟧, colornSrcList⟦13⟧, "1/25"}, {nSrcList⟦14⟧, colornSrcList⟦14⟧, "1/30"} };
     ρNomColorKey = Table[{ρNomi[i], colorρList⟦i⟧}, {i, Length[ρNomList]}];
```

Display code	ρ (Nominal)
	1/7 °
	1/6 °
	1/5 °
	1/4 °
	1/3 °
	1/2 °
	1 °
	3/2 °
	2 °
	3 °
	4 °
	6 °
	8 °
	10 °
	15 °
	16 °
	24 °
	32 °
	42 °
	52 °
	64 °

Display code	N
	9
	16
	25
	36
	49
	64
	81
	100
	121
	225
	256
	324
	625
	900

Table A1: Color codes for N (*left*) and
ρNominal (*Right*), the properties used to label the Library data.

```
In[61]:= Table[idsFORnSrc[iN] = {}, {iN, Length[nSrcList]}];
(* ids give fitData table entries with nSrc sources.*)
Table[For[i = 1, i ≤ Length[fitData], i++, If[(nSrci[i] - nSrcList[[iN]])^2 < 1,
    AppendTo[idsFORnSrc[iN], i]]], {iN, Length[nSrcList]}];
iN = 3;
Print["For example, idsFORnSrc[", iN,
 "] gives the ids for the records in the fitData table for iN = ", iN,
 " and nSrc = ", nSrcList[[iN]], " They are ids = ", idsFORnSrc[iN], "."]
Clear[iN]
```

For example, idsFORnSrc[3
 ] gives the ids for the records in the fitData table for iN = 3 and nSrc = 25 They are ids =
 {43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63}.

In[66]:= `Table[idsFORρNom[iρ] = {}, {iρ, Length[ρNomList]}];`

`Table[For[i = 1, i ≤ Length[fitData], i++, If[Abs[`$\frac{(ρNomi[i] - ρNomList[\![iρ]\!])}{Min[ρNomi[i], ρNomList[\![iρ]\!]]}$`] < 10`$^{-2}$`,`

`AppendTo[idsFORρNom[iρ], i]]], {iρ, Length[ρNomList]}];`

`iρ = 3;`
`Print["For example, idsFORρNom[", iρ,`
`"] gives the ids for the records in the fitData table for iρ = ",`
`iρ, " with nominal radius ρNom = ", ρNomList[\![iρ]\!],`
`" degrees.  idsFORρNom[", iρ, "] = ", idsFORρNom[iρ], "."]`
`Clear[iρ]`

For example, idsFORρNom[3] gives the ids for the records in the fitData table for iρ =
  3 with nominal radius ρNom = 0.2 degrees. idsFORρNom[3] =
  {3, 24, 45, 66, 87, 108, 129, 150, 171, 192, 213, 234, 255, 276}.

In[73]:= `lpRhoη0min[iN_] := ListPlot[Table[{ρRMSi[i], η0mini[i]}, {i, idsFORnSrc[iN]}],`
`PlotRange → {{0., 80.}, {25., 45.}}, FrameLabel → {"ρRMS, degrees ", "η0min, degrees"},`
`PlotLabel → "η0min vs ρRMS", PlotStyle → colornSrcList[\![iN]\!] ,`
`GridLines → Automatic, Frame → True, ImageSize → 72 × 6]`
`eta0MinVSRhoData = Show[Table[lpRhoη0min[iN], {iN, Length[wList]}]]`
`nSrcColorKey`
`Print["Figure A1: Library data. The parameter `η_0^{min}` in Eq. A1`
`  plotted versus the RMS radius ρRMS, sorted by number of sources N. "]`

Out[74]=

Out[75]= `{{9, ▨}, {16, ▨}, {25, ▨}, {36, ▨}, {49, ▨}, {64, ▨}, {81, ▨},`
`{100, ▨}, {121, ▨}, {225, ▨}, {256, ▨}, {324, ▨}, {625, ▨}, {900, ▨}}`

Figure A1: Library data. The parameter η_0^{min} in Eq. A1
  plotted versus the RMS radius ρRMS, sorted by number of sources *N*.

In[77]:= `lpρSigmamin[iN_] := ListPlot[Table[{ρRMSi[i],`

`    If[8.5 < Mod[i, Length[ρNomList]] < 9.5, Around[σmini[i], dσmini[i]], σmini[i]]},`

`    {i, idsFORnSrc[iN]}], PlotRange → {{0.0, 80}, {0.0, 8.}}, FrameLabel → {"ρ", "σ"},`

`    FrameTicks → { {{0, 0 °}, {2, 2 °}, {4, 4 °}, {6, 6 °}, {8, 8 °}}, None} ,`

`    {{{0, 0 °}, {10, 10 °}, {20, 20 °}, {30, 30 °}, {40, 40 °}, {50, 50 °},`

`      {60, 60 °}, {70, 70 °}, {80, 80 °}}, None} }, PlotLabel → "σ`^{min}` vs ρRMS",`

`    PlotStyle → colornSrcList〚iN〛 , GridLines → Automatic, Frame → True,`

`    IntervalMarkersStyle → colornSrcList〚iN〛, ImageSize → 72 × 6];`

`  sigmaMINvsRadiusData = Show[Table[lpρSigmamin[iN], {iN, Length[wList]}]]`

`  Print["Figure A2: Library data. The parameter σ`^{min}` in Eq. A1`

`    plotted versus the RMS radius ρRMS, sorted by number of sources N. "]`

Out[78]=
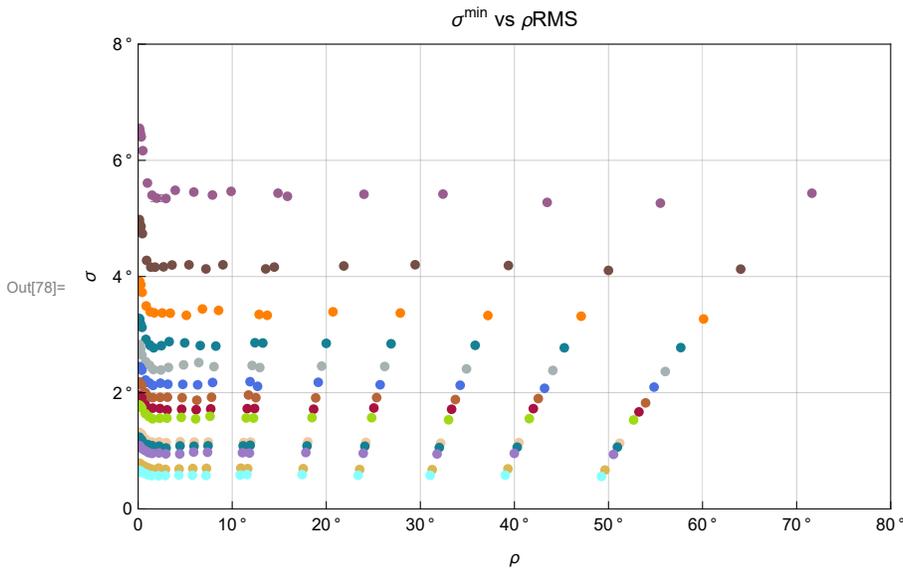
Figure A2: Library data. The parameter σ^{min} in Eq. A1
  plotted versus the RMS radius ρRMS, sorted by number of sources *N*.

In[434]:= `lpRhoη0max[iN_] := ListPlot[Table[{ρRMSi[i], η0maxi[i]}, {i, idsFORnSrc[iN]}],`

`    PlotRange → {{0., 80.}, {45., 65.}}, FrameLabel → {"ρRMS, degrees ", "η0max, degrees"},`

`    PlotLabel → "η0max vs ρRMS", PlotStyle → colornSrcList〚iN〛 ,`

`    GridLines → Automatic, Frame → True, ImageSize → 72 × 6]`

In[435]:= **eta0maxVSRhoData = Show[Table[lpRhoη0max[iN], {iN, Length[wList]}]]**

Out[435]=

Out[436]= {{9, ■}, {16, ■}, {25, ■}, {36, ■}, {49, ■}, {64, ■}, {81, ■},
{100, ■}, {121, ■}, {225, ■}, {256, ■}, {324, ■}, {625, ■}, {900, ■}}
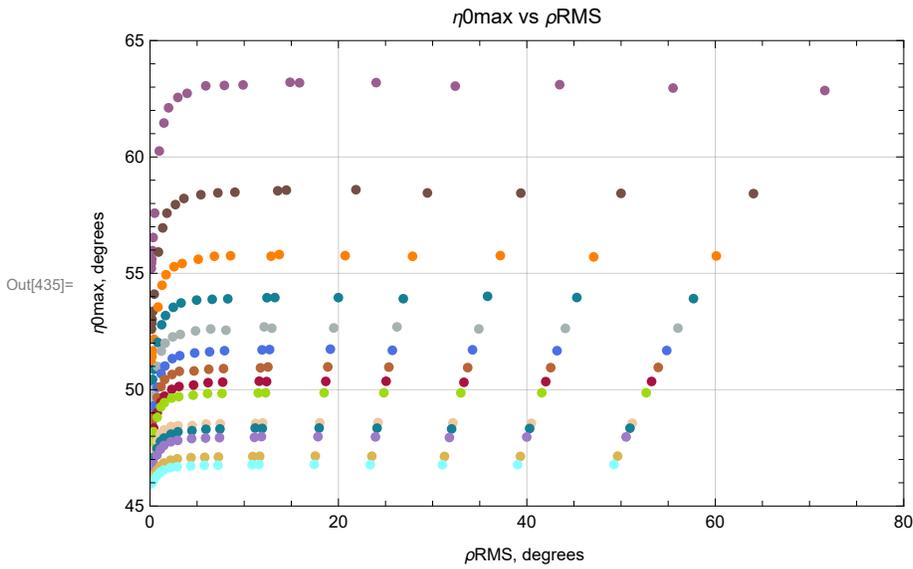
Figure A3: Library data. The parameter η_0^{max} in Eq. A2
plotted versus the RMS radius ρRMS, sorted by number of sources *N*.

In[438]:= **lpRhoσmax[iN_] := ListPlot[Table[{ρRMSi[i], σmaxi[i]}, {i, idsFORnSrc[iN]}],**
PlotRange → {{0., 80.}, {0., 7.}}, FrameLabel → {"ρRMS, degrees ", "σmax, degrees"},
PlotLabel → "σmax vs ρRMS", PlotStyle → colornSrcList〚iN〛 ,
GridLines → Automatic, Frame → True, ImageSize → 72 × 6]

In[439]:= **SigmaMaxVSRhoData = Show[Table[lpRhoσmax[iN], {iN, Length[wList]}]]**

Out[439]=

Out[440]= {{9, ■}, {16, ■}, {25, ■}, {36, ■}, {49, ■}, {64, ■}, {81, ■},
   {100, ■}, {121, ■}, {225, ■}, {256, ■}, {324, ■}, {625, ■}, {900, ■}}
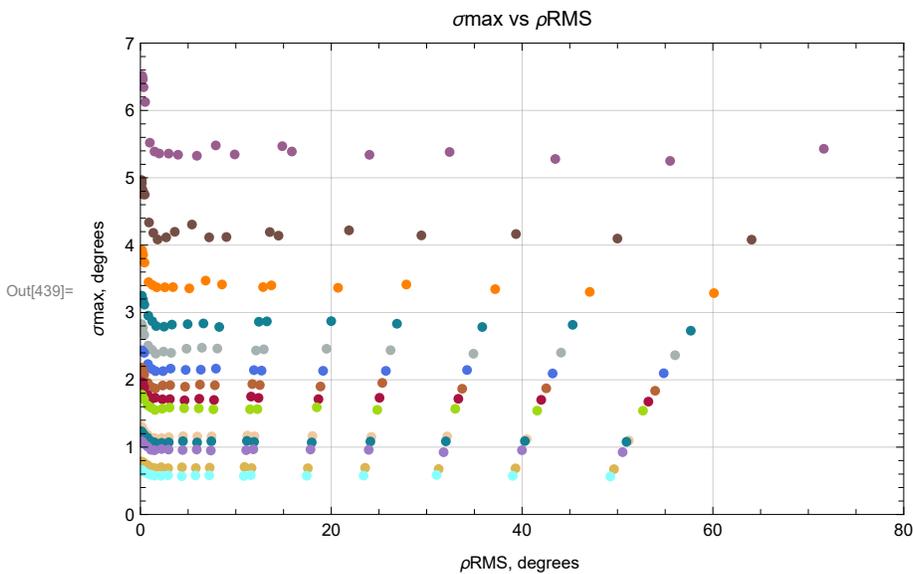
Figure A4: Library data. The parameter σ^{max} in Eq. A2
  plotted versus the RMS radius ρRMS, sorted by number of sources *N*.

A3. Interpolation Method B

The Library constructed in Sec. A2 is a table of the values of the four parameters, η_0^{min}, σ^{min}, η_0^{max}, and σ^{max}, needed to determine the probability distributions and significances in Eqs. (A1-A4).

We consider the four parameters as functions of the number of sources *N* and the root-mean-square radius ρRMS. It is sometimes convenient to denote the inverse square root of *N* by *w* and the ratio of ρRMS to 90° by rRMS,

$$w \;=\; N^{-1/2} \quad \text{and} \;\; \text{rRMS} \;=\; \rho\text{RMS}/90° . \tag{A5}$$

The Library data for the smallest alignment angle $\overline{\eta}_{min}$ is graphed in Figs. 9 and A1. The data for avoidance, *i.e.* the largest avoidance angle $\overline{\eta}_{max}$, is plotted in Fig. A4 in terms of *w* and ρRMS.

Definitions:

Tables: w$\rho\eta$0minLib, wρdη0minLib, w$\rho\eta$0maxLib, wρdη0maxLib, w$\rho\sigma$minLib, wρdσminLib, w$\rho\sigma$maxLib, wρdσmaxLib
The tables w$\rho\eta$0minLib ... put Library data in the form (*w*, ρRMS, *quantity*), where "quantity" is a parameter or its standard error: η_0^{min}, dη_0^{min}, σ^{min}, dσ^{min}, η_0^{max}, dη_0^{max}, σ^{max} , dσ^{max}

The associated interpolation functions are η0minBint, dη0minBint , η0maxBint , dη0maxBint, σminBint, dσminBint, σmaxBint , dσmaxBint

nExample, ρRMSexample   Values of the number of sources and the RMS radius are used to illustrate the interpolation of the Library data with Interpolation Method B

Setting up the interpolations takes two steps. First a tables of the data are constructed. Each entry of a table has the form {*w*, ρRMS, parameter}. Second, the interpolation for each parameter is defined. There are four parameters η_0^{min}, σ^{min}, η_0^{max}, and σ^{max} and each one has a standard error dη_0^{min}, dσ^{min}, dη_0^{max}, and dσ^{max} developed in the fitting process. Thus there are 8 tables of data and 8 interpolation functions.

