Asymmetrical uncertainties

Antonio Possolo\(^1\), Christos Merkatas\(^1\) and Olha Bodnar\(^2\)

\(^1\) National Institute of Standards and Technology, Gaithersburg, MD, United States of America
\(^2\) Unit of Statistics, School of Business, Örebro University, SE-70182 Örebro, Sweden, and National Institute of Standards and Technology, Gaithersburg, MD, United States of America

E-mail: antonio.possolo@nist.gov, christos.merkatas@nist.gov and olha.bodnar@mdh.se

Received 29 April 2019, revised 10 June 2019
Accepted for publication 18 June 2019
Published 9 July 2019

Abstract

In several disciplines, measurement results occasionally are expressed using coverage intervals that are asymmetric relative to the measured value. The conventional treatment of such results, when there is the need to propagate their uncertainties to derivative quantities, is to replace the asymmetric uncertainties by ‘symmetrized’ versions thereof. We show that such simplification is unnecessary, illustrate how asymmetry may be modeled and recognized explicitly, and propagated using standard Monte Carlo methods. We present three distributions (Fechner, skew-normal, and generalized extreme value), among many available alternatives, that can be used as models for asymmetric uncertainties associated with scalar input quantities, in the context of the measurement model considered in the GUM. We provide an example where such uncertainties are propagated to the uncertainty of a ratio of mass fractions. We also show how a similar, model-based approach can be used in the context of data reductions from interlaboratory studies and other consensus building exercises where the reported uncertainties are expressed asymmetrically, illustrating the approach to obtain consensus estimates of the absorption cross-section of ozone, and of the distance to galaxy M83 in the Virgo cluster.

Keywords: uncertainty, ozone, black hole, skew-normal, interlaboratory study, symmetrization, consensus

[Supplementary material for this article is available online](https://doi.org/10.1088/1681-7575/ab2a8d)

(Some figures may appear in colour only in the online journal)

\[1\] Introduction

The following are but a few illustrative measurement results that comprise uncertainty evaluations expressed in the form of coverage regions that are not centered at, hence are asymmetrical relative to the measured value:

- Gavriluyk \textit{et al.}[19] report a measurement of the half-life of the double-\(\beta\)-decay of \(^{78}\text{Kr}\), as \((9.2^{+3.5}_{-2.6} \pm 1.3) \times 10^{21}\) year [12, page 769]. The ‘9.2’, qualified with ‘left’ (‘\(-2.6\)’) and ‘right’ (‘\(+5.5\)’) uncertainties, expresses contributions from statistical errors, while the ‘1.3’ expresses contributions from systematic errors [12, section 5.2.1]. The interval \([9.2 - \sqrt{2.6^2 + 1.3^2}, 9.2 + \sqrt{5.5^2 + 1.3^2}]\) may then be regarded as expressing the belief that, with approximately 68% probability, it includes the true value of the half-life: that is, as a ±1-sigma interval assuming that the uncertainty is expressed using a Gaussian distribution.

- The Event Horizon Telescope Collaboration [41] estimate the distance from Earth to the black hole at the center of the M87 galaxy in the Virgo cluster as \(16.8^{+0.8}_{-0.7}\) Mpc.

- Crowther \textit{et al.}[16] report an asymmetric uncertainty associated with their estimate of the age of the R136 star cluster in the Tarantula nebula as \(1.5^{+0.3}_{-0.7} \times 10^6\) year.

- The measurement result that Yoshino \textit{et al.}[44] obtained for the absorption cross-section of ozone at 253.65\,nm was reported as \(1145^{+71}_{-14} \times 10^{-20}\) cm\(^2\) molecule\(^{-1}\).

- Nelson \textit{et al.}[32, page 8657] use an asymmetric probability distribution to characterize the uncertainty associated with the purity, 95.97%, of 3-epi-25(OH)D\(_3\).

Original content from this work may be used under the terms of the Creative Commons Attribution 3.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.
determined using quantitative $^1$H-nuclear magnetic resonance spectroscopy, and report a 95% coverage interval that is asymmetric relative to the measured value, $[94.73\% , 97.59\% ]$. Nelson et al. [33, table 6] report an asymmetric, 95% coverage interval for the mass fraction of benzoic acid in a primary standard for quantitative NMR, as $0.999\,92\,g/g - 0.000\,06\,g/g + 0.000\,04\,g/g$.

- The Bank of England uses the Fechner distribution—discussed in section 2.1—to characterize the uncertainty associated with forecasts of future inflation, thus capturing perceived asymmetries between upside and downside risks [15]. For example, in August 2016, the Bank’s Monetary Policy Committee forecast the annual inflation of the consumer price index, for the third quarter of 2017, in the form of a Fechner distribution with mode 1.93%, uncertainty 1.34%, and difference between the mean and the mode (skew) of 0.10% [29].

- On September 22, 2017, the IceCube Neutrino Observatory detected a high-energy neutrino originating from a source with right ascension $77.43^{+0.95}_{-0.65}$ and declination $+5.72^{+0.30}_{-0.30}$ (degrees of arc, J2000), which together define a box that includes the true source with 90% probability [22].

This contribution addresses the problem of propagating measurement uncertainties expressed asymmetrically, when the corresponding measured values are combined using a suitable function to produce an estimate of a derivative quantity, or when comparable, independent measurements of the same quantity are combined to form a consensus value.

Rather than ‘symmetrizing’ the uncertainties reported asymmetrically before combining the measurement results, for example as proposed by Barlow [7, 8] and Audi et al. [4, appendix A], we address the problem directly by preserving and modeling such uncertainties, first (section 2) in the context of the measurement model considered in the Guide to the expression of uncertainty in measurement (GUM) [26], and second (section 3) in the context of statistical models used in consensus building exercises, for example in interlaboratory studies or meta-analyses [27]. Section 4 highlights some conclusions and recommendations.

In section 2 we present three simple models, selected from among the plethora of asymmetric distributions that may be used for the same purpose: the Fechner distribution, the skew-normal distribution, and the generalized extreme value distribution. In section 2.4, we apply all three to the estimation and evaluation of the uncertainty associated with a ratio of mass fractions of two rare earth elements, and compare the results.

In section 3 we demonstrate how the skew-normal distribution may be used to model measurement uncertainties reported asymmetrically, in a meta-analysis of results for the absorption cross-section of ozone, first when evaluating the uncertainty associated with the DerSimonian–Laird estimator of the consensus value, and second when a Bayesian statistical model is used to compute both an estimate of the consensus value and an evaluation of the associated uncertainty. We also demonstrate how the same techniques can be used to blend three determinations of the distance to galaxy M38 in the Virgo cluster, whose associated uncertainties are expressed asymmetrically.

2. Asymmetric uncertainties in measurement equations

The measurement model considered in the GUM expresses an output quantity $y$ as a known function of several input quantities, $y = f(x_1, \ldots, x_n)$, and handles the uncertainties associated with the input quantities by modeling them as random variables.

The approach to uncertainty propagation described in the GUM uses only the standard deviations of these random variables (and any correlations between them) to produce an approximate evaluation of the standard uncertainty $u(y)$, by application of what is variously known as the delta method [14, section 5.5.28], Gauss’s formula [35, section A.2], or the law of propagation of uncertainty [25, section 7.1.1].

Alternatively, a Monte Carlo method may be used that propagates the full distributions of those random variables, as described by Morgan and Henrion [30] and by the GUM Supplement 1 (GUM-S1) [24]. Both approaches are implemented in the NIST Uncertainty Machine (uncertainty.nist.gov) [28].

The conventional technique described in the GUM requires that the random variables used as models for the input quantities should have finite standard deviations, but not necessarily that their probability distributions be symmetrical. However, even when the input quantities have symmetrical distributions, the distribution of the output quantity need not be either Gaussian or even symmetric.

The Monte Carlo method will recognize and express any asymmetry that may be present, because it uses the full distributions, not just their standard deviations, and its results can be used to produce valid coverage intervals regardless of whether the distribution of the output quantity is symmetric or asymmetric [36].