```
In[88]:= wρη0minLib = Table[{{wi[i], ρRMSi[i]}, η0mini[i]}, {i, Length[fitData]}];
        wρdη0minLib = Table[{wi[i], ρRMSi[i], dη0mini[i]}, {i, Length[fitData]}];
        wρη0maxLib = Table[{wi[i], ρRMSi[i], η0maxi[i]}, {i, Length[fitData]}];
        wρdη0maxLib = Table[{wi[i], ρRMSi[i], dη0maxi[i]}, {i, Length[fitData]}];
        wρσminLib = Table[{wi[i], ρRMSi[i], σmini[i]}, {i, Length[fitData]}];
        wρdσminLib = Table[{wi[i], ρRMSi[i], dσmini[i]}, {i, Length[fitData]}];
        wρσmaxLib = Table[{wi[i], ρRMSi[i], σmaxi[i]}, {i, Length[fitData]}];
        wρdσmaxLib = Table[{wi[i], ρRMSi[i], dσmaxi[i]}, {i, Length[fitData]}];
```

In[96]:= `η0minBint = Interpolation[wρη0minLib];(* int - interpolation function*)`
`dη0minBint = Interpolation[wρdη0minLib];`
`η0maxBint = Interpolation[wρη0maxLib];`
`dη0maxBint = Interpolation[wρdη0maxLib];`
`σminBint = Interpolation[wρσminLib];`
`dσminBint = Interpolation[wρdσminLib];`
`σmaxBint = Interpolation[wρσmaxLib];`
`dσmaxBint = Interpolation[wρdσmaxLib];`

⋯ Interpolation: Interpolation on unstructured grids is currently only supported for InterpolationOrder–>1 or InterpolationOrder–>All. Order will be reduced to 1.

⋯ Interpolation: Interpolation on unstructured grids is currently only supported for InterpolationOrder–>1 or InterpolationOrder–>All. Order will be reduced to 1.

⋯ Interpolation: Interpolation on unstructured grids is currently only supported for InterpolationOrder–>1 or InterpolationOrder–>All. Order will be reduced to 1.

⋯ Interpolation: Interpolation on unstructured grids is currently only supported for InterpolationOrder–>1 or InterpolationOrder–>All. Order will be reduced to 1.

⋯ Interpolation: Interpolation on unstructured grids is currently only supported for InterpolationOrder–>1 or InterpolationOrder–>All. Order will be reduced to 1.

⋯ Interpolation: Interpolation on unstructured grids is currently only supported for InterpolationOrder–>1 or InterpolationOrder–>All. Order will be reduced to 1.

⋯ Interpolation: Interpolation on unstructured grids is currently only supported for InterpolationOrder–>1 or InterpolationOrder–>All. Order will be reduced to 1.

⋯ Interpolation: Interpolation on unstructured grids is currently only supported for InterpolationOrder–>1 or InterpolationOrder–>All. Order will be reduced to 1.

⋯ Interpolation: Interpolation on unstructured grids is currently only supported for InterpolationOrder–>1 or InterpolationOrder–>All. Order will be reduced to 1.

By the rules of interpolations, when the variables w and ρ are in the range of the Library data, then Mathematica finds an average value from the surrounding Library data points. The range of the variable w is $\frac{1}{15} \le w \le \frac{1}{3}$, while the range of ρRMS depends on the number of sources, decreasing as the source density increases.

In[104]:= `(*An example of using the interpolating functions.*)`
`nExample = 155(*RandomInteger[{9,900}]*);`
`$\rho$RMSexample = 1.42077(*RandomReal[{0.15,64}]*);`
`Print["The most likely value of the two random`
`    run distribution parameters $\eta_0^{min}$ and $\sigma^{min}$ (alignment) for ",`
` nExample, " sources in a region with RMS radius ", $\rho$RMSexample,`
` "° is found by interpolation of Library data:   "]`
`Print["Given N = ", nExample, " and $\rho$RMS = ", $\rho$RMSexample,`
` "°, we find that w = N$^{-1/2}$ = ", N[nExample$^{-1/2}$], ", and $\rho$ = $\rho$RMS = ", $\rho$RMSexample,`
` ". The interpolations give $\eta_0^{min}$ = ", $\eta$0minBint[nExample$^{-1/2}$, $\rho$RMSexample], "°, $\sigma^{min}$ = ",`
` $\sigma$minBint[nExample$^{-1/2}$, $\rho$RMSexample], "°, $\eta_0^{max}$ = ", $\eta$0maxBint[nExample$^{-1/2}$, $\rho$RMSexample],`
` "°, $\sigma^{max}$ = ", $\sigma$maxBint[nExample$^{-1/2}$, $\rho$RMSexample], "°."]`
`Print["Compare these values with the Library data plotted in Figs. A1 to A4."]`

The most likely value of the two random
    run distribution parameters η_0^{min} and σ^{min} (alignment) for 155
  sources in a region with RMS radius 1.42077° is found by interpolation of Library data:

Given N = 155 and ρRMS = 1.42077°, we find that w = N$^{-1/2}$ = 0.0803219, and ρ = ρRMS = 1.42077
 . The interpolations give η_0^{min} = 41.295°, σ^{min} = 1.35777°, η_0^{max} = 48.78°, σ^{max} = 1.36664°.

Compare these values with the Library data plotted in Figs. A1 to A4.

A4. Fitting the Library data with formulas

Definitions:

tables of Library, i.e. fitData, info (w, rRMS, *quantity*)
wrEtamin for η_0^{min}, wrEtamax for η_0^{max}, wrSigmamin for σ^{min}, wrSigmamax for σ^{max} , the four distribution parameters.

 (η0minAtBIGr[iN], η0maxAtBIGr[iN])   the values of η_0^{min}, η_0^{max} for the largest radius sample and for each *N*.

(wrEtaminFlatTor1, wrEtamaxFlatTor1)   Flat to r = 1 = $\rho/90$. We extend wrEtamin, wrEtamax by assuming the value of η_0^{min},
η_0^{max} for a sample with radius 90° is the same as the value at the largest radius sample, for each *N*.

NonlinearModelFits to the data tables wrEtamin, ... defined above
nlmrEtamin for η_0^{min}, nlmrEtamax   for η_0^{max}, nlmrSigmamindegA9 for σ^{min}, nlmrSigmamaxdegA9 for σ^{max}

values of the parameters of the fits
 { etamink0,etamina0, etaminb0,etamina2, etaminb3, etaminb4,etaminc4,etaminbb45} for η_0^{min}
 { etamaxk0,etamaxa0, etamaxb0,etamaxa2, etamaxb3, etamaxb4,etamaxc4,etamaxbb45} for η_0^{max}
 { sigmamina0, sigmaminb0,sigmamina2, sigmaminb2, sigmamina4, sigmaminb4} for σ^{min}
 { sigmamaxa0, sigmamaxb0,sigmamaxa2, sigmamaxb2, sigmamaxa4, sigmamaxb4} for σ^{max}

standard errors of the parameters of the fits
 { detamink0,detamina0, detaminb0,detamina2, detaminb3, detaminb4,detaminc4,detaminbb45} for η_0^{min}
 { detamaxk0,detamaxa0, detamaxb0,detamaxa2, detamaxb3, detamaxb4,detamaxc4,detamaxbb45}   for η_0^{max}

{ dsigmamina0, dsigmaminb0,dsigmamina2, dsigmaminb2,dsigmamina4, dsigmaminb4} for σ^{min}

{ dsigmamaxa0, dsigmamaxb0,dsigmamaxa2, dsigmamaxb2,dsigmamaxa4, dsigmamaxb4} for σ^{max}

formulas of the fits in terms of w, ρRMS

etaminRhoFit for η_0^{min}, etamaxRhoFit for η_0^{max}, sigmaminRhoFit for σ^{min}, sigmamaxRhoFit for σ^{max} , the four distribution parameters.

plots of the fits in terms of w, ρRMS

plotRhoEtamin[iN] (for one value of N), for all N: eta0MinVSRhoFit for η_0^{min},

plotRhoEtamax[iN] (for one value of N), for all N: eta0MaxVSRhoFit for η_0^{max},

plotRhoSigmamin[iN] (for one value of N), for all N: sigma0MinVSRhoFit for σ^{min},

plotRhoSigmamax[iN] (for one value of N), for all N: sigma0MaxVSRhoFit   for σ^{max}

Percent Differences of the Library data to the fits

sortPercentDiffEtaminfit   for η_0^{min}, sortPercentDiffEtamaxfit for η_0^{max}, sortPercentDiffSigmaminfit for σ^{min}, sortPercentDif-fSigmamaxfit for σ^{max}

A4a. Fit η_0^{min}

In[108]:= ```
(*To avoid spurious behavior for samples beyond the size of Library data,
we assign the Library data at the largest ρRMS value to the point at ρRMS = 90°. *)
wrEtamin = Sort[Table[{wi[i], rRMSi[i], η0mini[i]}, {i, Length[fitData]}]];
η0minAtBIGr[iN_] := Sort[Table[{rRMSi[i], η0mini[i]} , {i, idsFORnSrc[iN]}]][[-1, 2]]
wrEtaminFlatTor1 =
    Join[wrEtamin, Table[{wList[[iN]], 1., η0minAtBIGr[iN]}, {iN, Length[wList]}]];
{wrEtamin[[1]], wrEtamin[[-1]]}
```

Out[111]= $\left\{\left\{\dfrac{1}{30}, 0.00114368, 44.0362\right\}, \left\{\dfrac{1}{3}, 0.795843, 27.0844\right\}\right\}$


In[112]:= ```
nlmrEtamin = NonlinearModelFit[wrEtaminFlatTor1,
      (k0 - x (a0 + b0 x - (a2 + b3 y ) Tanh[ (b4 y + c4 y² + bb45 x y )])),
      {{k0, 45}, {a0, 28.2}, {b0, 4.61}, {a2, -22.2}, {b3, -25.6},
       {b4, 71.8}, {c4, -590.}, {bb45, 47.2}}, {x, y}, MaxIterations → 1000];
{ etamink0, etamina0, etaminb0, etamina2, etaminb3, etaminb4, etaminc4, etaminbb45} =
   {k0, a0, b0, a2, b3, b4, c4, bb45} /. nlmrEtamin["BestFitParameters"];
{ detamink0, detamina0, detaminb0, detamina2, detaminb3,
   detaminb4, detaminc4, detaminbb45} = nlmrEtamin["ParameterErrors"];
```

Fitting function: $f(x=w, y=\rho RMS/90.)$ =
$$k0 - x\left(a0 + b0\,x - (a2 + b3\,y)\,\text{Tanh}\left[b4\,y + bb45\,x\,y + c4\,y^2\right]\right) \qquad (A5)$$

Fit : $\eta_0^{\min}(w, \rho RMS)$ = $44.9585 - w\,(25.8207 + 4.70015\,w -$
$$(-26.3517 - 0.00451303\,\rho RMS)\,\text{Tanh}\left[0.60349\,\rho RMS + 0.408717\,w\,\rho RMS - 0.0066049\,\rho RMS^2\right]) \qquad (A6)$$

Estimated variance = 0.00476533

⋯ General: Exp[−1590.29] is too small to represent as a normalized machine number; precision may be lost.

⋯ General: Exp[−872.607] is too small to represent as a normalized machine number; precision may be lost.

	Estimate	Standard Error	t-Statistic	P-Value
k0	44.9585	0.0130775	3437.84	0.
a0	25.8207	0.201108	128.393	3.14999×10^{-264}
b0	4.70015	0.556505	8.44583	1.31592×10^{-15}
Parameter Table: a2	−26.3517	0.0839719	−313.815	0.
b3	−0.406173	0.12215	−3.32519	0.000993314
b4	54.3141	1.12601	48.2361	2.21467×10^{-143}
c4	−53.4997	1.31496	−40.6855	3.84433×10^{-124}
bb45	36.7845	4.69193	7.83995	7.93001×10^{-14}

In[119]:= **etaminRhoFit[w_, ρRMS_] := Normal[nlmrEtamin] /. {x → w, y → ρRMS / 90.}**

In[120]:= **plotRhoEtamin[iN_] :=**
   Plot[etaminRhoFit[wList〚iN〛, ρ], {ρ, 0., 80.}, PlotPoints → 200, PlotRange → All,
   PlotStyle → colornSrcList〚iN〛];

In[423]:= **iN = 1;**
Show[{lpRhoη0min[iN], plotRhoEtamin[iN]}]
Clear[iN]
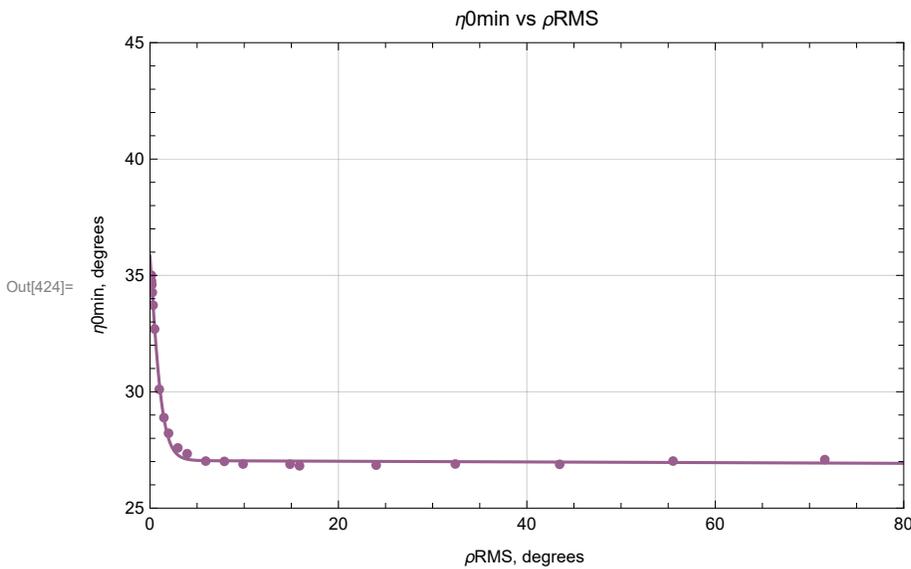
Out[424]=

Out[428]= {{9, ■}, {16, ■}, {25, ■}, {36, ■}, {49, ■}, {64, ■}, {81, ■},
   {100, ■}, {121, ■}, {225, ■}, {256, ■}, {324, ■}, {625, ■}, {900, ■}}

Figure A5: Fit to Library data. For an active choice of
   N, the parameter $\eta_0^{\min}$ in Eq. A1 plotted versus the RMS radius ρRMS.

In[447]:= **eta0MinVSRhoFit =**
Show[Table[Show[{lpRhoη0min[iN], plotRhoEtamin[iN]}], {iN, Length[nSrcList]}],
ImageSize → 72 × 5]

Out[447]=

η0min vs ρRMS

Out[449]= {{9, ⬛}, {16, ⬛}, {25, ⬛}, {36, ⬛}, {49, ⬛}, {64, ⬛}, {81, ⬛},
{100, ⬛}, {121, ⬛}, {225, ⬛}, {256, ⬛}, {324, ⬛}, {625, ⬛}, {900, ⬛}}
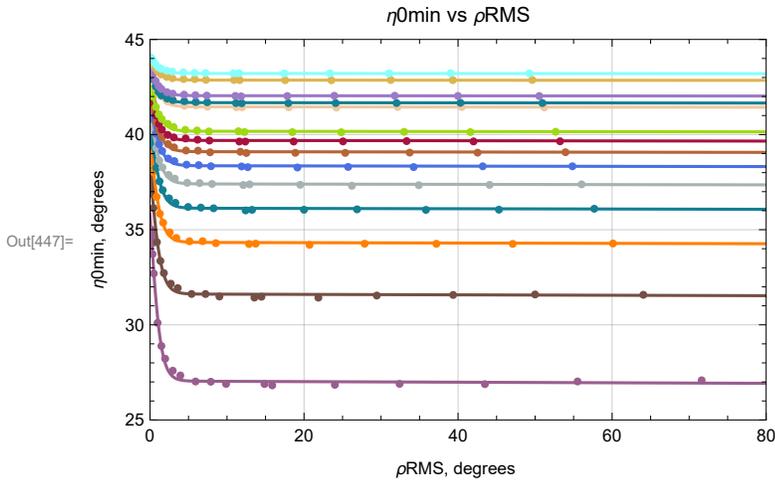
Figure A6: Fit to Library data for the parameter η_0^{min} in Eq. A1, sorted by number of sources N.

In[443]:= **sortPercentDiffEtaminfit =**
Sort[Flatten[Table[100. Abs[(η0mini[i] − etaminRhoFit[wi[i], ρRMSi[i]]) / η0mini[i]],
{i, Length[fitData]}]]] ;
ListPlot[sortPercentDiffEtaminfit , PlotRange → All, FrameLabel → {"", " % diff η_0^{min} "},
PlotLabel → "% diff. of the η_0^{min} fit ", GridLines → Automatic, Frame → True]

% diff. of the η_0^{min} fit

Out[444]=
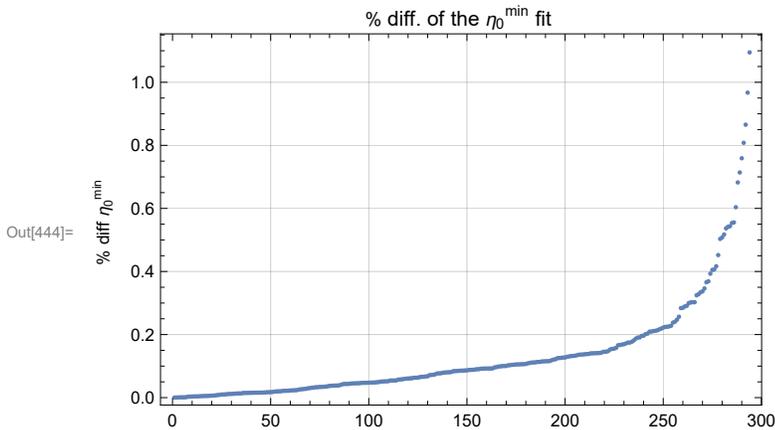
Figure A7: The percent difference of the
fit to Library data for η_0^{min}. The mean %diff of the fit to data is
0.130918%, with a standard dev. of 0.162754% . The largest %diff = 1.09474%.

A4b. Fit $\eta_0{}^{max}$

```
In[134]:= (*Artificially extend the Library data to ρRMS = 90°. *)
      wrEtamax = Sort[Table[{wi[i], rRMSi[i], η0maxi[i]}, {i, Length[fitData]}]];
      η0maxAtBIGr[iN_] := Sort[Table[{rRMSi[i], η0maxi[i]}, {i, idsFORnSrc[iN]}]]〚-1, 2〛
      wrEtamaxFlatTor1 =
        Join[wrEtamax, Table[{wList〚iN〛, 1., η0maxAtBIGr[iN]}, {iN, Length[wList]}]];
      {wrEtamax〚1〛, wrEtamax〚-1〛}
```

$$Out[137]= \left\{\left\{\frac{1}{30}, 0.00114368, 45.9625\right\}, \left\{\frac{1}{3}, 0.795843, 62.8535\right\}\right\}$$

```
In[138]:= nlmrEtamax = NonlinearModelFit[wrEtamaxFlatTor1,
          (k0 + x (a0 + b0 x - (a2 + b3 y) Tanh[ (b4 y + c4 y² + bb45 x y )])),
          {{k0, 45}, {a0, 28.2}, {b0, 4.61}, {a2, -22.2}, {b3, -25.6},
            {b4, 71.8}, {c4, -590.}, {bb45, 47.2}}, {x, y}, MaxIterations → 1000];
      { etamaxk0, etamaxa0, etamaxb0, etamaxa2, etamaxb3, etamaxb4, etamaxc4, etamaxbb45} =
        {k0, a0, b0, a2, b3, b4, c4, bb45} /. nlmrEtamax["BestFitParameters"];
      { detamaxk0, detamaxa0, detamaxb0, detamaxa2, detamaxb3,
          detamaxb4, detamaxc4, detamaxbb45} = nlmrEtamax["ParameterErrors"];
```

Fitting function: $f(x=w, y=\rho RMS/90.) =$
$k0 + x \left(a0 + b0\, x - (a2 + b3\, y)\, \text{Tanh}\left[b4\, y + bb45\, x\, y + c4\, y^2\right]\right)$ (A7)

Fit : $\eta_0{}^{max}(w, \rho RMS) = 44.9986 + w\, (26.9406 + 3.08965\, w -$
$(-25.9516 - 0.00244035\, \rho RMS)\, \text{Tanh}\left[0.718833\, \rho RMS + 0.25621\, w\, \rho RMS - 0.00771145\, \rho RMS^2\right])$ (A8)

Estimated variance = 0.00494574

⋯ General: Exp[−1584.82] is too small to represent as a normalized machine number; precision may be lost.

⋯ General: Exp[−854.093] is too small to represent as a normalized machine number; precision may be lost.

	Estimate	Standard Error	t-Statistic	P-Value
k0	44.9986	0.0133304	3375.64	0.
a0	26.9406	0.207206	130.018	7.72215×10^{-266}
b0	3.08965	0.568653	5.43328	1.1459×10^{-7}
a2	−25.9516	0.087979	−294.975	0.
b3	−0.219631	0.12378	−1.77437	0.0770161
b4	64.6949	1.35264	47.8288	2.10577×10^{-142}
c4	−62.4627	5.36713	−11.638	4.38453×10^{-26}
bb45	23.0589	5.52874	4.17073	0.0000397903

Parameter Table:

```
In[145]:= etamaxRhoFit[w_, ρRMS_] := Normal[nlmrEtamax] /. {x → w, y → ρRMS / 90.}
```

```
In[146]:= plotRhoEtamax[iN_] :=
        Plot[etamaxRhoFit[wList〚iN〛, ρ], {ρ, 0., 80.}, PlotPoints → 200, PlotRange → All,
          PlotStyle → colornSrcList〚iN〛 ];
```

In[451]:= **eta0MaxVSRhoFit =**
  Show[Table[Show[{lpRho𝜂0max[iN], plotRhoEtamax[iN]}], {iN, Length[nSrcList]}],
   ImageSize → 72 × 5]
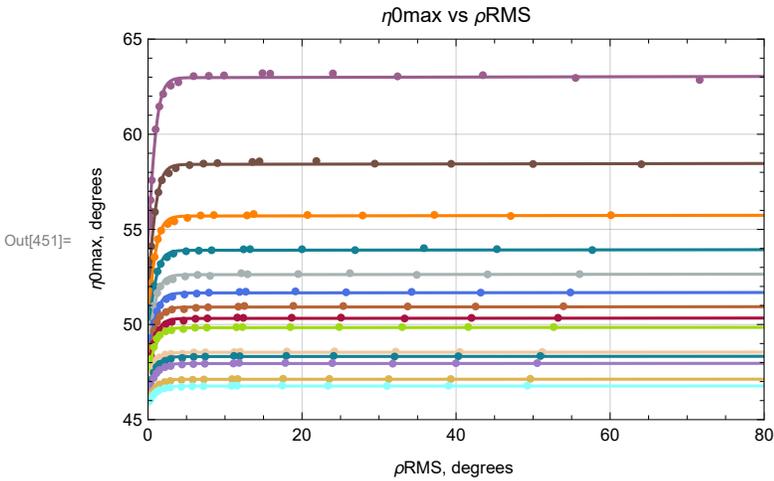
Out[451]=

Out[636]= {{9, ⬛}, {16, ⬛}, {25, ⬛}, {36, ⬛}, {49, ⬛}, {64, ⬛}, {81, ⬛},
  {100, ⬛}, {121, ⬛}, {225, ⬛}, {256, ⬛}, {324, ⬛}, {625, ⬛}, {900, ⬛}}

Figure A8: Fit to Library data for the parameter η_0^{max} in Eq. A2, sorted by number of sources N.

In[457]:= **sortPercentDiffEtamaxfit =**
   Sort[Flatten[Table[100. Abs[(𝜂0maxi[i] − etamaxRhoFit[wi[i], 𝜌RMSi[i]]) / 𝜂0maxi[i]],
    {i, Length[fitData]}]]] ;
  ListPlot[sortPercentDiffEtamaxfit , PlotRange → All, FrameLabel → {"", " % diff η_0^{max} "},
   PlotLabel → "% diff of the η_0^{max} fit ", GridLines → Automatic, Frame → True]

Out[458]=
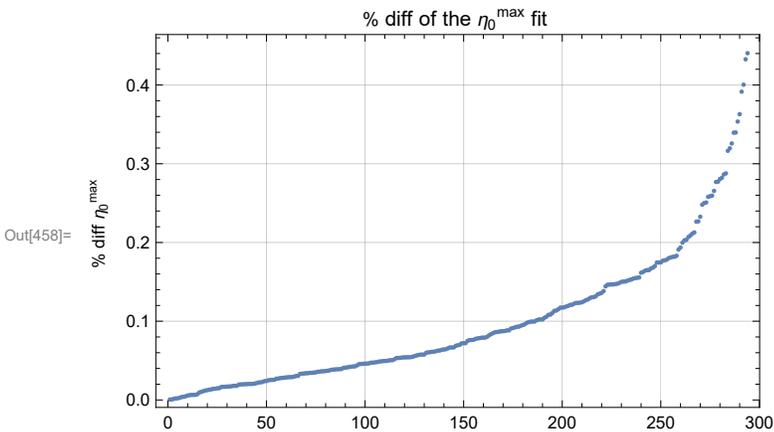

Figure A9: The percent difference of fit
   to Library data for η_0^{max}. The mean %diff of the fit to data is
 0.0972924%, with a standard dev. of 0.0858119% . The largest %diff = 0.44043%.

A4c. Fit $\sigma^{\min}$

In[153]:= **wrSigmamin = Table[{wi[i], rRMSi[i], σmini[i]}, {i, Length[fitData]}];**

In[154]:= **nlmrSigmamindegA9 = NonlinearModelFit$\Big[$wrSigmamin,**

$\qquad$ **+ $\Big(\dfrac{x}{4.}$ (a0 + b0 x − (a2 + b2 x) Tanh[(a4 + b4 y)])$\Big)$, {{a0, 74.5}, {b0, −5.84}, {a2, 4.08},**

$\qquad$ **{b2, 10.76}, {a4, −1.338}, {b4, 208.}}, {x, y}, MaxIterations → 1000$\Big]$;**

{ sigmamina0, sigmaminb0, sigmamina2, sigmaminb2, sigmamina4, sigmaminb4} =
$\quad$ **{a0, b0, a2, b2, a4, b4} /. nlmrSigmamindegA9["BestFitParameters"];**

{ dsigmamina0, dsigmaminb0, dsigmamina2, dsigmaminb2, dsigmamina4, dsigmaminb4} =
$\quad$ **nlmrSigmamindegA9["ParameterErrors"];**

Fitting function: $f(x=w, y=\rho\mathrm{RMS}/90.) = 0.25\,x\,(a0 + b0\,x − (a2 + b2\,x)\,\mathrm{Tanh}[a4 + b4\,y])$ $\qquad$ (A9)

Fit : $\sigma^{\min}(w, \rho\mathrm{RMS}) = 0.25\,w$
$\quad$ $(74.5043 − 5.83586\,w + (4.07678 + 10.7564\,w)\,\mathrm{Tanh}[1.33841 − 2.31346\,\rho\mathrm{RMS}])$ $\qquad$ (A10)

Estimated variance = 0.000841484

⚬⚬⚬ General: Exp[−917.306] is too small to represent as a normalized machine number; precision may be lost.

		Estimate	Standard Error	t-Statistic	P-Value
	a0	74.5043	0.183742	405.483	0.
	b0	−5.83586	0.73164	−7.97641	3.58036×10^{-14}
Parameter Table:	a2	4.07678	0.200891	20.2935	1.81979×10^{-57}
	b2	10.7564	0.793409	13.5571	1.01429×10^{-32}
	a4	−1.33841	0.136367	−9.81476	8.37837×10^{-20}
	b4	208.211	15.1007	13.7881	1.47936×10^{-33}

In[161]:= **Clear[sigmaminRhoFit]**
sigmaminRhoFit[w_, ρRMS_] := Normal[nlmrSigmamindegA9] /. {x → w, y → ρRMS / 90.}

In[163]:= **plotRhoSigmamin[iN_] := Plot[sigmaminRhoFit[wList⟦iN⟧, ρ], {ρ, 0., 90.}, PlotPoints → 500,**
$\qquad$ **PlotStyle → colornSrcList⟦iN⟧, PlotRange → All];**

In[639]:=
```
iN = 2;
lpρSigmamin[iN];
plotRhoSigmamin[iN];
Show[{lpρSigmamin[iN], plotRhoSigmamin[iN]}]
nSrcColorKey
Print["Figure A10: Fit to Library data. For an active choice of N, here N = ",
  nSrcList[[iN]], ", the parameter σ^min in Eq. A1 plotted versus the RMS radius ρRMS. "]
Clear[iN]
```
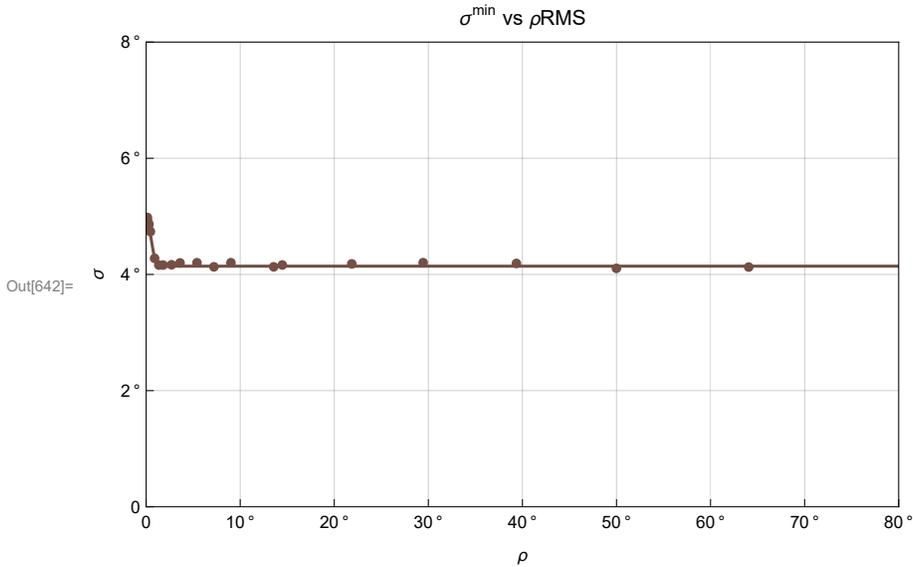
Out[642]=

Out[643]= {{9, ▦}, {16, ▦}, {25, ▦}, {36, ▦}, {49, ▦}, {64, ▦}, {81, ▦},
    {100, ▦}, {121, ▦}, {225, ▦}, {256, ▦}, {324, ▦}, {625, ▦}, {900, ▦}}

Figure A10: Fit to Library data. For an active choice of *N*, here *N* =
16, the parameter σ^{min} in Eq. A1 plotted versus the RMS radius ρRMS.

In[460]:=
```
sigma0MinVSRhoFit =
  Show[Table[Show[{lpρSigmamin[iN], plotRhoSigmamin[iN]}], {iN, Length[nSrcList]}],
    ImageSize → 72 × 5]
```
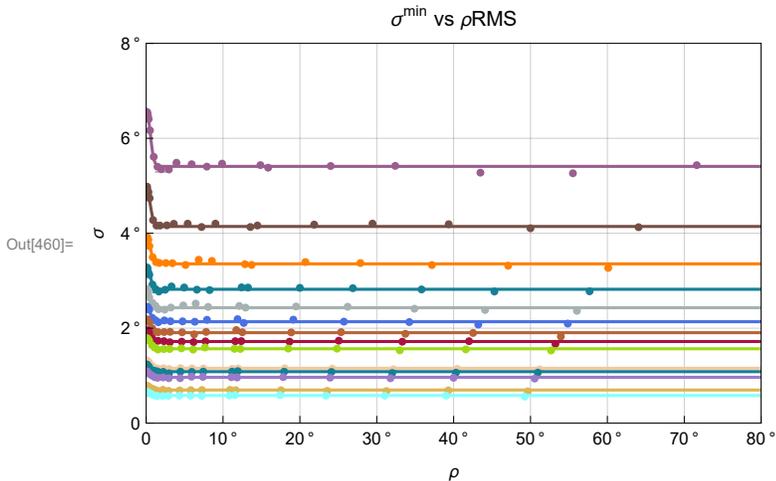
Out[460]=

Out[682]= {{9, ■}, {16, ■}, {25, ■}, {36, ■}, {49, ■}, {64, ■}, {81, ■},
 {100, ■}, {121, ■}, {225, ■}, {256, ■}, {324, ■}, {625, ■}, {900, ■}}

Figure A11: Fit to Library data of σ^{min} curves. Sorted by number of
sources N, the parameter σ^{min} in Eq. A1 plotted versus the RMS radius ρRMS.

In[463]:= **sortPercentDiffSigmaminfit =**
 Sort[Flatten[Table[100. Abs[(σmini[i] - sigmaminRhoFit[wi[i], ρRMSi[i]]) / σmini[i]],
 {i, Length[fitData]}]]] ;
ListPlot[sortPercentDiffSigmaminfit , PlotRange → All, FrameLabel → {"", " % diff σ^{min} "},
 PlotLabel → "% diff. of the σ^{min} fit", GridLines → Automatic, Frame → True]

Out[464]=
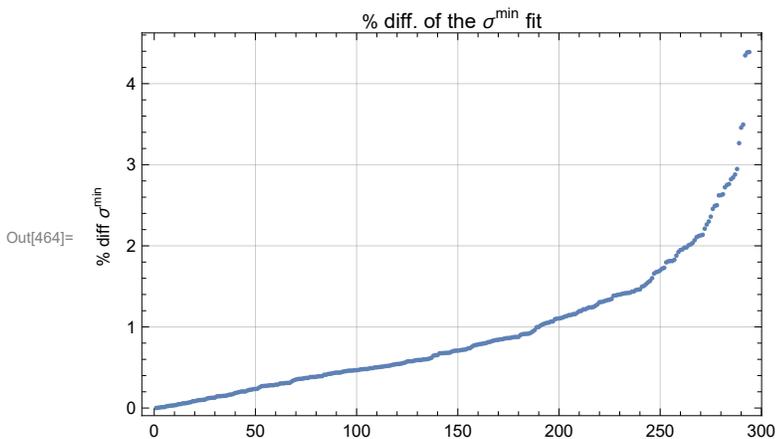

Figure A12: The percent difference of the
fit to Library data for σ^{min}. The mean % diff of the fit to data is
0.928308%, with a standard dev. of 0.807157% . The largest % diff = 4.38974%.

A4d. Fit σ^{max}

In[177]:= **wrSigmamax = Table[{wi[i], rRMSi[i], σmaxi[i]}, {i, Length[fitData]}];**

In[178]:= **nlmrSigmamaxdegA9 = NonlinearModelFit[wrSigmamax,**

 $+ \left(\dfrac{x}{4.} \ (a0 + b0 \ x - (a2 + b2 \ x) \ Tanh[\ (a4 + b4 \ y \ )]) \right)$, {{a0, 74.5}, {b0, -5.84}, {a2, 4.08},

 {b2, 10.76}, {a4, -1.338}, {b4, 208.}}, {x, y}, MaxIterations → 1000];
{ sigmamaxa0, sigmamaxb0, sigmamaxa2, sigmamaxb2, sigmamaxa4, sigmamaxb4} =
 {a0, b0, a2, b2, a4, b4} /. nlmrSigmamaxdegA9["BestFitParameters"];
{ dsigmamaxa0, dsigmamaxb0, dsigmamaxa2, dsigmamaxb2, dsigmamaxa4, dsigmamaxb4} =
 nlmrSigmamaxdegA9["ParameterErrors"];

Fitting function: $f(x=w, y=\rho\text{RMS}/90.) = 0.25\, x\, (a0 + b0\, x - (a2 + b2\, x)\, \text{Tanh}[a4 + b4\, y])$ (A11)

Fit : $\sigma^{max}(w, \rho\text{RMS}) = 0.25\, w$
$(74.4814 - 5.83469\, w + (3.86191 + 12.2034\, w)\, \text{Tanh}[1.25969 - 2.27575\, \rho\text{RMS}])$ (A12)

Estimated variance = 0.00100172

> ⋯ General: Exp[−884.035] is too small to represent as a normalized machine number; precision may be lost.

Parameter Table:

	Estimate	Standard Error	t-Statistic	P-Value
a0	74.4814	0.206226	361.163	0.
b0	−5.83469	0.866266	−6.73545	8.866×10^{-11}
a2	3.86191	0.226334	17.0629	1.35261×10^{-45}
b2	12.2034	0.9357	13.042	7.23988×10^{-31}
a4	−1.25969	0.147816	−8.52201	8.92593×10^{-16}
b4	204.817	16.519	12.3989	1.40601×10^{-28}

In[185]:= `Clear[sigmamaxRhoFit]`

`sigmamaxRhoFit[w_, ρRMS_] := Normal[nlmrSigmamaxdegA9] /. {x → w, y → ρRMS / 90.}`

In[187]:= `plotRhoSigmamax[iN_] := Plot[sigmamaxRhoFit[wList〚iN〛, ρ], {ρ, 0., 80.}, PlotPoints → 500,`
`    PlotStyle → colornSrcList〚iN〛, PlotRange → All ];`

In[469]:= `sigma0MaxVSRhoFit =`
`  Show[Table[Show[{lpRhoσmax[iN], plotRhoSigmamax[iN]}], {iN, Length[nSrcList]}],`
`   ImageSize → 72 × 5]`
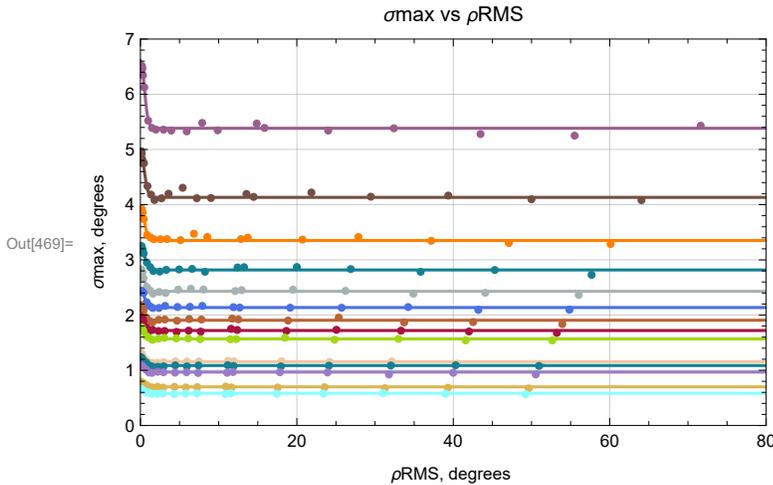
Out[469]=

Out[649]= $\{\{9, \blacksquare\}, \{16, \blacksquare\}, \{25, \blacksquare\}, \{36, \blacksquare\}, \{49, \blacksquare\}, \{64, \blacksquare\}, \{81, \blacksquare\},$
$\{100, \blacksquare\}, \{121, \blacksquare\}, \{225, \blacksquare\}, \{256, \blacksquare\}, \{324, \blacksquare\}, \{625, \blacksquare\}, \{900, \blacksquare\}\}$

Figure A13: Fit to Library data of σ^{max} curves. Sorted by number of
   sources N, the parameter σ^{max} in Eq. A2 plotted versus the RMS radius ρRMS.

In[472]:= **sortPercentDiffSigmamaxfit =**
    Sort[Flatten[Table[100. Abs[(σmaxi[i] – sigmamaxRhoFit[wi[i], ρRMSi[i]]) / σmaxi[i]],
        {i, Length[fitData]}]]] ;
  ListPlot[sortPercentDiffSigmamaxfit , PlotRange → All, FrameLabel → {"", " % diff σ^{max} "},
   PlotLabel → "% diff. of the σ^{max} fit", GridLines → Automatic, Frame → True]

Out[473]=
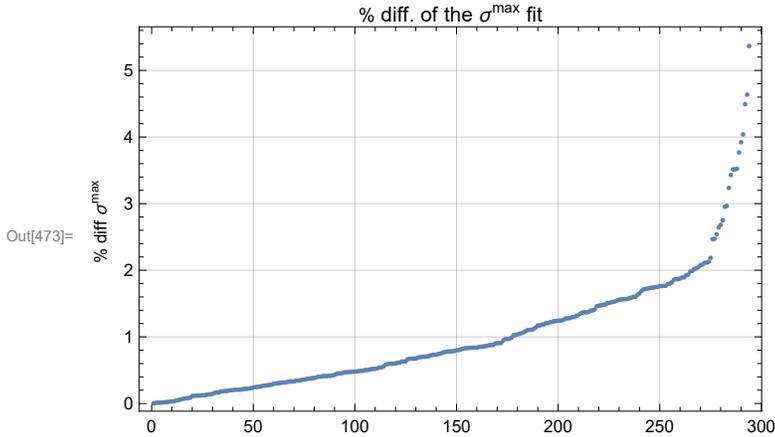

Figure A14: The percent difference of the fit
   to the Library data for σ^{max}. The mean % diff of the fit to data is
1.0031%, with a standard dev. of 0.877799% . The largest % diff = 5.36604%.

A5. Uncertainty of the fitting formulas

    In this section, we provide estimates of the uncertainties of the fitting functions for the parameters η_0^{min}, σ^{min}, η_0^{max}, σ^{max} found in the previous section. The uncertainties reflect the standard errors for the parameters like k0, a0, b0, a1, a2, b2, a3, ..., determined by Mathematica when it determined the fitting functions.

    Let us use the following fitting function for the η_0^{min} data to illustrate the process.

    f(x = w, y = ρRMS/90) = k0 - x(a0 + b0 x - (a2 + b3 y) Tanh[b4 y + c4 y^2 + bb45 x y]) .

The parameters, k0,a0,b0,a2,b3,b4,c4,bb45, were varied to make f fit the η_0^{min} data. One gets best values of the parameters and their standard errors with the Mathematica operation "NonlinearModelFit". Denote the standard errors by dk0, da0, db0, da1, da2, db2, da3. Each standard error drives its associated parameter to make the distribution decrease to some standard fraction of its peak value.

    We want a one-sigma fit. To get curves above and below the best curve that enclose 68% of the data points, we adjusted the standard errors by a common factor. See Fig. 16 for range of fits for η_0^{min}.

    The results are three functions, big, best and small. The range from the small to the big value curve contains 68% of the data points.

Definitions:

w1; ρ1 ; ρ2 values to use in examples

% differences between Library data and fits
percentDiffEtaminfitnSrc for η_0^{min}
percentDiffEtamaxfitnSrc for η_0^{max}
percentDiffsigmaminfitnSrc for σ^{min}

percentDiffsigmamaxfitnSrc for σ^{max}

derivative of fits with respect to the fit's parameters
detamindp for $\eta_0{}^{min}$
detamaxdp for $\eta_0{}^{max}$
dsigmamindp for σ^{min}
dsigmamaxdp for σ^{max}

sign* the sign of *

One-sigma fits. Between the larger formula, e.g. etaminFitbig, and the smaller, e.g. etaminFitsmall, there are 68% of the data points included.
etaminrFitbig and etaminFitsmall   for $\eta_0{}^{min}$
etamaxFitbig and etamaxFitsmall   for $\eta_0{}^{max}$
sigmaminFitbig and sigmaminFitsmall for σ^{min}
sigmamaxFitbig and sigmamaxFitsmall for σ^{max}

inbounds   A data point is "inbounds" if it is between the big and small values allowed by the fit.

A5a. Uncertainty of the $\eta_0{}^{min}$ formulas

In[194]:= **w1 = 16$^{-1/2}$; ρ1 = 24(*degrees*); ρ2 = 3(*degrees*);**

In[195]:= **percentDiffEtaminfitnSrc[iN_] := Table[**
   Abs[100. (η0mini[i] − etaminRhoFit[wi[i], ρRMSi[i]]) / η0mini[i]], {i, idsFORnSrc[iN]}]

In[485]:= **ListPlot[**
   {Table[{nSrcList〚iN〛, Max[percentDiffEtaminfitnSrc[iN]]}, {iN, Length[nSrcList]}],
    Table[{nSrcList〚iN〛, mean[percentDiffEtaminfitnSrc[iN]]}, {iN, Length[nSrcList]}]},
   Ticks → {Table[m^2, {m, {5, 10, 15, 18, 25, 30}}], Automatic},
   PlotRange → {{0., 1000.}, All}, PlotLabel → "$\eta_0{}^{min}$: Compare fit to data",
   AxesLabel → {"N sources", "%"}, PlotLegends → {"Max % diff.", "mean % diff."}]

Out[485]=
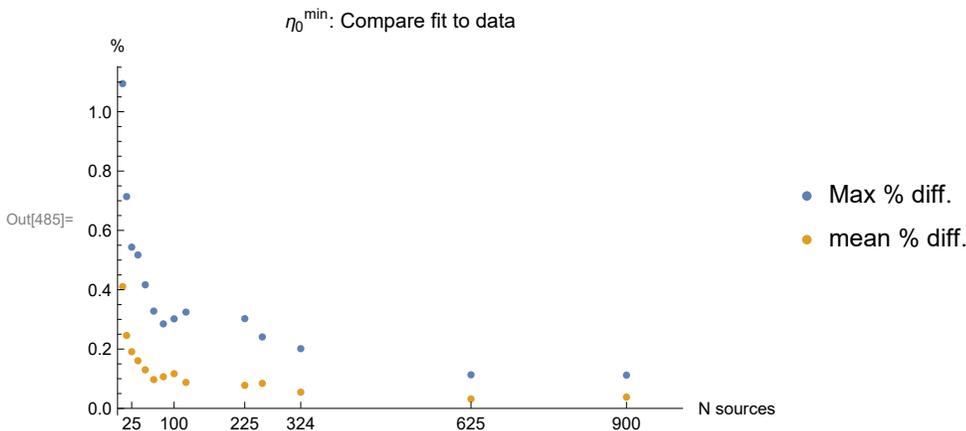
Figure A15: Percent difference stats of the fit with the Library data as a function of the number of sources *N*. For each *N*, this graph shows the mean % diff and the largest %diff of the fitting function for $\eta_0^{\min}$, over the set of all sample sizes ρRMS.

In[198]:= **(*Determine whether to add or subtract the standard errors of each parameter*)**
(*by finding the sign of the derivative wrt each parameter.*)
detamindp[x_, y_, p_] :=
 D[(k0 - x (a0 + (*5.4*) b0 x - (a2 + b3 y) Tanh[(b4 y + c4 y² + bb45 x y)])), p]
signdetamindp[x_, y_, p_] :=
 Simplify[Sign[detamindp[x, y, p] /. nlmrEtamin["BestFitParameters"]], {x > 0, y > 0}]
detamindp[x, y, b4];
signdetamindp[x, y, b4];
Print["Table of derivatives and their sign, {parameter p, df/dp, sign(df/dp)}:"]
Table[{p, detamindp[x, y, p], signdetamindp[x, y, p]}, {p, {k0, a0, b0, a2, b3, b4, c4, bb45}}]

Table of derivatives and their sign, {parameter p, df/dp, sign(df/dp)}:

Out[203]= $\{\{k0, 1, 1\}, \{a0, -x, -1\}, \{b0, -x^2, -1\},$
 $\{a2, x \, \text{Tanh}[b4 \, y + bb45 \, x \, y + c4 \, y^2], \text{Sign}[\text{Tanh}[(54.3141 + 36.7845 \, x - 53.4997 \, y) \, y]]\},$
 $\{b3, x \, y \, \text{Tanh}[b4 \, y + bb45 \, x \, y + c4 \, y^2], \text{Sign}[\text{Tanh}[(54.3141 + 36.7845 \, x - 53.4997 \, y) \, y]]\},$
 $\{b4, x \, y \, (a2 + b3 \, y) \, \text{Sech}[b4 \, y + bb45 \, x \, y + c4 \, y^2]^2, -1\},$
 $\{c4, x \, y^2 \, (a2 + b3 \, y) \, \text{Sech}[b4 \, y + bb45 \, x \, y + c4 \, y^2]^2, -1\},$
 $\{bb45, x^2 \, y \, (a2 + b3 \, y) \, \text{Sech}[b4 \, y + bb45 \, x \, y + c4 \, y^2]^2, -1\}\}$

In[204]:= **(*Adjust the range of the big and small curves to contain 68% of the data.*)**
adjust = 0.8
(*Clear[etaminFitbig,etaminFitsmall]*)

Out[204]= 0.8

In[205]:= **etaminFitbig[w_, ρ_] := etaminFitbig[w, ρ] =**
 ((k0 - x (a0 + (*5.4*) b0 x - (a2 + b3 y) Tanh[(b4 y + c4 y² + bb45 x y)])) /. {
  k0 → etamink0 + adjust signdetamindp[x, y, k0] detamink0,
  a0 → etamina0 + adjust signdetamindp[x, y, a0] detamina0,
  b0 → etaminb0 + adjust signdetamindp[x, y, b0] detaminb0,
  a2 → etamina2 + adjust signdetamindp[x, y, a2] detamina2,
  b3 → etaminb3 + adjust signdetamindp[x, y, b3] detaminb3,
  b4 → etaminb4 + adjust signdetamindp[x, y, b4] detaminb4,
  c4 → etaminc4 + adjust signdetamindp[x, y, c4] detaminc4, bb45 →
   etaminbb45 + adjust signdetamindp[x, y, bb45] detaminbb45}) /. {x → w, y → ρ / 90.}
etaminFitbig[w1, ρ2];