We assume that each scalar input $x$ is qualified with a left uncertainty $u_l(x)$ and a right uncertainty $u_R(x)$, both positive, but the underlying probability distributions are left unspecified. Very often, these are interpreted so that the interval $[x - u_L(x), x + u_R(x)]$ covers the true value of $x$ with probability approximately 68%; in other words, it is a 1-sigma interval.

In our general treatment, we assume that probabilities $0 < p_L < p_M < p_R < 1$ are specified such that $p_L = \Pr\{X < x - u_L(x)\}$, $p_M = \Pr\{X < x\}$, and $p_R = \Pr\{X > x + u_R(x)\}$, where $X$ denotes the random variable modeling the uncertainty associated with the input $x$. In the examples below, $p_M = 0.5$ (that is, $x$ is regarded as an estimate of the median of the underlying distribution), $p_L$ is either 0.025 or 0.16, and $p_R$ is either 0.84 or 0.975.

In sections 2.1–2.3 we review three different models for asymmetric probability distributions that reproduce the aforementioned, limited information provided about each input quantity. In the online supplementary materials (stacks.iop.org/MET/56/045009/mmedia) we present R code [37] that
finds the values of the parameters of the model that best reproduce $x - \mu_R(x)$, $x$, and $x + \mu_L(x)$ as specified percentiles of the corresponding distribution, using a non-linear, weighted least-squares method. In section 2.4 we present an illustrative example and compare the accuracy that the different models achieve.

2.1. Fechner distribution

The Fechner distribution, also known as the split normal or two-piece normal distribution [43], consists of two half-normal distributions with the same mode, one concentrated to the left of the mode, the other to the right of the mode, and with their respective densities suitably rescaled so that the resulting probability density is continuous. Nelson et al [32] use it to report uncertainties in studies of chemical purity of vitamin D metabolites.

The distribution is determined by three parameters: the mode, $\mu$, and the left and right uncertainties, $\sigma_L > 0$ and $\sigma_R > 0$. The mean is $\mu + \sqrt{2/\pi}(\sigma_R - \sigma_L)$, and the variance is $(1 - 2/\pi)(\sigma_R - \sigma_L^2) + \sigma_L^2\sigma_R$. This distribution can approximate left or right skewed, asymmetric distributions whose Pearson’s moment coefficient of skewness has absolute value no larger than 0.9953, approximately. Figure 1 shows two examples of the Fechner probability density, and algorithm F describes how a Fechner distribution may be chosen to reproduce specified percentiles. R package faniplot [1] provides computational facilities supporting the Fechner distribution.

The Fechner distribution may be used for the ‘symmetrization’ of asymmetric uncertainties very much along the same lines as in Method 2 of Audi et al [4, appendix A]. The idea is to fit an asymmetric distribution to the measurement result (measured value and associated uncertainty, expressed using an asymmetric interval), and then use the mean and standard deviation of the fitted distribution as estimate of the measured value and as standard measurement uncertainty, respectively.

![Figure 1.](image)

Table 1. Examples of two different symmetrizations of asymmetric half-life uncertainties.

| NUCLIDE  | ORIGINAL $T_{1/2}$ | Audi et al [4] | FECHNER |
|----------|---------------------|----------------|---------|
| $^{83}$Mo | $6_{-3}^{+30}$ ms  | 20 ± 17 ms     | 16 ± 17 ms |
| $^{100}$Kr | $7_{-3}^{+11}$ ms | 11 ± 7 ms      | 10 ± 7 ms |
| $^{264}$Hs | $327_{-120}^{+448}$ µs | 490 ± 280 µs  | 454 ± 290 µs |
| $^{266}$Mt | $1.01_{-0.24}^{+0.47}$ ms | 1.1 ± 0.4 ms   | 1.11 ± 0.36 ms |

Table 2. Mean ($R$), standard deviation ($u(R)$), and lower ($L_{95\%}(R)$) and upper ($U_{95\%}(R)$) endpoints of a 95% coverage interval for the true value of the ratio $w_L(\text{La})/w_C(\text{Yb})$. The values under GAM/WEI pertain to the ‘reference’ model for the ratio that is described in the caption of figure 2.

|  | F | SN | GEV | GAM/WEI |
|---|---|----|-----|---------|
| $R$ | 14.5 | 14.5 | 14.4 | 14.0 |
| $u(R)$ | 4.7 | 4.7 | 4.7 | 4.7 |
| $L_{95\%}(R)$ | 6.4 | 6.4 | 6.4 | 6.4 |
| $U_{95\%}(R)$ | 24.7 | 24.6 | 24.6 | 24.6 |

Table 1 displays the symmetrization produced by Audi et al [4], and the symmetrization based on the mean and standard deviation of the Fechner distribution, fitted by non-linear weighted least squares to the original measurement results for the half-life $T_{1/2}$, of each of four different radioactive isotopes.