```
In[207]:= etaminFitsmall[w_, ρ_] := etaminFitsmall[w, ρ] =
          ((k0 - x (a0 + (*5.4*) b0 x - (a2 + b3 y) Tanh[ (b4 y + c4 y² + bb45 x y)])) ) /. {
             k0 → etamink0 - adjust signdetamindp[x, y, k0] detamink0,
             a0 → etamina0 - adjust signdetamindp[x, y, a0] detamina0,
             b0 → etaminb0 - adjust signdetamindp[x, y, b0] detaminb0,
             a2 → etamina2 - adjust signdetamindp[x, y, a2] detamina2,
             b3 → etaminb3 - adjust signdetamindp[x, y, b3] detaminb3,
             b4 → etaminb4 - adjust signdetamindp[x, y, b4] detaminb4,
             c4 → etaminc4 - adjust signdetamindp[x, y, c4] detaminc4, bb45 →
               etaminbb45 - adjust signdetamindp[x, y, bb45] detaminbb45}) /. {x → w, y → ρ / 90.}
          etaminFitsmall[w1, ρ2];
```

```
In[209]:= (*A data point is "inbounds" if it is
          between the big and small values allowed by the fit.*)
          inbounds = {};
          For[i = 1, i ≤ Length[fitData], i++,
           If[ (etaminFitsmall[wi[i], ρRMSi[i]] < η0mini[i] < etaminFitbig[wi[i], ρRMSi[i]]) ,
             AppendTo[inbounds, i] ] ]
          N[Length[inbounds] / Length[fitData]];
          Round[100. N[Length[inbounds] / Length[fitData]]];
```

For example, the big, best, and small values
  of η^{min} determined by the Library fitting formulas at N = w^{-2} =
16. and ρRMS = 90 rRMS = 3° are {31.9531, 31.83, 31.7098}, in degrees.

For example, the big, best, and small values
  of η^{min} determined by the Library fitting formulas at N = w^{-2} =
16. and ρRMS = 90 rRMS = 24° are {31.6963, 31.5945, 31.4927}, in degrees.

To represent a 'one sigma' range, the big and small values of η^{min} determined by
  the Library fitting formulas should enclose 68% of the data points. There are
 69% of the data points enclosed between the big and small curves.

```
In[654]:= iN = 1;
        Show[{Plot[
            {etaminFitsmall[wList[[iN]], ρ], etaminRhoFit[wList[[iN]], ρ], etaminFitbig[wList[[iN]], ρ]},
            {ρ, 0.001, 80.}, PlotStyle → {Red, Green, Blue}, PlotRange → {{0, 80},
              {Round[etaminRhoFit[wList[[iN]], 50]] - 1.5, Round[etaminRhoFit[wList[[iN]], 0]] + 1.0}},
            GridLines → Automatic, Frame → True, FrameLabel → {"ρ, deg", "η", "η₀ᵐⁱⁿ vs ρRMS"},
            FrameTicks → { {Table[{j, ToString[j] "°"}, {j, Round[etaminRhoFit[wList[[iN]], 50.] -
                    1.5], Round[etaminRhoFit[wList[[iN]], 0] + 1.5]}], None}  ,
              {Table[i, {i, 0, 90, 10}], None} }, ImageSize → 72 × 4], lpRhoη0min[iN],
          Graphics[{Text[StyleForm[ "N = ", FontSize → 12, FontWeight → "Plain"],
              {50, etaminRhoFit[wList[[iN]], 50.] + 0.75}],
            Text[StyleForm[ ToString[nSrcList[[iN]], InputForm, NumberMarks → False],
              FontSize → 12, FontWeight → "Plain"], {60, etaminRhoFit[wList[[iN]], 50.] + 0.75}]}]}]
        Print["Figure A16: Plus/minus range of the fit
            to Library data for η₀ᵐⁱⁿ. For an active choice of N, here N = ",
          nSrcList[[iN]], ", the fit to parameter η₀ᵐⁱⁿ in Eq. A1 plotted versus the
            RMS radius ρRMS. Big, Best, Small are shaded Blue, Green, Red. "]
        Clear[iN]
```



```
Out[655]=
```
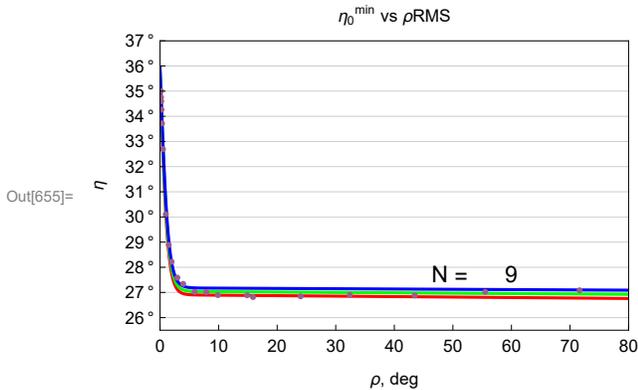
Figure A16: Plus/minus range of the fit
  to Library data for η_0^{min}. For an active choice of *N*, here *N* =
9, the fit to parameter η_0^{min} in Eq. A1 plotted versus the RMS radius
  ρRMS. Big, Small, Best are shaded Blue, Green, Red.

A5b. Uncertainty of the η_0^{max} formulas

```
In[220]:= w1 = 16⁻¹/²·; ρ1 = 24(*degrees*); ρ2 = 3(*degrees*);

In[221]:= percentDiffEtamaxfitnSrc[iN_] := Table[
          Abs[100. (η0maxi[i] - etamaxRhoFit[wi[i], ρRMSi[i]]) / η0maxi[i]], {i, idsFORnSrc[iN]}]
```

In[487]:= **ListPlot[**

    {Table[{nSrcList〚iN〛, Max[percentDiffEtamaxfitnSrc[iN]]}, {iN, Length[nSrcList]}],

     Table[{nSrcList〚iN〛, mean[percentDiffEtamaxfitnSrc[iN]]}, {iN, Length[nSrcList]}]},

    Ticks → {Table[m², {m, {10, 15, 18, 25, 30}}], Automatic},

    PlotLabel → "η_0^{max}: Compare fit to data", AxesLabel → {"N sources", "%"},

    PlotRange → {{0., 1000.}, All}, PlotLegends → {"Max % diff.", "mean % diff."}]

Out[487]=
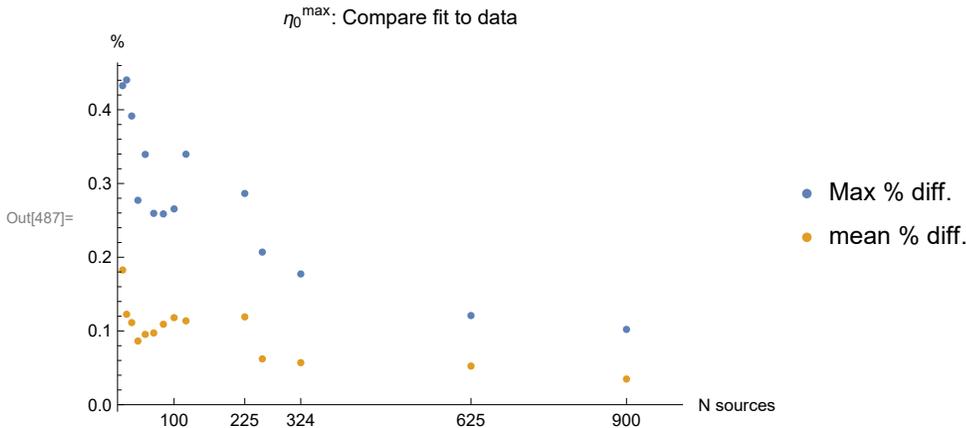
Figure A17: The Library data to fit %diff stats as a function of source
  number *N*. For each *N*, this graph shows the mean % difference and the largest
  %diff of the fitting function for η_0^{max}, over the set of sample sizes ρRMS for *N*.

In[224]:= **(*Determine whether to add or subtract the standard errors of each parameter*)**

(*by finding the sign of the derivative wrt each parameter.*)

detamaxdp[x_, y_, p_] := D[(k0 + x (a0 + b0 x − (a2 + b3 y) Tanh[(b4 y + c4 y² + bb45 x y)])), p]

signdetamaxdp[x_, y_, p_] :=

 Simplify[Sign[detamaxdp[x, y, p] /. nlmrEtamax["BestFitParameters"]], {x > 0, y > 0}]

detamaxdp[x, y, b4];

signdetamaxdp[x, y, b4];

Print["Table of derivatives and their sign, {parameter p, df/dp, sign(df/dp)}:"]

Table[{p, detamaxdp[x, y, p], signdetamaxdp[x, y, p]}, {p, {k0, a0, b0, a2, b3, b4, c4, bb45}}]

Table of derivatives and their sign, {parameter p, df/dp, sign(df/dp)}:

Out[229]= $\Big\{$ {k0, 1, 1}, {a0, x, 1}, {b0, x², 1},

    {a2, − x Tanh[b4 y + bb45 x y + c4 y²], − Sign[Tanh[(64.6949 + 23.0589 x − 62.4627 y) y]]},

    {b3, − x y Tanh[b4 y + bb45 x y + c4 y²], − Sign[Tanh[(64.6949 + 23.0589 x − 62.4627 y) y]]},

    {b4, − x y (a2 + b3 y) Sech[b4 y + bb45 x y + c4 y²]², 1},

    {c4, − x y² (a2 + b3 y) Sech[b4 y + bb45 x y + c4 y²]², 1},

    {bb45, − x² y (a2 + b3 y) Sech[b4 y + bb45 x y + c4 y²]², 1}$\Big\}$

```
In[230]:= (*Adjust the range of the big and small curves to contain 68% of the data.*)
        adjust = 0.933
        Clear[etamaxFitbig, etamaxFitsmall]

Out[230]= 0.933


In[232]:= etamaxFitbig[w_, ρ_] := etamaxFitbig[w, ρ] =
            ((k0 + x (a0 + (*5.4*)b0 x - (a2 + b3 y) Tanh[ (b4 y + c4 y² + bb45 x y )]))) /. {
                k0 → etamaxk0 + adjust signdetamaxdp[x, y, k0] detamaxk0,
                a0 → etamaxa0 + adjust signdetamaxdp[x, y, a0] detamaxa0,
                b0 → etamaxb0 + adjust signdetamaxdp[x, y, b0] detamaxb0,
                a2 → etamaxa2 + adjust signdetamaxdp[x, y, a2] detamaxa2,
                b3 → etamaxb3 + adjust signdetamaxdp[x, y, b3] detamaxb3,
                b4 → etamaxb4 + adjust signdetamaxdp[x, y, b4] detamaxb4,
                c4 → etamaxc4 + adjust signdetamaxdp[x, y, c4] detamaxc4, bb45 →
                  etamaxbb45 + adjust signdetamaxdp[x, y, bb45] detamaxbb45}) /. {x → w, y → ρ / 90.}
        etamaxFitbig[w1, ρ2];


In[234]:= etamaxFitsmall[w_, ρ_] := etamaxFitsmall[w, ρ] =
            ((k0 + x (a0 + (*5.4*)b0 x - (a2 + b3 y) Tanh[ (b4 y + c4 y² + bb45 x y )])) ) /. {
                k0 → etamaxk0 - adjust signdetamaxdp[x, y, k0] detamaxk0,
                a0 → etamaxa0 - adjust signdetamaxdp[x, y, a0] detamaxa0,
                b0 → etamaxb0 - adjust signdetamaxdp[x, y, b0] detamaxb0,
                a2 → etamaxa2 - adjust signdetamaxdp[x, y, a2] detamaxa2,
                b3 → etamaxb3 - adjust signdetamaxdp[x, y, b3] detamaxb3,
                b4 → etamaxb4 - adjust signdetamaxdp[x, y, b4] detamaxb4,
                c4 → etamaxc4 - adjust signdetamaxdp[x, y, c4] detamaxc4, bb45 →
                  etamaxbb45 - adjust signdetamaxdp[x, y, bb45] detamaxbb45}) /. {x → w, y → ρ / 90.}
        etamaxFitsmall[w1, ρ2];


In[236]:= inbounds = {};
        For[i = 1, i ≤ Length[fitData], i++,
         If[ (etamaxFitsmall[wi[i], ρRMSi[i]] < η0maxi[i] < etamaxFitbig[wi[i], ρRMSi[i]]) ,
           AppendTo[inbounds, i] ] ]
        N[Length[inbounds] / Length[fitData]];
        Round[100. N[Length[inbounds] / Length[fitData]]];
```

For example, the big, best, and small values
   of η^{max} determined by the Library fitting formulas at N = w^{-2} =
 16. and ρRMS = 90 rRMS = 3° are {58.4191, 58.282, 58.1407}, in degrees.

For example, the big, best, and small values
   of η^{max} determined by the Library fitting formulas at N = w^{-2} =
 16. and ρRMS = 90 rRMS = 24° are {58.5515, 58.4294, 58.3072}, in degrees.

To represent a 'one sigma' range, the big and small values of η^{max} determined by
   the Library fitting formulas should enclose 68% of the data points. There are
 68% of the data points enclosed between the big and small curves.

```
In[659]:=  iN = 8;
       Show[{Plot[
           {etamaxFitsmall[wList[[iN]], ρ], etamaxRhoFit[wList[[iN]], ρ], etamaxFitbig[wList[[iN]], ρ]},
           {ρ, 0.001, 80.}, PlotStyle → {Red, Green, Blue}, PlotRange → {{0, 80},
              {Round[etamaxRhoFit[wList[[iN]], 0]] - 1.5, Round[etamaxRhoFit[wList[[iN]], 50.]] + 1.0}},
           GridLines → Automatic, Frame → True, FrameLabel → {"ρ, deg", "η", "η₀^max vs ρRMS"},
           FrameTicks → { {Table[{j, ToString[j] "°"}, {j, Round[etamaxRhoFit[wList[[iN]], 0.] - 1.5],
                 Round[etamaxRhoFit[wList[[iN]], 50.] + 1.5]}], None}  ,
              {Table[i, {i, 0, 90, 10}], None} }, ImageSize → 72 × 4], lpRhoη0max[iN],
         Graphics[{Text[StyleForm[ "N = ", FontSize → 12, FontWeight → "Plain"],
             {50, etamaxRhoFit[wList[[iN]], 50.] - 0.75}],
            Text[StyleForm[ ToString[nSrcList[[iN]], InputForm, NumberMarks → False],
              FontSize → 12, FontWeight → "Plain"], {60, etamaxRhoFit[wList[[iN]], 50.] - 0.75}]}]}]
       Print["Figure A18: Plus/minus range of the fit to
           Library data for η₀^max. For an active choice of N , here N = ",
          nSrcList[[iN]], ", the fit to parameter η₀^max in Eq. A2 plotted versus the
           RMS radius ρRMS. Big, Best, Small are shaded Blue, Green, Red. "]
       Clear[iN]
```


Out[660]=
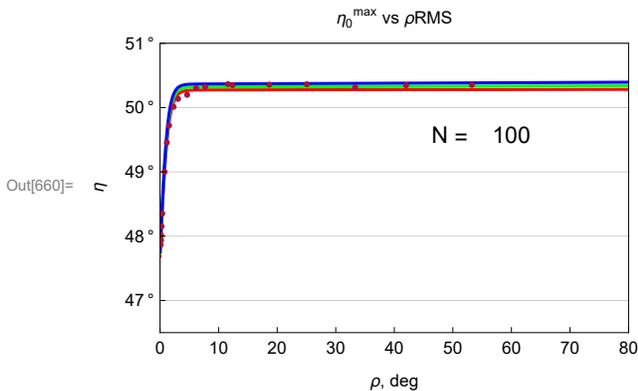
Figure A18: Plus/minus range of the fit
   to Library data for η_0^{max}. For an active choice of *N* , here *N* =
  100, the fit to parameter η_0^{max} in Eq. A2 plotted versus the RMS radius
   ρRMS. Big, Best, Small are shaded Blue, Green, Red.

A5c. Uncertainty of the σ^{min} formulas


```
In[247]:=  w1 = 16^(-1/2.); ρ1 = 24(*degrees*); ρ2 = 3(*degrees*);

In[248]:=  percentDiffsigmaminfitnSrc[iN_] := Table[
           Abs[100. (σmini[i] - sigmaminRhoFit[wi[i], ρRMSi[i]]) / σmini[i]], {i, idsFORnSrc[iN]}]
```

In[489]:= `ListPlot[`

    `{Table[{nSrcList[[iN]], Max[percentDiffsigmaminfitnSrc[iN]]}, {iN, Length[nSrcList]}],`

     `Table[{nSrcList[[iN]], mean[percentDiffsigmaminfitnSrc[iN]]}, {iN, Length[nSrcList]}]},`

    `Ticks → {Table[m`^{2}`, {m, {10, 15, 18, 25, 30}}], Automatic},`

    `PlotLabel → "σ`^{min}`: Compare fit to data", AxesLabel → {"N sources", "%"},`

    `PlotRange → {{0., 1000.}, All}, PlotLegends → {"Max % diff.", "mean % diff."}]`

Out[489]=
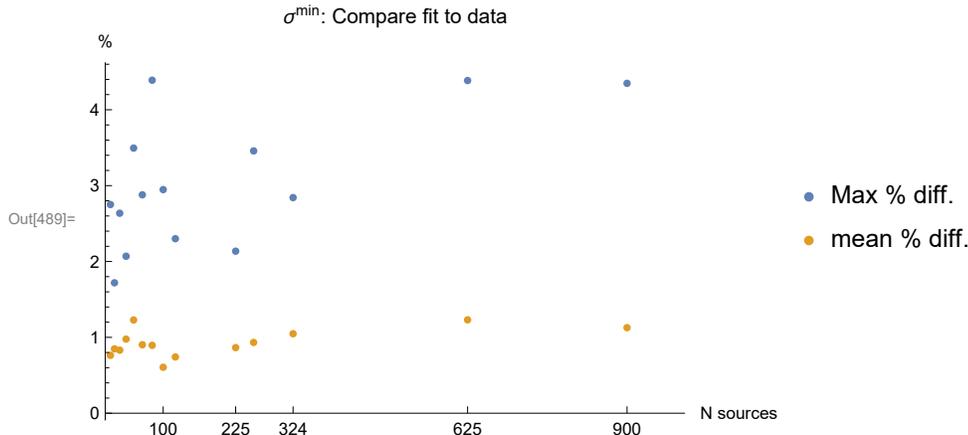
Figure A19: The Library data to fit %diff stats as a function of source
  number *N*. For each *N*, this graph shows the mean % difference and the largest
  %diff of the fitting function for σ^{min}, over the set of sample sizes ρRMS for *N*.

In[251]:= `(*Determine whether to add or subtract the standard errors of each parameter*)`

`(*by finding the sign of the derivative wrt each parameter.*)`

`dsigmamindp[x_, y_, p_] := D[(`$\frac{x}{4.}$` (a0 + b0 x - (a2 + b2 x) Tanh[ (a4 + b4 y )])), p]`

`signdsigmamindp[x_, y_, p_] := Simplify[`

  `Sign[dsigmamindp[x, y, p] /. nlmrSigmamindegA9["BestFitParameters"]], {x > 0, y > 0}]`

`dsigmamindp[x, y, b4];`

`signdsigmamindp[x, y, b4];`

`Print["Table of derivatives and their sign, {parameter p, df/dp, sign(df/dp)}:"]`

`Table[{p, dsigmamindp[x, y, p], signdsigmamindp[x, y, p]}, {p, {a0, b0, a2, b2, a4, b4}}]`

Table of derivatives and their sign, {parameter p, df/dp, sign(df/dp)}:

Out[256]= $\{\{a0, 0.25\,x, 1\}, \{b0, 0.25\,x^2, 1\},$

    $\{a2, -0.25\,x\,\text{Tanh}[a4 + b4\,y], \text{Sign}[\text{Tanh}[1.33841 - 208.211\,y]]\},$

    $\{b2, -0.25\,x^2\,\text{Tanh}[a4 + b4\,y], \text{Sign}[\text{Tanh}[1.33841 - 208.211\,y]]\},$

    $\{a4, -0.25\,x\,(a2 + b2\,x)\,\text{Sech}[a4 + b4\,y]^2, -1\}, \{b4, -0.25\,x\,(a2 + b2\,x)\,y\,\text{Sech}[a4 + b4\,y]^2, -1\}\}$

In[257]:= `(*Adjust the range of the big and small curves to contain 68% of the data.*)`

`adjust = 1.13`

`Clear[sigmaminFitbig, sigmaminFitsmall]`

Out[257]= `1.13`

In[259]:= `sigmaminFitbig[w_, ρ_] :=`

$$\text{sigmaminFitbig}[w, \rho] = \left(\left(\frac{x}{4.} (a0 + b0\ x - (a2 + b2\ x)\ \text{Tanh}[\ (a4 + b4\ y\ )])\right)\right) /. \{$$

`        a0 → sigmamina0 + adjust signdsigmamindp[x, y, a0] dsigmamina0,`
`        b0 → sigmaminb0 + adjust signdsigmamindp[x, y, b0] dsigmaminb0,`
`        a2 → sigmamina2 + adjust signdsigmamindp[x, y, a2] dsigmamina2,`
`        b2 → sigmaminb2 + adjust signdsigmamindp[x, y, b2] dsigmaminb2,`
`        a4 → sigmamina4 + adjust signdsigmamindp[x, y, a4] dsigmamina4,`

$$b4 \to \text{sigmaminb4} + \text{adjust signdsigmamindp}[x, y, b4]\ \text{dsigmaminb4}\Big) /. \{x \to w, y \to \rho\ /\ 90.\}$$

`    sigmaminFitbig[w1, ρ2]`

Out[260]= `4.1966`

In[261]:= `sigmaminFitsmall[w_, ρ_] :=`

$$\text{sigmaminFitsmall}[w, \rho] = \left(\left(\frac{x}{4.} (a0 + b0\ x - (a2 + b2\ x)\ \text{Tanh}[\ (a4 + b4\ y\ )])\right)\right) /. \{$$

`        a0 → sigmamina0 - adjust signdsigmamindp[x, y, a0] dsigmamina0,`
`        b0 → sigmaminb0 - adjust signdsigmamindp[x, y, b0] dsigmaminb0,`
`        a2 → sigmamina2 - adjust signdsigmamindp[x, y, a2] dsigmamina2,`
`        b2 → sigmaminb2 - adjust signdsigmamindp[x, y, b2] dsigmaminb2,`
`        a4 → sigmamina4 - adjust signdsigmamindp[x, y, a4] dsigmamina4,`

$$b4 \to \text{sigmaminb4} - \text{adjust signdsigmamindp}[x, y, b4]\ \text{dsigmaminb4}\Big) /. \{x \to w, y \to \rho\ /\ 90.\}$$

`    sigmaminFitsmall[w1, ρ2]`

Out[262]= `4.08838`

In[263]:= `inbounds = {};`
`    For[i = 1, i ≤ Length[fitData], i++,`
`     If[ (sigmaminFitsmall[wi[i], ρRMSi[i]] < σmini[i] < sigmaminFitbig[wi[i], ρRMSi[i]]) ,`
`       AppendTo[inbounds, i] ]  ]`
`    N[Length[inbounds] / Length[fitData]];`
`    Round[100. N[Length[inbounds] / Length[fitData]]];`

The big, best, and small values of σ^{min} determined by the Library fitting formulas at $N = w^{-2} =$
 16. and ρRMS = 90 rRMS = 3° are {4.1966, 4.14248, 4.08838}, in degrees.

The big, best, and small values of σ^{min} determined by the Library fitting formulas at $N = w^{-2} =$
 16. and ρRMS = 90 rRMS = 24° are {4.19656, 4.14247, 4.08837}, in degrees.

To represent a 'one sigma' range, the big and small values of σ^{min} determined by
   the Library fitting formulas should enclose 68% of the data points. There are
  68% of the data points enclosed between the big and small curves.

```
In[664]:= iN = 1;
    Show[{Plot[{sigmaminFitsmall[wList[[iN]], ρ], sigmaminRhoFit[wList[[iN]], ρ],
        sigmaminFitbig[wList[[iN]], ρ]}, {ρ, 0.001, 80.}, PlotStyle → {Red, Green, Blue},
       PlotRange → {{0, 80}, {Round[sigmaminRhoFit[wList[[iN]], 50]] - 0.5,
          Round[sigmaminRhoFit[wList[[iN]], 0]] + 0.5}}, GridLines → Automatic,
       Frame → True, FrameLabel → {"ρ, deg", "σ", "σ^min vs ρRMS"}, FrameTicks →
        { {Table[{j, ToString[j] "°"}, {j, Round[sigmaminRhoFit[wList[[iN]], 50.] - 1.5],
           Round[sigmaminRhoFit[wList[[iN]], 0] + 1.5], 0.5}], None} ,
          {Table[i, {i, 0, 90, 10}], None} }, ImageSize → 72 × 4], lpρSigmamin[iN],
      Graphics[{Text[StyleForm[ "N = ", FontSize → 12, FontWeight → "Plain"],
          {50, sigmaminRhoFit[wList[[iN]], 50.] + 0.35}],
         Text[StyleForm[ ToString[nSrcList[[iN]], InputForm, NumberMarks → False], FontSize → 12,
           FontWeight → "Plain"], {60, sigmaminRhoFit[wList[[iN]], 50.] + 0.35}]}]}]
    Print["Figure A20: Plus/minus range of the fit
      to Library data for σ^min. For an active choice of N , here N = ",
     nSrcList[[iN]], ", the fit to parameter σ^min in Eq. A1 plotted versus the
       RMS radius ρRMS. Big, Best, Small are shaded Blue, Green, Red. "]
    Clear[iN]
```


Out[665]=
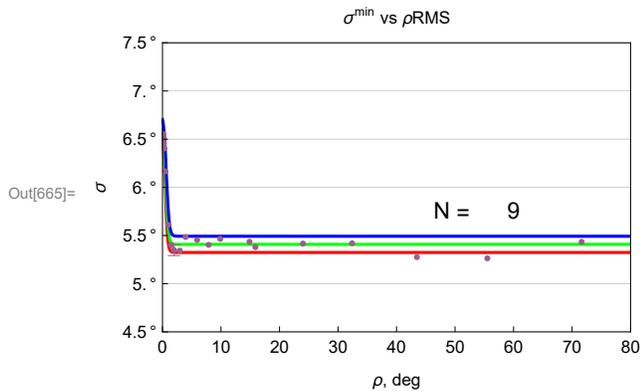
Figure A20: Plus/minus range of the fit
   to Library data for σ^{min}. For an active choice of N , here N =
 9, the fit to parameter σ^{min} in Eq. A1 plotted versus the RMS radius
   ρRMS. Big, Best, Small are shaded Blue, Green, Red.

A5d. Uncertainty of the σ^{max} formulas