The choice of values for the weights corresponding to the left, central, and right percentiles depends on where the greatest modeling accuracy is required. In most cases where the left and right percentiles are not far out in the tails, equal weights for all three percentiles produce reasonable results. But when one wishes to ensure that the tails are modeled accurately because they have the greatest impact on the uncertainty, possibly at the price of a less accurate reproduction of the central percentile, then left and right weights larger than the central weight may prove helpful.
Algorithm F

F-1: If the interval $[x - u_L(x), x + u_R(x)]$ has coverage probability about 68%, and $0.645 < u_R(x)/u_L(x) < 1.55$, then the Fechner distribution may be a suitable model for the reported asymmetry. If the coverage probability is approximately 95%, then the model may be suitable when $0.410 < u_R(x)/u_L(x) < 2.44$.

F-2: Use R function parFechner defined in Listing 1 of the online supplementary materials to find the values of $\mu$, $\sigma_L$, and $\sigma_R$ of the best fitting Fechner distribution.

2.2. Skew-normal distribution

The skew-normal distribution [6], with computational facilities implemented in R package sn [5], generalizes the Gaussian distribution to accommodate a modicum of either left or right skewness (figure 1). It is specified by three real-valued parameters: location $\xi$, scale $\omega > 0$, and shape $\alpha$. The Gaussian distribution corresponds to $\alpha = 0$. For distributions in this family, Pearson’s moment coefficient of skewness does not exceed 0.995.

Given a measured value, $x$, its associated left and right uncertainties, $u_L(x)$ and $u_R(x)$, and the probabilities $p_L$, $p_X$ and $p_R$ such that $x - u_L(x), x, \text{ and } x + u_R(x)$ are the 100$p_L$, 100$p_X$, and 100$p_R$ percentiles of the fitted distribution, the best fitting skew-normal distribution is determined by the following procedure.

Algorithm SN

SN-1: If the interval $[x - u_L(x), x + u_R(x)]$ has coverage probability about 68%, and $0.645 < u_R(x)/u_L(x) < 1.55$, then the skew-normal distribution may be a suitable model for the reported asymmetry. If the coverage probability is approximately 95%, then the model may be suitable when $0.410 < u_R(x)/u_L(x) < 2.44$.

SN-2: Use R function parSkewNormal defined in Listing 2 of the online supplementary materials to find the values of $\xi$, $\omega$, and $\alpha$ of the best fitting skew-normal distribution.

2.3. Generalized extreme value distribution

The family of generalized extreme value (GEV) distributions [23, chapter 22], with computational facilities implemented in R package evd [39], includes the Gumbel, Fréchet, and reverse Weibull distributions as special cases. The members of this family are specified by real-valued location ($\mu$), scale ($\sigma > 0$), and shape ($\xi$) parameters. When $\xi > 0$, the support of the distribution is the semi-axis extending from $\mu - \sigma/\xi$ to plus infinity, and when $\xi < 0$ it is the semi-axis extending from minus infinity to $\mu - \sigma/\xi$. The Gumbel distribution corresponds to $\xi = 0$, and its support extends from minus infinity to plus infinity.

Given a measured value, $x$, its associated left and right uncertainties, $u_L(x)$ and $u_R(x)$, and the probabilities $p_L$, $p_X$ and $p_R$ such that $x - u_L(x), x, \text{ and } x + u_R(x)$ are the 100$p_L$, 100$p_X$, and 100$p_R$ percentiles of the fitted distribution, the best fitting GEV distribution is determined by the following procedure. For the GEV densities depicted in figure 1, the fitted values of $\xi$ both are negative, $-0.1465$ for $w_{La}$ and $-0.4967$ for $w_{Yb}$.

Algorithm GEV

GEV-1: The GEV may be an appropriate model for asymmetric distributions with a very wide range of values of the ratio $u_R(x)/u_L(x)$, regardless of whether the coverage probability of $[x - u_L(x), x + u_R(x)]$ is around 68% or 95%.

GEV-2: Use R function parGEV defined in Listing 3 of the supplementary materials to find the values of $\mu$, $\sigma$, and $\xi$ of the best fitting GEV distribution.

2.4. Example—chondrite-normalized mass fraction of Lanthanum

Anders and Grevesse [2] estimate the average mass fraction of ytterbium in carbonaceous chondrites of type CI as

\[ w_{C}(Yb) = 162.5 \mu \text{g kg}^{-1}. \]

Suppose that the associated uncertainty is expressed by an asymmetric, 95% coverage interval ranging from $L_{95\%}(w_{C}(Yb)) = 154.8 \mu \text{g kg}^{-1}$ to $U_{95\%}(w_{C}(Yb)) = 167.3 \mu \text{g kg}^{-1}$.

Suppose also that the measured mass fraction of lanthanum in samples of a marine sediment has median

\[ w_{S}(La) = 2277 \mu \text{g kg}^{-1}, \]

and that, with 95% probability, its true value lies between $L_{95\%}(w_{S}(La)) = 1041 \mu \text{g kg}^{-1}$ and $U_{95\%}(w_{S}(La)) = 3988 \mu \text{g kg}^{-1}$, suggesting an asymmetric probability distribution skewed to the right.

Since mass fractions of rare earths in terrestrial samples are often reported in the form of ratios to the mass fraction of ytterbium in carbonaceous chondrites [3], this example focuses on the ratio $w_{S}(La)/w_{C}(Yb)$ and evaluates the associated uncertainty, which involves propagating the two asymmetrical uncertainties given above.

Application of each of the three models described in section 2 involved the same steps, and produced the results listed in table 2 and depicted in figure 2:

1. Fit the model distribution to the measurement result for $w_{S}(La)$, matching its 2.5th, 50th, and 97.5th percentiles to $L_{95\%}(w_{S}(La)), w_{S}(La)$, and $U_{95\%}(w_{S}(La))$;
2. Fit the model distribution to the measurement result for $w_{C}(Yb)$, matching its 2.5th, 50th, and 97.5th percentiles to $L_{95\%}(w_{C}(Yb)), w_{C}(Yb)$, and $U_{95\%}(w_{C}(Yb))$;
3. Draw a large sample $w_{S,1}(La), \ldots, w_{S,K}(La)$ from the distribution fitted in step (1);
4. Draw a large sample $w_{C,1}(Yb), \ldots, w_{C,K}(Yb)$ from the distribution fitted in step (2);
5. Use the sample of ratios $w_{S,1}(La)/w_{C,1}(Yb), \ldots, w_{S,K}(La)/w_{C,K}(Yb)$ to produce an estimate of the true value of $w_{S}(La)/w_{C}(Yb)$, and an evaluation of the associated uncertainty.

Listing 4 of the online supplementary materials shows R code used to apply the steps above using the skew-normal model. Similar codes produce the corresponding results for the Fechner and GEV models.
Table 3. Determinations of the ozone absorption cross-section \( \sigma \) at 253.65 nm, and evaluations of the associated uncertainty produced by Hodges et al [20], expressed asymmetrically in the form of left and right uncertainties \( u_L(\sigma) \) and \( u_R(\sigma) \). The probability distribution used to model the reported uncertainty is either the Gaussian distribution (G) when \( u_L(\sigma) = u_R(\sigma) \), or the skew-normal distribution (SN) in the other cases, with location \( \xi \), scale \( \omega \), and shape \( \alpha \). In all cases, \( u(\sigma) \) denotes the standard deviation of the model used to describe the reported uncertainty. The labels under LAB are as described by Hodges et al [20], who fully reference the corresponding publications.