```
In[274]:= w1 = 16^(-1/2.); ρ1 = 24(*degrees*); ρ2 = 3(*degrees*);
```

```
In[275]:= percentDiffsigmamaxfitnSrc[iN_] := Table[
      Abs[100. (σmaxi[i] - sigmamaxRhoFit[wi[i], ρRMSi[i]]) / σmaxi[i]], {i, idsFORnSrc[iN]}]
```

In[491]:= `ListPlot[`

    `{Table[{nSrcList[[iN]], Max[percentDiffsigmamaxfitnSrc[iN]]}, {iN, Length[nSrcList]}],`

     `Table[{nSrcList[[iN]], mean[percentDiffsigmamaxfitnSrc[iN]]}, {iN, Length[nSrcList]}],`

     `Table[{nSrcList[[iN]], stanDev[percentDiffsigmamaxfitnSrc[iN]]},`

      `{iN, Length[nSrcList]}]}, Ticks → {Table[m², {m, {10, 15, 18, 25, 30}}], Automatic},`

    `PlotLabel → "σ`max`: Compare fit to data", AxesLabel → {"N sources", "%"},`

    `PlotRange → {{0., 1000.}, All},`

    `PlotLegends → {"Max % diff.", "mean % diff.", "StanDev % diff"}]`

Out[491]=
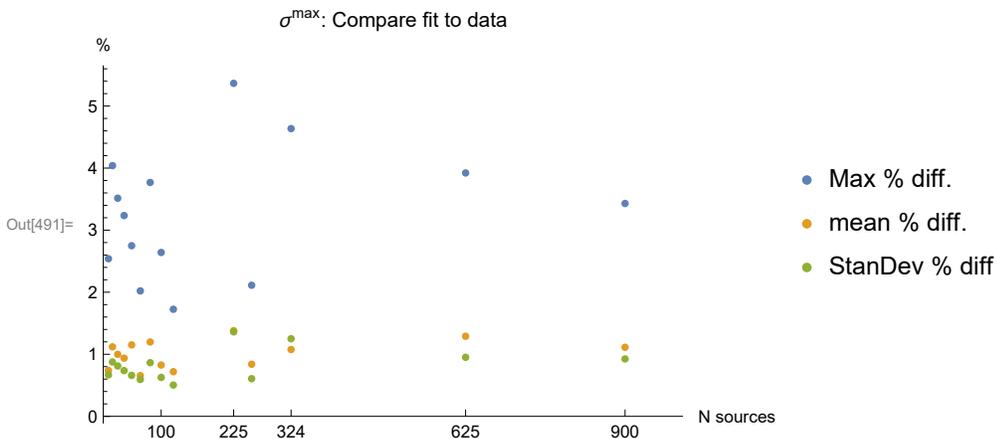
Figure A21: The Library data to fit %diff stats as a function of source
   number N. For each N, this graph shows the mean % difference and the largest
   %diff of the fitting function for σ^{max}, over the set of sample sizes ρRMS for N.

In[278]:= `(*Determine whether to add or subtract the standard errors of each parameter*)`

    `(*by finding the sign of the derivative wrt each parameter.*)`

    `dsigmamaxdp[x_, y_, p_] := D[(`$\frac{x}{4.}$` (a0 + b0 x - (a2 + b2 x) Tanh[ (a4 + b4 y )])), p]`

    `signdsigmamaxdp[x_, y_, p_] := Simplify[`

     `Sign[dsigmamaxdp[x, y, p] /. nlmrSigmamaxdegA9["BestFitParameters"]], {x > 0, y > 0}]`

    `dsigmamaxdp[x, y, b4];`

    `signdsigmamaxdp[x, y, b4];`

    `Print["Table of derivatives and their sign, {parameter p, df/dp, sign(df/dp)}:"]`

    `Table[{p, dsigmamaxdp[x, y, p], signdsigmamaxdp[x, y, p]}, {p, {a0, b0, a2, b2, a4, b4}}]`

    Table of derivatives and their sign, {parameter p, df/dp, sign(df/dp)}:

Out[283]= $\{\{a0, 0.25\,x, 1\}, \{b0, 0.25\,x^2, 1\},$

    $\{a2, -0.25\,x\,\text{Tanh}[a4 + b4\,y], \text{Sign}[\text{Tanh}[1.25969 - 204.817\,y]]\},$

    $\{b2, -0.25\,x^2\,\text{Tanh}[a4 + b4\,y], \text{Sign}[\text{Tanh}[1.25969 - 204.817\,y]]\},$

    $\{a4, -0.25\,x\,(a2 + b2\,x)\,\text{Sech}[a4 + b4\,y]^2, -1\}, \{b4, -0.25\,x\,(a2 + b2\,x)\,y\,\text{Sech}[a4 + b4\,y]^2, -1\}\}$

In[284]:= `(*Adjust the range of the big and small curves to contain 68% of the data.*)`
`adjust = 1.1`
`Clear[sigmamaxFitbig, sigmamaxFitsmall]`

Out[284]= `1.1`

In[286]:= `sigmamaxFitbig[w_, ρ_] :=`

$$\text{sigmamaxFitbig}[w, ρ] = \left(\left(\frac{x}{4.}\ (a0 + b0\ x - (a2 + b2\ x)\ \text{Tanh}[\ (a4 + b4\ y\ )])\right)\right) /. \{$$

`        a0 → sigmamaxa0 + adjust signdsigmamaxdp[x, y, a0] dsigmamaxa0,`
`        b0 → sigmamaxb0 + adjust signdsigmamaxdp[x, y, b0] dsigmamaxb0,`
`        a2 → sigmamaxa2 + adjust signdsigmamaxdp[x, y, a2] dsigmamaxa2,`
`        b2 → sigmamaxb2 + adjust signdsigmamaxdp[x, y, b2] dsigmamaxb2,`
`        a4 → sigmamaxa4 + adjust signdsigmamaxdp[x, y, a4] dsigmamaxa4,`

$$\text{b4} → \text{sigmamaxb4} + \text{adjust signdsigmamaxdp}[x, y, b4]\ \text{dsigmamaxb4}\Big) /. \{x → w, y → ρ / 90.\}$$

`sigmamaxFitbig[w1, ρ2];`

In[288]:= `sigmamaxFitsmall[w_, ρ_] :=`

$$\text{sigmamaxFitsmall}[w, ρ] = \left(\left(\frac{x}{4.}\ (a0 + b0\ x - (a2 + b2\ x)\ \text{Tanh}[\ (a4 + b4\ y\ )])\right)\right) /. \{$$

`        a0 → sigmamaxa0 - adjust signdsigmamaxdp[x, y, a0] dsigmamaxa0,`
`        b0 → sigmamaxb0 - adjust signdsigmamaxdp[x, y, b0] dsigmamaxb0,`
`        a2 → sigmamaxa2 - adjust signdsigmamaxdp[x, y, a2] dsigmamaxa2,`
`        b2 → sigmamaxb2 - adjust signdsigmamaxdp[x, y, b2] dsigmamaxb2,`
`        a4 → sigmamaxa4 - adjust signdsigmamaxdp[x, y, a4] dsigmamaxa4,`

$$\text{b4} → \text{sigmamaxb4} - \text{adjust signdsigmamaxdp}[x, y, b4]\ \text{dsigmamaxb4}\Big) /. \{x → w, y → ρ / 90.\}$$

`sigmamaxFitsmall[w1, ρ2];`

In[290]:= `inbounds = {};`
`For[i = 1, i ≤ Length[fitData], i++,`
` If[ (sigmamaxFitsmall[wi[i], ρRMSi[i]] < σmaxi[i] < sigmamaxFitbig[wi[i], ρRMSi[i]]) ,`
`   AppendTo[inbounds, i] ]  ]`
`N[Length[inbounds] / Length[fitData]];`
`Round[100. N[Length[inbounds] / Length[fitData]]];`

The big, best, and small values of σ^{max} determined by the Library fitting formulas at N = w^{-2} =
 16. and ρRMS = 90(rRMS) = 3° are {4.19264, 4.13188, 4.07117}, in degrees.

The big, best, and small values of σ^{max} determaxed by the Library fitting formulas at N = w^{-2} =
 16. and ρRMS = 90(rRMS) = 24° are {4.19258, 4.13187, 4.07116}, in degrees.

To represent a 'one sigma' range, the big and small values of σ^{max} determined by
   the Library fitting formulas should enclose 68% of the data points. There are
  68% of the data points enclosed between the big and small curves.

In[568]:= `iN = 8;`
```
Show[{Plot[{sigmamaxFitsmall[wList[[iN]], ρ],
    sigmamaxRhoFit[wList[[iN]], ρ], sigmamaxFitbig[wList[[iN]], ρ]}, {ρ, 0.001, 80.},
   PlotStyle → {Red, Green, Blue}, PlotRange → {{0., 80.}, {0., 3.5}},
   (*PlotRange→{{0,80},{Round[sigmamaxRhoFit[wList[[iN]],50]]-0.5,
     Round[sigmamaxRhoFit[wList[[iN]],0]]+0.5}},*)GridLines → Automatic,
   Frame → True, FrameLabel → {"ρ, deg", "σ", "σ^max vs ρRMS"}, FrameTicks →
    { {Table[{j, ToString[j] "°"}, {j, Round[sigmamaxRhoFit[wList[[iN]], 50.] - 1.5],
        Round[sigmamaxRhoFit[wList[[iN]], 0] + 1.5], 0.5]}, None} ,
     {Table[i, {i, 0, 90, 10}], None} }, ImageSize → 72 × 4], lpRhoσmax[iN],
  Graphics[{Text[StyleForm[ "N = ", FontSize → 12, FontWeight → "Plain"],
     {50, sigmamaxRhoFit[wList[[iN]], 50.] + 0.35}],
    Text[StyleForm[ ToString[nSrcList[[iN]], InputForm, NumberMarks → False], FontSize → 12,
      FontWeight → "Plain"], {60, sigmamaxRhoFit[wList[[iN]], 50.] + 0.35}]}]}]
Clear[iN]
```

Out[569]=
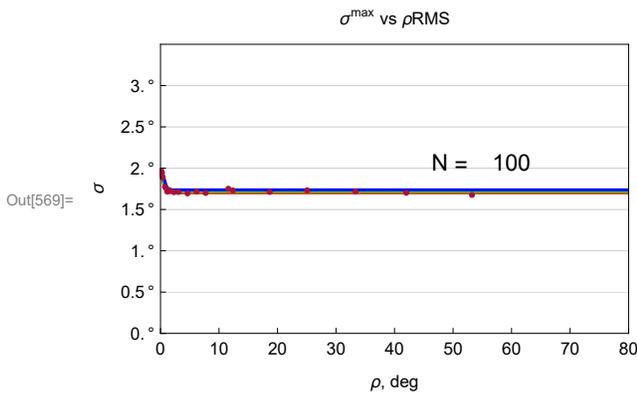
Figure A22: Plus/minus range of the fit to Library data for σ^{max}.
  For an active choice of *N* , the fit to parameter σ^{max} in Eq. A1 plotted
  versus the RMS radius ρRMS. Big, Best, Small are shaded Blue, Green, Red.

A6. Three observed samples. How do the methods compare?

Previously, we investigated the alignment and avoidance correlations of three observed samples, 27 QSOs, 13 QSOs, 99 Stars. With each sample, we applied the Hub Test, calculating the alignment angle function $\overline{\eta}(H)$ in Eq. (1) on a grid of points on the Celestial Sphere. We found and saved the extremes, $\overline{\eta}_{min}$ and $\overline{\eta}_{max}$, of $\overline{\eta}(H)$. The smallest alignment angle $\overline{\eta}_{min}$ reflects the best alignment of the vectors with points on the Sphere and the largest alignment angle $\overline{\eta}_{max}$ shows their mutual avoidance of points on the Sphere. The $\overline{\eta}_{min}$ and $\overline{\eta}_{max}$ measure the correlated behavior of the vectors. See Table A2 for the values of these quantities for the three samples.

Judging the correlations requires estimating the significance of the $\overline{\eta}_{min}$ and $\overline{\eta}_{max}$. How likely is it that random directions at the sources would do better? To answer this question, we applied Direct Method A. We made thousands of random runs by replacing the polarization directions with random directions and calculating $\overline{\eta}(H)$ and recording the extremes, $\overline{\eta}_{min}$ and $\overline{\eta}_{max}$, of $\overline{\eta}(H)$. Histograms of the two extremes yields their probability distributions by fitting with a suitable function. The fitting function formula has two parameters, the location of the distribution's peak, $\eta_0{}^{min}$ and the distribution's half-width σ^{min}. Two more parameters, $\eta_0{}^{max}$ and

$\sigma^{\max}$, determine the probability distribution for the avoidance angle $\overline{\eta}_{\max}$. See Table A3 for the parameters associated with the various probability distributions for the three samples. The method is Direct Method A. These three samples have significance determined by the most reliable method, Method A, and we can use the three samples to check the accuracy of Interpolation Method B and Formula Method C.

      Interpolation Method B and Formula Method C depend on a Library of parameters $\eta_0{}^{\min}$, $\sigma^{\min}$, $\eta_0{}^{\max}$ and $\sigma^{\max}$ determined for various numbers of sources N and region size ρRMS, by applying Direct Method A to each of the simulation samples (N,ρRMS). Methods B and C have been described above. In this section we compare Methods A, B, C, using the previously investigated samples of 27 QSOs, 13 QSOs, 99 Stars.

threeSamplesObservedData:

1. Sample   2. N, number of sources 3. ρRMS, region's root-mean-square radius 4. smallest alignment angle $\overline{\eta}_{\min}$ in degrees 5. largest avoidance angle $\overline{\eta}_{\max}$ in degrees

threeSamplesRandomRunParameters:

1. Sample   2. $\eta_0{}^{\min}$, most likely in random runs   3. $d\eta_0{}^{\min}$, standard error   4. $\sigma^{\min}$, half-width of random run distribution   5. $d\sigma^{\min}$, standard error   6. $\eta_0{}^{\max}$, most likely in random runs   7. $d\eta_0{}^{\max}$, standard error   8. $\sigma^{\max}$, half-width of random run distribution   9. $d\sigma^{\max}$, standard error

The standard errors are determined by Mathematica when the histograms are fit.

A6a. Three observed samples, properties and calculations from data

Definitions:

threeSamplesObservedColumns, threeSamplesObserved   header and data for samples studied previously, Refs. 3,4,5
threeSamplesRandomRunColumns, threeSamplesRandomRunParameters   header and parameters found in previous work
threeSamplesPropertiesmin, threeSamplesDistParametersmin   a portion of the above tables needed for Part I the Article

In[301]:= `threeSamplesObservedColumns = {"Sample", "N", "ρRMS°", "η̄ₘᵢₙ°", "η̄ₘₐₓ°"};`

```
                              "27 QSOs"  27 6.82492 21.094 66.660
threeSamplesObserved =  "13 QSOs"  13 4.72812 10.865 62.665;
                              "99 Stars" 99 6.83803  7.007 83.122
```

In[303]:= `threeSamplesRandomRunColumns =`
`{"Sample", "η₀ᵐⁱⁿ°", "dη₀ᵐⁱⁿ°", "σᵐⁱⁿ°", "dσᵐⁱⁿ°", "η₀ᵐᵃˣ°", "dη₀ᵐᵃˣ°", "σᵐᵃˣ°", "dσᵐᵃˣ°"};`

`threeSamplesRandomRunParameters =`
```
  "27 QSOs" 34.923 0.029 3.272 0.034 55.200 0.031 3.279 0.037
  "13 QSOs" 30.256 0.165 4.638 0.197 60.168 0.171 4.811 0.204;
  "99 Stars" 39.947 0.018 1.737 0.021 50.074 0.011 1.765 0.013
```

Sample	N	ρRMS, deg	$\overline{\eta}_{min}$, deg	$\overline{\eta}_{max}$, deg
27 QSOs	27	6.82492	21.094	66.66
13 QSOs	13	4.72812	10.865	62.665
99 Stars	99	6.83803	7.007	83.122

Table A2: Number of sources, root-mean-square radius, and extremes of the alignment angle
function $\overline{\eta}$(H) for the three samples from previously 'published' work. See Refs. 3,4,5.

Sample	η_0^{min}	$d\eta_0^{min}$	σ^{min}	$d\sigma^{min}$	η_0^{max}	$d\eta_0^{max}$	σ^{max}	$d\sigma^{max}$
27 QSOs	34.923	0.029	3.272	0.034	55.2	0.031	3.279	0.037
13 QSOs	30.256	0.165	4.638	0.197	60.168	0.171	4.811	0.204
99 Stars	39.947	0.018	1.737	0.021	50.074	0.011	1.765	0.013

Table A3: Parameters and their standard errors for the fits to the random run distributions Eqs.
(A1) and (A2), obtained with Direct Method A. All of the parameter values are in degrees.

The following section creates table for the article, Part I the Article. The article in Part I does not include the topic of avoidance, so
"max" quantities are excluded.

In[309]:= **(*This table is created for Part I the Article*)threeSamplesPropertiesmin =**
Partition[Flatten[Table[threeSamplesObserved〚i, j〛, {i, 3}, {j, 4}]], 4]

Out[309]= {{27 QSOs, 27, 6.82492, 21.094},
 {13 QSOs, 13, 4.72812, 10.865}, {99 Stars, 99, 6.83803, 7.007}}

In[310]:= **(*This table is created for Part I the Article*)threeSamplesDistParametersmin =**
Partition[Flatten[Table[{threeSamplesRandomRunParameters〚i, 1〛,
Around[threeSamplesRandomRunParameters〚i, 2〛, threeSamplesRandomRunParameters〚i, 3〛],
Around[threeSamplesRandomRunParameters〚i, 4〛,
threeSamplesRandomRunParameters〚i, 5〛]}, {i, 3}]], 3]

Out[310]= {{27 QSOs, 34.923 ± 0.029, 3.272 ± 0.034},
 {13 QSOs, 30.26 ± 0.17, 4.64 ± 0.20}, {99 Stars, 39.947 ± 0.018, 1.737 ± 0.021}}

Sample	N	ρRMS, deg	$\overline{\eta}_{min}$, deg
27 QSOs	27	6.82492	21.094
13 QSOs	13	4.72812	10.865
99 Stars	99	6.83803	7.007

Sample	η_0^{min}, deg	σ^{min}, deg
27 QSOs	34.923 ± 0.029	3.272 ± 0.034
13 QSOs	30.26 ± 0.17	4.64 ± 0.20
99 Stars	39.947 ± 0.018	1.737 ± 0.021

Table A4: *Left*: Number of sources N, root-mean-square radius ρRMS, and minimum
value of $\overline{\eta}$(H), Eq. (1), of the three samples from earlier articles. *Right*:
Location of the peak η_0^{min} and the half-width σ^{min} of the distribution of the
$\overline{\eta}_{min}$ with random polarization vectors. This is Table 1 in Part I the Article.

A6b. Three observed samples, distribution parameters $\eta_0{}^{\min}$, $\eta_0{}^{\max}$, $\sigma^{\min}$, $\sigma^{\max}$

Definitions:

Probability Distribution Parameters $\eta_0{}^{\min}$, $\sigma^{\min}$, $\eta_0{}^{\max}$, $\sigma^{\max}$ and their standard errors obtained by the 3 Methods

Method A
η0minA , dη0minA , σminA , dσminA , η0maxA , dη0maxA , σmaxA , dσmaxA
Method B
η0minB , dη0minB , σminB , dσminB , η0maxB , dη0maxB , σmaxB , dσmaxB
Method C
η0minC , dη0minC , σminC , dσminC , η0maxC , dη0maxC , σmaxC , dσmaxC

Collect the values by parameter, including the standard error as a $\pm$
compareη0min, compareσmin, compareη0max, compareσmax

Set the range of the vertical axis for the figures
lowyη0min, lowyη0max, lowyσmin, lowyσmax   for $\eta_0{}^{\min}$, $\sigma^{\min}$, $\eta_0{}^{\max}$, $\sigma^{\max}$, respectively

Graph the parameters obtained by the three methods, Methods A, B, C
lpη0minCompare, lpη0maxCompare, lpσminCompare, lpσmaxCompare

For each method, collect the significances of the $\overline{\eta}_{\min}$ for the three samples
sigminA, sigminB, sigminC

For each method, collect the plus/minus range of the significances of the $\overline{\eta}_{\min}$ for the three samples
dsigminAPLUS and dsigminAMINUS for Method A
dsigminBPLUS and dsigminBMINUS for Method B
dsigminCPLUS and dsigminCMINUS for Method C

lpSigminCompare   Graph the significances including their plus/minus values

lpLogSigminCompare   Graph the significance exponents a including their plus/minus values, where $p = 10^{-a}$ and p is
                        the significance


```
(*Collect the 4 distribution parameters and their standard errors by Method A.*)
η0minA = Table[threeSamplesRandomRunParameters〚i, 2〛,
    {i, Length[threeSamplesRandomRunParameters]}] ;(*degrees*)
dη0minA = Table[threeSamplesRandomRunParameters〚i, 3〛,
    {i, Length[threeSamplesRandomRunParameters]}] ;

σminA = Table[threeSamplesRandomRunParameters〚i, 4〛,
   {i, Length[threeSamplesRandomRunParameters]}] ;
(*degrees*)
dσminA = Table[threeSamplesRandomRunParameters〚i, 5〛,
   {i, Length[threeSamplesRandomRunParameters]}] ;
```

```
η0maxA = Table[threeSamplesRandomRunParameters〚i, 6〛,
    {i, Length[threeSamplesRandomRunParameters]}]; (*degrees*)
dη0maxA = Table[threeSamplesRandomRunParameters〚i, 7〛,
    {i, Length[threeSamplesRandomRunParameters]}];

σmaxA = Table[threeSamplesRandomRunParameters〚i, 8〛,
    {i, Length[threeSamplesRandomRunParameters]}]; (*degrees*)
dσmaxA = Table[threeSamplesRandomRunParameters〚i, 9〛,
    {i, Length[threeSamplesRandomRunParameters]}];

(*Collect the 4 distribution parameters and their standard errors by Method B.*)
η0minB = Table[η0minBint[threeSamplesObserved〚i, 2〛^(-1/2), threeSamplesObserved〚i, 3〛],
    {i, Length[threeSamplesObserved]}]; (*degrees*)
dη0minB = Table[dη0minBint[threeSamplesObserved〚i, 2〛^(-1/2), threeSamplesObserved〚i, 3〛],
    {i, Length[threeSamplesObserved]}];

σminB = Table[σminBint[threeSamplesObserved〚i, 2〛^(-1/2), threeSamplesObserved〚i, 3〛],
    {i, Length[threeSamplesObserved]}]; (*degrees*)
dσminB = Table[dσminBint[threeSamplesObserved〚i, 2〛^(-1/2), threeSamplesObserved〚i, 3〛],
    {i, Length[threeSamplesObserved]}];

η0maxB = Table[η0maxBint[threeSamplesObserved〚i, 2〛^(-1/2), threeSamplesObserved〚i, 3〛],
    {i, Length[threeSamplesObserved]}];
(*degrees*)
dη0maxB = Table[dη0maxBint[threeSamplesObserved〚i, 2〛^(-1/2), threeSamplesObserved〚i, 3〛],
    {i, Length[threeSamplesObserved]}];