| LAB          | \( \sigma \) | \( u_L(\sigma) \) | \( u_R(\sigma) \) | \( u(\sigma) \) | MODEL | \( \xi \) | \( \omega \) | \( \alpha \) |
|--------------|--------------|-----------------|-----------------|----------------|-------|-----------|-----------|-----------|
| AFCRC-59     | 1.141        | 0.033           | 0.037           | 0.0350         | SN    | 1.1407    | 0.0351    | 0.0731    |
| Hearm-61     | 1.147        | 0.022           | 0.022           | 0.0220         | G     |           |           |           |
| JPL-64       | 1.157        | 0.046           | 0.046           | 0.0460         | G     |           |           |           |
| Griggs-68    | 1.129        | 0.022           | 0.023           | 0.0225         | SN    | 1.1287    | 0.0225    | 0.0429    |
| JPL-86       | 1.157        | 0.012           | 0.009           | 0.0105         | SN    | 1.1553    | 0.0105    | 0.0487    |
| UniMin-87    | 1.136        | 0.008           | 0.006           | 0.0070         | SN    | 1.1348    | 0.0070    | 0.0610    |
| HSCA-88      | 1.145        | 0.009           | 0.011           | 0.0100         | SN    | 1.1453    | 0.0100    | 0.0787    |
| UniReims-93  | 1.130        | 0.016           | 0.015           | 0.0155         | SN    | 1.1292    | 0.0155    | 0.0328    |
| UniBremen-99 | 1.150        | 0.030           | 0.030           | 0.0300         | G     |           |           |           |
| UPMC-04      | 1.150        | 0.022           | 0.023           | 0.0225         | SN    | 1.1496    | 0.0225    | 0.0482    |
| NIES-06      | 1.122        | 0.009           | 0.009           | 0.0090         | G     |           |           |           |
| UniBremen-14 | 1.120        | 0.020           | 0.023           | 0.0215         | SN    | 1.1195    | 0.0216    | 0.1090    |
| BIPM-15      | 1.127        | 0.005           | 0.005           | 0.0050         | G     |           |           |           |
| BIPM-16      | 1.124        | 0.010           | 0.010           | 0.0100         | G     |           |           |           |

To validate the proposed methods, some underlying probability distributions have to be assigned to \( w_S(La) \) and to \( w_C(Yb) \). Since we are using the Fechner, skew-normal, or GEV models for the asymmetric uncertainties for these mass fractions, to prove their effectiveness, albeit only in one particular case, we chose to model \( w_S(La) \) and \( w_C(Yb) \) as random variables with asymmetric distributions different from the three that we use as approximants: \( w_S(La) \) as value of a random variable with a gamma probability distribution with shape 9 and scale 253 \( \mu \text{g kg}^{-1} \), and \( w_C(Yb) \) as value of a random variable with a Weibull distribution with shape 64 and scale 163.94 \( \mu \text{g kg}^{-1} \).

Under these assumptions, we know the ‘right’ answer for the distribution of the ratio, and can then compare the approximations obtained using the Fechner, skew-normal, and GEV, against it. The probability density of the corresponding ratio is depicted using a dotted line in figure 2, and its mean, standard deviation, and 2.5th and 97.5th percentiles are listed under GAM/WEI in table 2.

3. Asymmetric uncertainties in interlaboratory studies and meta-analyses

3.1. Absorption cross-section of ozone

Table 3 lists the measurement results for the ozone absorption cross-section \( \sigma \) at 253.65 nm that Hodges et al [20, table 2] used to produce an estimate of \( \sigma \) as a consensus value. Application of algorithm SN from section 2.2 to the cases whose model is SN, under the assumption that each interval ranging from \( \sigma - u_L(\sigma) \) to \( \sigma + u_R(\sigma) \) has approximately 68% coverage and that \( \sigma \) is the median of the corresponding skew-normal distribution, produced the location \( \xi \), scale \( \omega \), and shape \( \alpha \) of this distribution, hence \( u(\sigma) = \omega \sqrt{1 - 2\alpha^2/(\pi(1 + \alpha^2))} \) [5].