σmaxB = Table[σmaxBint[threeSamplesObserved〚i, 2〛^(-1/2), threeSamplesObserved〚i, 3〛],
    {i, Length[threeSamplesObserved]}]; (*degrees*)
dσmaxB = Table[dσmaxBint[threeSamplesObserved〚i, 2〛^(-1/2), threeSamplesObserved〚i, 3〛],
    {i, Length[threeSamplesObserved]}];

(*Collect the 4 distribution parameters and their standard errors by Method C.*)
η0minC = Table[etaminRhoFit[threeSamplesObserved〚i, 2〛^(-1/2), threeSamplesObserved〚i, 3〛],
    {i, Length[threeSamplesObserved]}]; (*degrees*)
dη0minC = Table[etaminFitbig[threeSamplesObserved〚i, 2〛^(-1/2), threeSamplesObserved〚i, 3〛] -
    η0minC〚i〛, {i, Length[threeSamplesObserved]}];

σminC = Table[sigmaminRhoFit[threeSamplesObserved〚i, 2〛^(-1/2), threeSamplesObserved〚i, 3〛],
    {i, Length[threeSamplesObserved]}]; (*degrees*)
dσminC = Table[sigmaminFitbig[threeSamplesObserved〚i, 2〛^(-1/2), threeSamplesObserved〚i, 3〛] -
    σminC〚i〛, {i, Length[threeSamplesObserved]}];
```

```
η0maxC = Table[etamaxRhoFit[threeSamplesObserved[[i, 2]]^(-1/2), threeSamplesObserved[[i, 3]]],
    {i, Length[threeSamplesObserved]}] ;(*degrees*)
dη0maxC = Table[etamaxFitbig[threeSamplesObserved[[i, 2]]^(-1/2), threeSamplesObserved[[i, 3]]] -
    η0maxC[[i]], {i, Length[threeSamplesObserved]}];

σmaxC = Table[sigmamaxRhoFit[threeSamplesObserved[[i, 2]]^(-1/2), threeSamplesObserved[[i, 3]]],
    {i, Length[threeSamplesObserved]}]; (*degrees*)
dσmaxC = Table[sigmamaxFitbig[threeSamplesObserved[[i, 2]]^(-1/2), threeSamplesObserved[[i, 3]]] -
    σmaxC[[i]], {i, Length[threeSamplesObserved]}];
```

In[337]:=
```
compareη0min = Table[{threeSamplesObserved[[i, 1]], Around[η0minA[[i]], dη0minA[[i]] ],
    Around[η0minB[[i]], dη0minB[[i]] ], Around[η0minC[[i]], dη0minC[[i]] ]}, {i, 3}];
compareσmin = Table[{threeSamplesObserved[[i, 1]], Around[σminA[[i]], dσminA[[i]] ],
    Around[σminB[[i]], dσminB[[i]] ], Around[σminC[[i]], dσminC[[i]] ]}, {i, 3}];
```

Probability distribution parameters for $\overline{\eta}_{min}$ with the three Methods:

(a) peak η_0^{min} :

Sample	A (degrees)	B (degrees)	C (degrees)
27 QSOs	34.923 ± 0.029	34.797 ± 0.015	34.74 ± 0.10
13 QSOs	30.26 ± 0.17	30.250 ± 0.031	30.14 ± 0.16
99 Stars	39.947 ± 0.018	39.678 ± 0.012	39.67 ± 0.05

(b) half-width σ^{min} :

Sample	A (degrees)	B (degrees)	C (degrees)
27 QSOs	3.272 ± 0.034	3.306 ± 0.018	3.23 ± 0.04
13 QSOs	4.64 ± 0.20	4.61 ± 0.04	4.56 ± 0.06
99 Stars	1.737 ± 0.021	1.729 ± 0.015	1.728 ± 0.015

Table A5: Probability distribution parameters η_0^{min} and σ^{min} by three Methods, Direct Method A, Interpolation Method B, Formula Method C. This is Table 2 in Part I the Article.

In[342]:=
```
compareη0max = Table[{threeSamplesObserved[[i, 1]], Around[η0maxA[[i]], dη0maxA[[i]] ],
    Around[η0maxB[[i]], dη0maxB[[i]] ], Around[η0maxC[[i]], dη0maxC[[i]] ]}, {i, 3}];
compareσmax = Table[{threeSamplesObserved[[i, 1]], Around[σmaxA[[i]], dσmaxA[[i]] ],
    Around[σmaxB[[i]], dσmaxB[[i]] ], Around[σmaxC[[i]], dσmaxC[[i]] ]}, {i, 3}];
```

(a) peak η_0^{max}

Sample	A	B	C
27 QSOs	55.200 ± 0.031	55.321 ± 0.021	55.29 ± 0.10
13 QSOs	60.17 ± 0.17	59.780 ± 0.027	59.90 ± 0.16
99 Stars	50.074 ± 0.011	50.346 ± 0.015	50.35 ± 0.05

Sample	A	B	C
27 QSOs	3.28 ± 0.04	3.336 ± 0.025	3.23 ± 0.04
13 QSOs	4.81 ± 0.20	4.591 ± 0.032	4.55 ± 0.07
99 Stars	1.765 ± 0.013	1.728 ± 0.018	1.729 ± 0.017

(b) half-width σ^{max} :

Table A6: Probability distribution parameters $\eta_\theta{}^{max}$ and σ^{max} in
Eq. (A2) for the distributions of $\overline{\eta}_{max}$ from random runs. There are three
Methods: Direct Method A, Interpolation Method B, Formula Method C. (a) The
location of the peak $\eta_\theta{}^{max}$, in degrees. (b) The half-widths σ^{max}, in degrees.

lowyη0min = {34.3, 29.7, 39.3};

lpη0minCompare[i_] := ListPlot[{Around[η0minA[[i]], dη0minA[[i]]],
Around[η0minB[[i]], dη0minB[[i]]], Around[η0minC[[i]], dη0minC[[i]]]},
PlotRange → {{0.5, 3.5}, {lowyη0min[[i]], lowyη0min[[i]] + 1.}}, GridLines → Automatic,
Frame → True, FrameLabel → {"Method", "$\eta_\theta{}^{min}$, degrees"},
FrameTicks → {Automatic, {{{1., "A"}, {2., "B"}, {3., "C"}}, Automatic}},
PlotLabel → threeSamplesObserved[[i, 1]] ": $\eta_\theta{}^{min}$ (alignment)", ImageSize → 72 × 5]

In[571]:= **GraphicsRow[{lpη0minCompare[1], lpη0minCompare[2]}, ImageSize → 72 × 9]**
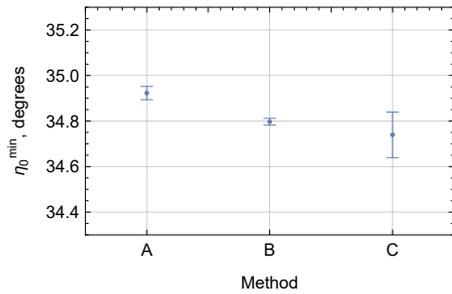
Out[571]=

Out[575]=

Figure A23: Graphs of the probability distribution parameter η_0^{min} in Eq. (A1)
  by Direct Method A, Interpolation Method B, Formula Method C. The ranges are the
  standard errors calculated by Mathematica when the distributions are fit by Eq. (A1).

```
lowyη0max = (*{54.75,59.4,49.75}*){54.5, 59.5, 49.8};
lpη0maxCompare[i_] := ListPlot[{Around[η0maxA〚i〛, dη0maxA〚i〛],
    Around[η0maxB〚i〛, dη0maxB〚i〛], Around[η0maxC〚i〛, dη0maxC〚i〛]},
   PlotRange → {{0.5, 3.5}, {lowyη0max〚i〛, lowyη0max〚i〛 + 1.}}, GridLines → Automatic,
   Frame → True, FrameLabel → {"Method", "η0max, degrees"},
   FrameTicks → {Automatic, {{{1., "A"}, {2., "B"}, {3., "C"}}, Automatic}},
   PlotLabel → threeSamplesObserved〚i, 1〛 ": η0max (avoidance)", ImageSize → 72 × 5]
```

In[576]:= `GraphicsRow[{lpη0maxCompare[1], lpη0maxCompare[2]}, ImageSize → 72 × 9]`

Out[576]=

Out[578]=
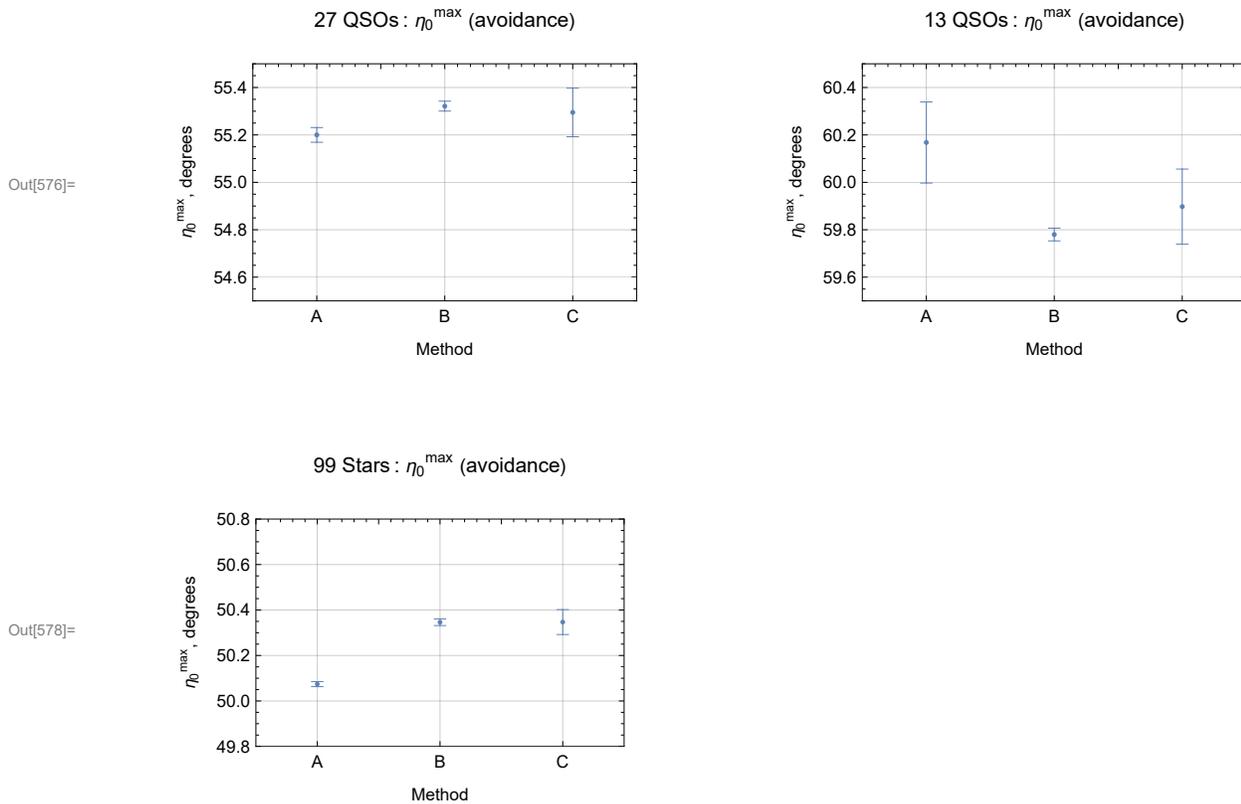
Figure A24: Compare parameter η_0^{max} in Eq. (A2)
  by the 3 methods for the 3 previously studied observed samples.

```
lowyσmin = {2.8, 4.2, 1.3};
lpσminCompare[i_] := ListPlot[{Around[σminA[[i]], dσminA[[i]]], Around[σminB[[i]], dσminB[[i]]],
    Around[σminC[[i]], dσminC[[i]]]}, PlotRange → {{0.5, 3.5}, {lowyσmin[[i]], lowyσmin[[i]] + 1.}},
   GridLines → Automatic, Frame → True, FrameLabel → {"Method", "σ^min, degrees"},
   FrameTicks → {Automatic, {{1., "A"}, {2., "B"}, {3., "C"}}, Automatic}},
   FrameTicks → {Automatic, {{1., "A"}, {2., "B"}, {3., "C"}}, Automatic}},
   PlotLabel → threeSamplesObserved[[i, 1]] ": σ^min (alignment)", ImageSize → 72 × 5]
```

In[581]:= `GraphicsRow[{lpσminCompare[1], lpσminCompare[2]}, ImageSize → 72 × 9]`

Out[581]=

Out[583]=
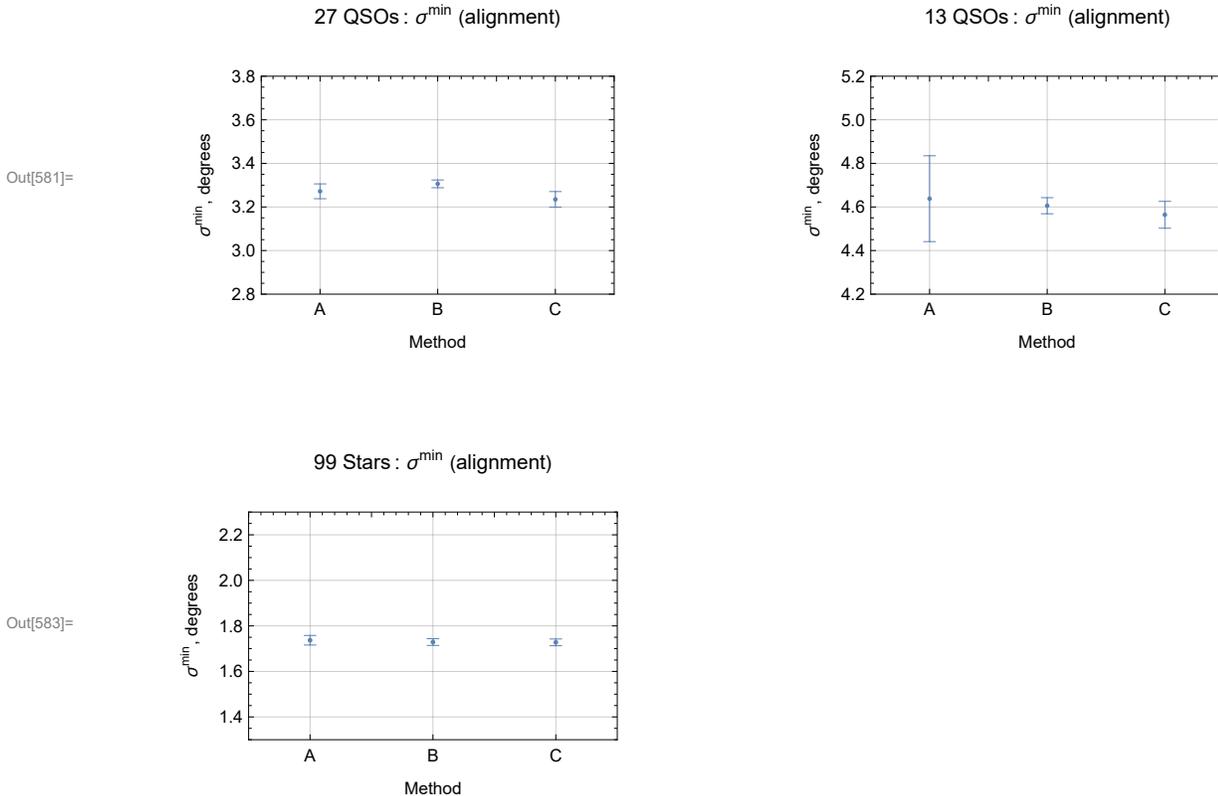
Figure A25: Compare parameter σ^{min} by the 3 methods for the 3 observed samples.