Hodges et al [20] computed the consensus value \( \sigma_{DL} = 1.1329 \times 10^{-17} \text{cm}^2 \text{molecule}^{-1} \) using the DerSimonian–Laird procedure for consensus building [17, 27]. Because this is a method-of-moments estimate involving only the estimates of \( \sigma \) and the standard measurement uncertainties \( u(\sigma) \), it is indifferent to any asymmetries of the reported uncertainties and is expected to perform well when the asymmetries are mild, as they are in this dataset.

The asymmetries were taken into account in the evaluation of \( u(\sigma_{DL}) = 0.0035 \text{cm}^2 \text{molecule}^{-1} \) by applying a Monte Carlo method similar to the method described by Koepke et al [27, section 5.2], except that the measurement ‘errors’ were sampled from skew-normal distributions for the laboratories whose probabilistic model (in table 3) was SN, and from Gaussian distributions for the laboratories whose probabilistic model was G. The output of this Monte Carlo method produced a 95% coverage interval for \( \sigma_{DL} \) ranging from \( 1.1260 \times 10^{-17} \text{cm}^2 \text{molecule}^{-1} \) to \( 1.1397 \times 10^{-17} \text{cm}^2 \text{molecule}^{-1} \).

An alternative approach, patterned after the hierarchical Bayesian procedure implemented in the NIST Consensus Builder [27, section 5.3], involves the following steps:

(a) Use a random-effects model for the measured values of the absorption cross-section, \( \sigma_j = \theta_j + \epsilon_j \) for \( j = 1, \ldots, n \), where \( n = 14 \) is the number of results in table 3, the \( \{\theta_j\} \) are a sample from a Gaussian distribution with mean \( \sigma_0 \) (the true value of the absorption cross-section) and standard deviation \( \tau \) (dark uncertainty [27, 42]), and the \( \{\epsilon_j\} \) are measurement ‘errors’ with either skew-normal or Gaussian probability distributions, whose parameter values are listed in table 3.

(b) Assign a vague Gaussian prior distribution to \( \sigma_0 \), with mean \( 0 \times 10^{-17} \text{cm}^2 \text{molecule}^{-1} \) and standard deviation \( 1000 \times 10^{-17} \text{cm}^2 \text{molecule}^{-1} \), and a mildly informative prior distribution to \( \tau \), in the spirit of empirical Bayes: a half-Cauchy distribution with median equal to one half of the MAD of the \( \{\epsilon_j\} \) (MAD, as defined in R function mad, is the median of the absolute deviations from the median, rescaled so that it reproduces the standard
deviation for Gaussian samples). This choice of prior median for τ reflects the fact that the DerSimonian–Laird estimate of τ is 0 cm² molecule⁻¹. 

(c) Use Markov Chain Monte Carlo sampling to draw a sample of values from the posterior distribution of the absorption cross-section, and use this sample to produce a Bayesian estimate \( \sigma_{\text{BAYES}} = 1.1342 \times 10^{-17} \text{ cm}^2 \text{ molecule}^{-1} \), an evaluation of \( u(\sigma_{\text{BAYES}}) = 0.0041 \times 10^{-17} \text{ cm}^2 \text{ molecule}^{-1} \), and a 95% coverage interval for \( \sigma_0 \) ranging from 1.1266\( \times 10^{-17} \text{ cm}^2 \text{ molecule}^{-1} \) to 1.1425\( \times 10^{-17} \text{ cm}^2 \text{ molecule}^{-1} \).

For MCMC we employed a Hamiltonian or Hybrid Monte Carlo (HMC) sampler [9, 18, 31]. HMC samples the parameters in the model by simulating a discretized trajectory of a Hamiltonian dynamical system whose potential function is the negative of the logarithm of the probability density of the posterior distribution of interest.

HMC samplers often produce samples with autocorrelations of smaller absolute value than a simple Metropolis sampler, and the Markov chain also tends to reach equilibrium faster than for Metropolis samplers. We implemented HMC sampling using facilities available in R package rstan [38]. In particular, we used the so-called No-U-Turn sampler [21] that chooses the number of steps taken in the simulated trajectory adaptively.