```
lowyσmax = {2.8, 4.2, 1.3};
lpσmaxCompare[i_] := ListPlot[{Around[σmaxA[[i]], dσmaxA[[i]]], Around[σmaxB[[i]], dσmaxB[[i]]],
    Around[σmaxC[[i]], dσmaxC[[i]]]}, PlotRange → {{0.5, 3.5}, {lowyσmax[[i]], lowyσmax[[i]] +
       1.2}}, GridLines → Automatic, Frame → True, FrameLabel → {"Method", "σ^max, degrees"},
   FrameTicks → {Automatic, {{1., "A"}, {2., "B"}, {3., "C"}}, Automatic}},
   FrameTicks → {Automatic, {{1., "A"}, {2., "B"}, {3., "C"}}, Automatic}},
   PlotLabel → threeSamplesObserved[[i, 1]] ": σ^max (avoidance)", ImageSize → 72 × 5]
```

In[586]:= **GraphicsRow[{lpσmaxCompare[1], lpσmaxCompare[2]}, ImageSize → 72 × 9]**

Out[586]=

27 QSOs : σ^{max} (avoidance)　　　　　13 QSOs : σ^{max} (avoidance)

Out[587]=

99 Stars : σ^{max} (avoidance)
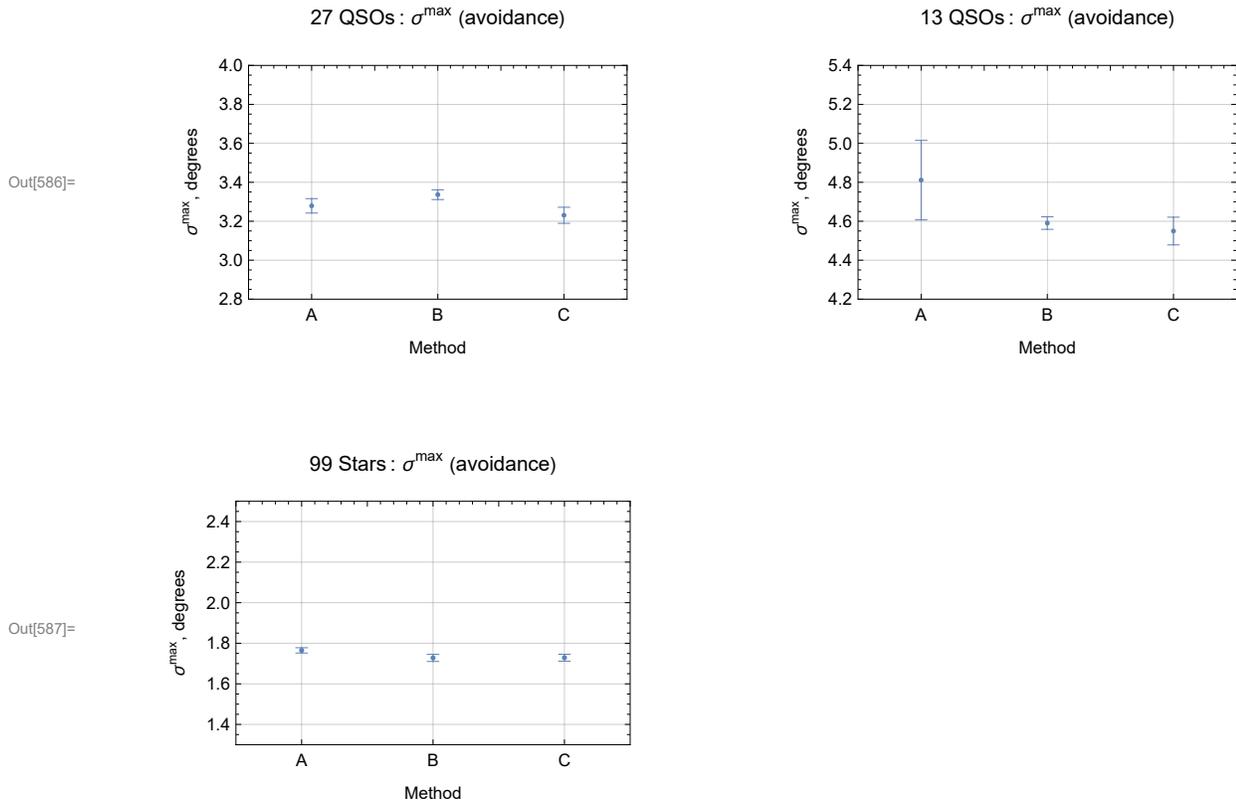
Figure A26: Compare parameter σ^{max} for the 3 methods with 3 observed samples.

A6c. Three observed samples, comparing significances obtained by the 3 Methods A, B, C

In[364]:= **(*Collect the significances of the $\overline{\eta}_{min}$ for the three samples 27 QSOs,**
13 QSOs, 99 Stars by Methods A, B, C.*)
sigminA = Table[signiMIN0[threeSamplesObserved⟦i, 4⟧, η0minA⟦i⟧, σminA⟦i⟧],
**　　{i, Length[threeSamplesObserved]}];**
sigminB = Table[signiMIN0[threeSamplesObserved⟦i, 4⟧, η0minB⟦i⟧, σminB⟦i⟧],
**　　{i, Length[threeSamplesObserved]}];**
sigminC = Table[signiMIN0[threeSamplesObserved⟦i, 4⟧, η0minC⟦i⟧, σminC⟦i⟧],
**　　{i, Length[threeSamplesObserved]}];**

In[367]:= `(*Collect the ± uncertainty range for significances of the $\overline{\eta}_{min}$ .*)`
`dsigminAPLUS[i_] := signiMIN0[threeSamplesObserved[[i, 4]],`
`    η0minA[[i]] - dη0minA[[i]], σminA[[i]] + dσminA[[i]] ] - sigminA[[i]];`
`dsigminAMINUS[i_] := sigminA[[i]] -`
`    signiMIN0[threeSamplesObserved[[i, 4]], η0minA[[i]] + dη0minA[[i]], σminA[[i]] - dσminA[[i]] ];`
`dsigminBPLUS[i_] := signiMIN0[threeSamplesObserved[[i, 4]],`
`    η0minB[[i]] - dη0minB[[i]], σminB[[i]] + dσminB[[i]] ] - sigminB[[i]];`
`dsigminBMINUS[i_] := sigminB[[i]] -`
`    signiMIN0[threeSamplesObserved[[i, 4]], η0minB[[i]] + dη0minB[[i]], σminB[[i]] - dσminB[[i]] ];`
`dsigminCPLUS[i_] := signiMIN0[threeSamplesObserved[[i, 4]],`
`    η0minC[[i]] - dη0minC[[i]], σminC[[i]] + dσminC[[i]] ] - sigminC[[i]];`
`dsigminCMINUS[i_] := sigminC[[i]] -`
`    signiMIN0[threeSamplesObserved[[i, 4]], η0minC[[i]] + dη0minC[[i]], σminC[[i]] - dσminC[[i]] ];`

`lpSigminCompare[i_] := ListPlot[{Around[sigminA[[i]], {dsigminAMINUS[i], dsigminAPLUS[i]}],`
`    Around[sigminB[[i]], {dsigminBMINUS[i], dsigminBPLUS[i]}],`
`    Around[sigminC[[i]], {dsigminCMINUS[i], dsigminCPLUS[i]}]},`
`  PlotRange → {{0.5, 3.5}, All}, GridLines → Automatic,`
`  Frame → True, FrameLabel → {"Method", "p-value"},`
`  FrameTicks → {Automatic, {{{1., "A"}, {2., "B"}, {3., "C"}}, Automatic}}, PlotLabel →`
`    threeSamplesObserved[[i, 1]] ": Significance of $\overline{\eta}_{min}$ (alignment)", ImageSize → 72 × 5]`

In[590]:= `GraphicsRow[{lpSigminCompare[1], lpSigminCompare[2]}, ImageSize → 72 × 9]`

Out[590]=

27 QSOs : Significance of $\overline{\eta}_{min}$ (alignment)

13 QSOs : Significance of $\overline{\eta}_{min}$ (alignment)

Out[591]=

99 Stars : Significance of $\overline{\eta}_{min}$ (alignment)
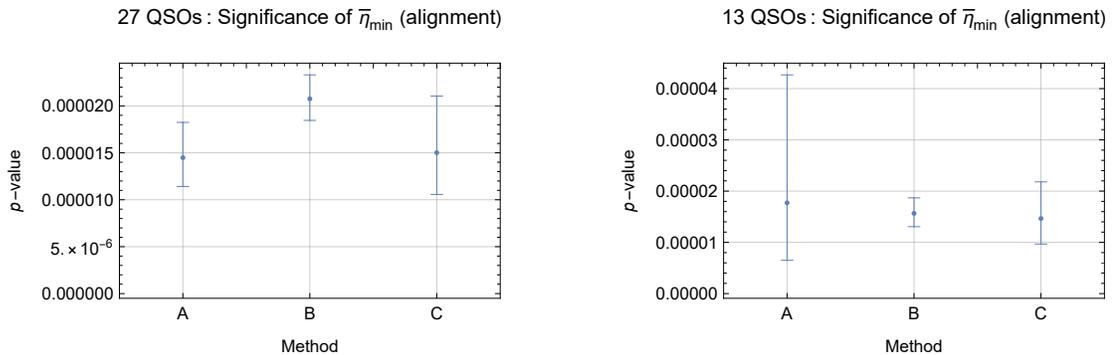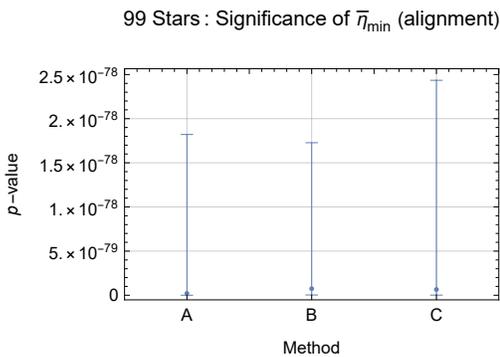
Figure A27: Compare the significance *p* of the observed $\overline{\eta}_{min}$
  (alignment) in Eq. (A3) for the 3 Methods with the 3 observed samples.

```
lpLogSigminCompare[i_] :=
 Show[{ListPlot[-Log[10, {Around[sigminA〚i〛, {dsigminAMINUS[i], dsigminAPLUS[i]}],
       Around[sigminB〚i〛, {dsigminBMINUS[i], dsigminBPLUS[i]}],
       Around[sigminC〚i〛, {dsigminCMINUS[i], dsigminCPLUS[i]}]}],
    PlotRange → {{0.5, 3.5}, All}, GridLines → Automatic, Frame → True,
    FrameLabel → {"Method", "a"},
    FrameTicks → {Automatic, {{{1., "A"}, {2., "B"}, {3., "C"}}, Automatic}},
    PlotLabel → "a for η̄min, where p = 10^-a", ImageSize → 72 × 5],
   Graphics[{Text[StyleForm[threeSamplesObserved〚i, 1〛, FontSize → 12,
       FontWeight → "Plain"], {2.5, -0.97 Log[10, sigminA〚i〛 ]}]}]}]
```

In[595]:= `GraphicsRow[{lpLogSigminCompare[1], lpLogSigminCompare[2]}, ImageSize → 72 × 9]`

Out[595]=

Out[596]=
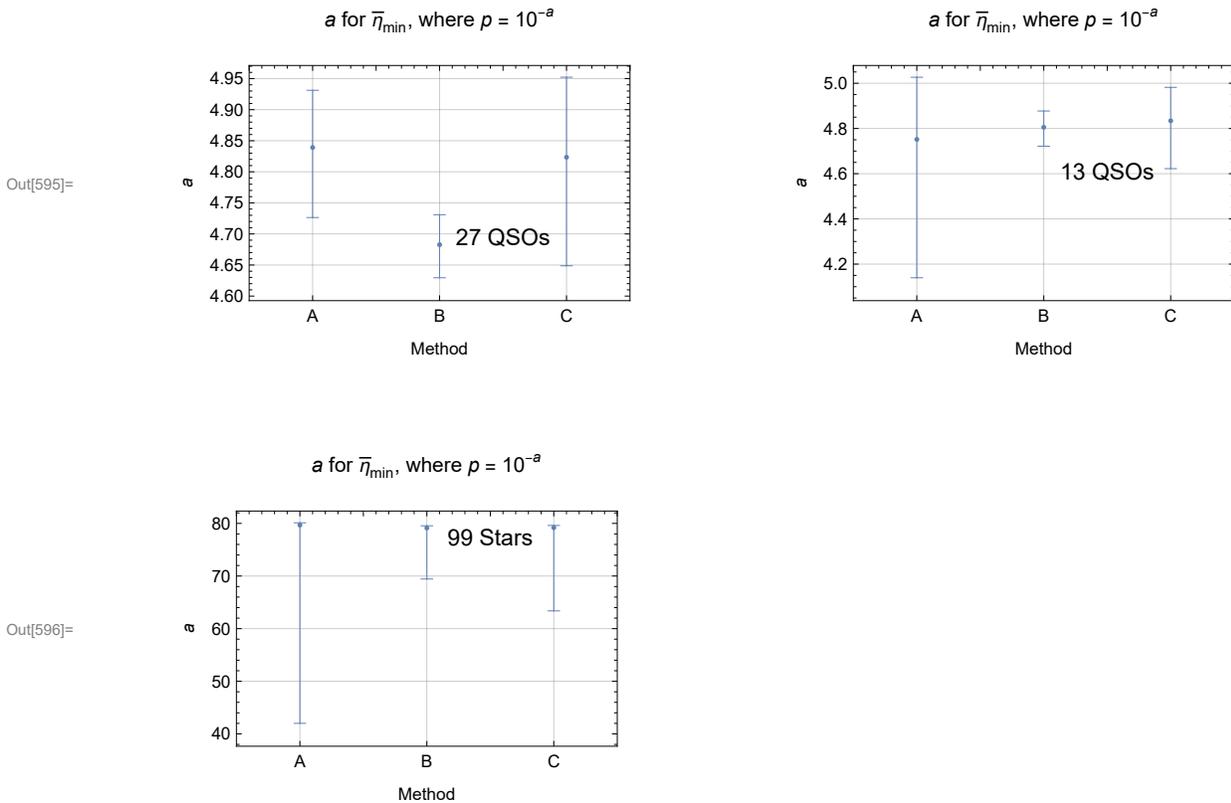
Figure A28: Compare the significance exponent *a,* with p = 10^{-a}, for the observed
  $\overline{\eta}_{min}$ (alignment) in Eq. (A3) by the 3 Methods with the 3 observed samples.

In[381]:= `(*Collect the significances of the` $\overline{\eta}_{max}$ `for the three samples 27 QSOs,`
`13 QSOs, 99 Stars by Methods A, B, C.*)`
`sigmaxA = Table[signiMAX0[threeSamplesObserved[[i, 5]],` η`0maxA[[i]],` σ`maxA[[i]]],`
`{i, Length[threeSamplesObserved]}];`
`sigmaxB = Table[signiMAX0[threeSamplesObserved[[i, 5]],` η`0maxB[[i]],` σ`maxB[[i]]],`
`{i, Length[threeSamplesObserved]}];`
`sigmaxC = Table[signiMAX0[threeSamplesObserved[[i, 5]],` η`0maxC[[i]],` σ`maxC[[i]]],`
`{i, Length[threeSamplesObserved]}];`

In[384]:= `(*Collect the ± uncertainty range for significances of the` $\overline{\eta}_{max}$ `.*)`
`dsigmaxAPLUS[i_] := signiMAX0[threeSamplesObserved[[i, 5]],`
`` η`0maxA[[i]] + d`η`0maxA[[i]],` σ`maxA[[i]] - d`σ`maxA[[i]] ] - sigmaxA[[i]];`
`dsigmaxAMINUS[i_] := sigmaxA[[i]] -`
`signiMAX0[threeSamplesObserved[[i, 5]],` η`0maxA[[i]] - d`η`0maxA[[i]],` σ`maxA[[i]] + d`σ`maxA[[i]] ];`
`dsigmaxBPLUS[i_] := signiMAX0[threeSamplesObserved[[i, 5]],`
`` η`0maxB[[i]] + d`η`0maxB[[i]],` σ`maxB[[i]] - d`σ`maxB[[i]] ] - sigmaxB[[i]];`
`dsigmaxBMINUS[i_] := sigmaxB[[i]] -`
`signiMAX0[threeSamplesObserved[[i, 5]],` η`0maxB[[i]] - d`η`0maxB[[i]],` σ`maxB[[i]] + d`σ`maxB[[i]] ];`
`dsigmaxCPLUS[i_] := signiMAX0[threeSamplesObserved[[i, 5]],`
`` η`0maxC[[i]] + d`η`0maxC[[i]],` σ`maxC[[i]] - d`σ`maxC[[i]] ] - sigmaxC[[i]];`
`dsigmaxCMINUS[i_] := sigmaxC[[i]] -`
`signiMAX0[threeSamplesObserved[[i, 5]],` η`0maxC[[i]] - d`η`0maxC[[i]],` σ`maxC[[i]] + d`σ`maxC[[i]] ];`

`lpSigmaxCompare[i_] := ListPlot[{Around[sigmaxA[[i]], {dsigmaxAMINUS[i], dsigmaxAPLUS[i]}],`
`Around[sigmaxB[[i]], {dsigmaxBMINUS[i], dsigmaxBPLUS[i]}],`
`Around[sigmaxC[[i]], {dsigmaxCMINUS[i], dsigmaxCPLUS[i]}]},`
`PlotRange → {{0.5, 3.5}, All}, GridLines → Automatic,`
`Frame → True, FrameLabel → {"Method", "`*p*`-value"},`
`FrameTicks → {Automatic, {{{1., "A"}, {2., "B"}, {3., "C"}}, Automatic}},`
`PlotLabel → threeSamplesObserved[[i, 1]] ": Significance of` $\overline{\eta}_{max}$ `", ImageSize → 72 × 5]`

In[604]:= `GraphicsRow[{lpSigmaxCompare[1], lpSigmaxCompare[2]}, ImageSize → 72 × 9]`

Out[604]=

27 QSOs : Significance of $\overline{\eta}_{max}$

13 QSOs : Significance of $\overline{\eta}_{max}$

Out[605]=

99 Stars : Significance of $\overline{\eta}_{max}$

Figure A29: Compare the significance p of the observed $\overline{\eta}_{max}$ (avoidance) in
  Eq. (A4) for the 3 Methods with the 3 previously analyzed observed samples.

```
lpLogSigmaxCompare[i_] :=
 Show[{ListPlot[-Log[10, {Around[sigmaxA〚i〛, {dsigmaxAMINUS[i], dsigmaxAPLUS[i]}],
        Around[sigmaxB〚i〛, {dsigmaxBMINUS[i], dsigmaxBPLUS[i]}],
        Around[sigmaxC〚i〛, {dsigmaxCMINUS[i], dsigmaxCPLUS[i]}]}],
     PlotRange → {{0.5, 3.5}, All}, GridLines → Automatic, Frame → True,
     FrameLabel → {"Method", "a"},
     FrameTicks → {Automatic, {{{1., "A"}, {2., "B"}, {3., "C"}}, Automatic}},
     PlotLabel → "a for η̅ₘₐₓ, where p = 10⁻ᵃ", ImageSize → 72 × 5],
    Graphics[{Text[StyleForm[threeSamplesObserved〚i, 1〛, FontSize → 12,
        FontWeight → "Plain"], {2.5, -(*0.97*) Log[10, sigmaxA〚i〛 ]}]}]}]
```

In[608]:= **GraphicsRow[{lpLogSigmaxCompare[1], lpLogSigmaxCompare[2]}, ImageSize → 72 × 9]**

Out[608]=

a for $\overline{\eta}_{max}$, where $p = 10^{-a}$

27 QSOs

a for $\overline{\eta}_{max}$, where $p = 10^{-a}$

13 QSOs
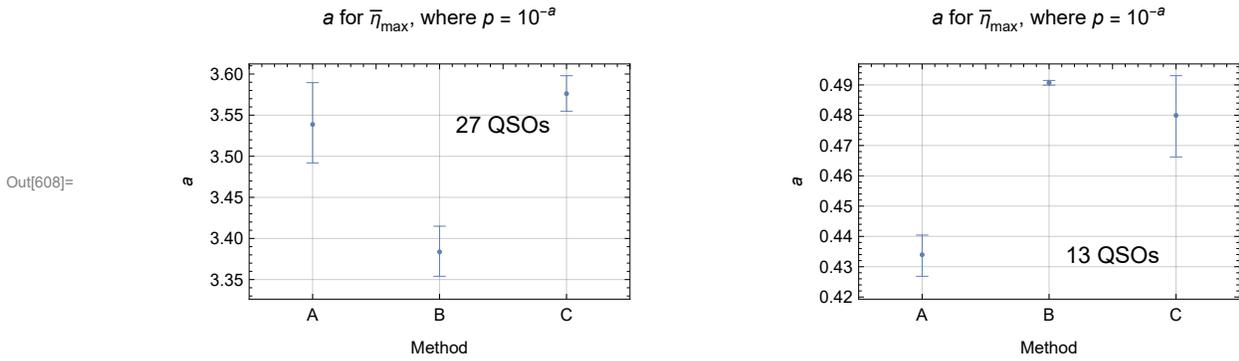
In[609]:= **GraphicsRow[{lpLogSigmaxCompare[3]}, ImageSize → 72 × 4.5]**

Out[609]=

a for $\overline{\eta}_{max}$, where $p = 10^{-a}$

99 Stars
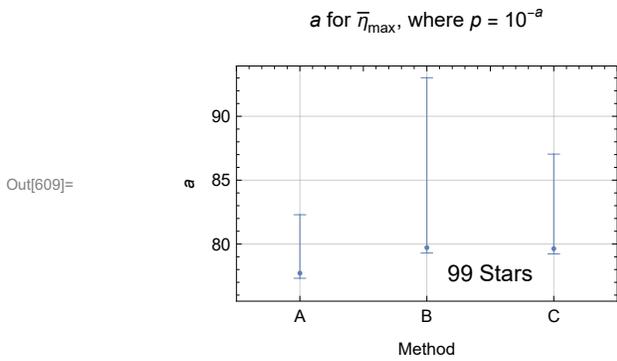
Figure A30: Compare the significance exponent *a*, with $p = 10^{-a}$, for the observed $\overline{\eta}_{max}$ (avoidance) in Eq. (A4) by the 3 Methods with the 3 observed samples.