Since the standardized difference, \( |\sigma_{\text{DL}} - \sigma_{\text{BAYES}}| / \sqrt{u^2(\sigma_{\text{DL}}) + u^2(\sigma_{\text{BAYES}})} \), between the DerSimonian–Laird and Bayes estimates of the absorption cross-section equals 0.24, the two estimates, \( \sigma_{\text{DL}} \) and \( \sigma_{\text{BAYES}} \), are judged not to be significantly different.

The posterior distribution of the dark uncertainty τ is not only concentrated near 0, but the posterior probability is 99.2% that τ is smaller than the median of the standard deviations of the measurement errors for the different laboratories. This is consistent with the fact that the DerSimonian–Laird procedure estimates τ to be 0 (that is, the measured values are not over-dispersed by comparison with their reported measurement uncertainties), which motivated the choice of value assigned to the median of the half-Cauchy prior distribution for τ.

3.2. Distance to black-hole in galaxy M87

The Event Horizon Telescope Collaboration [41, table 9] lists three measurements of the distance from Earth to galaxy M87 in the Virgo cluster, whose combination produced 16.8\( \pm 0.8 \) Mpc (approximately 55 million light-years) as estimate of that distance. At the center of this galaxy there lies the first black hole ever to be delineated in images, which were acquired by the Event Horizon Telescope [40]. The measurement of that distance is an input to the measurement of the mass of the black hole.

The reported distance measurements are 16.75\( \pm 1.04 \) Mpc [10], 16.67\( \pm 0.96 \) Mpc [11], and 16.98\( \pm 0.96 \) Mpc [13], considered to be independent. Each of these results was modeled using a skew-normal distribution. For example, for the first, (16.75 – 1.04) Mpc, 16.75 Mpc, and (16.75 + 1.11) Mpc were treated as the 16th, 50th, and 84th percentiles of the best fitting skew-normal distribution, and similarly for the other two. The 16th and 84th percentiles are chosen because we interpret these asymmetrically reported uncertainties as nominally equivalent to 1-sigma uncertainties. Figure 3 shows that these percentiles reproduce the measurement results exactly.

The same modeling choices and techniques described in section 3.1 produced a sample of size 16000, resulting from thinning and merging the samples drawn from the posterior distribution of the distance, by 4 MCMC chains of length 250000 each. The median of this sample is 16.88 Mpc. The corresponding, complete measurement result may then be expressed as 16.88\( \pm 0.55 \) Mpc, whose difference from the result reported in the original study is statistically insignificant.
4. Conclusions and suggestions

The problem of propagating uncertainties expressed asymmetrically can be solved, without replacing these uncertainties with symmetrized proxies, by modeling them using appropriate asymmetric probability distributions and sampling from these using a suitable Monte Carlo procedure.

We have described how the Fechner, the skew-normal, and the generalized extreme value distributions may be used as models for asymmetric uncertainties associated with mutually independent scalar input quantities of a measurement model of the sort considered in the GUM. Many other asymmetric distributions may be used for the same purpose.

When such input quantities are correlated, then either multivariate versions of these or other distributions may be used similarly, or a copula may be employed that imposes the required dependence structure as Possolo [34] illustrates.

The Fechner and skew-normal distributions can model only modest asymmetry while remaining close to the conventional, Gaussian model for measurement error. However, the generalized extreme value distribution can accommodate more extreme skewness, either to the left or to the right. All three performed comparably well in the example involving a ratio of two mass fractions, where associated uncertainties were expressed asymmetrically and whose ‘true’ models were different from those three.

We have also provided two alternative approaches to the combination of measurement results obtained in an interlaboratory study, using a collection of determinations of the absorption cross-section of ozone as an example. The model for the measured values comprised Gaussian random effects, and either skew-normal or Gaussian laboratory-specific, measurement error models. Both choices can easily be modified, for example, by replacing the Gaussian distribution with a Laplace distribution for the random effects, or by adopting still other asymmetric distributions for the measurement error models.

In one of the alternatives, based on the DerSimonian–Laird procedure, the asymmetries are recognized only during the evaluation of the uncertainty associated with the consensus value. The other alternative was a full-fledged hierarchical Bayesian model, hence was likelihood based, and the asymmetries were recognized fully both in the estimation of the consensus value and in the evaluation of the associated uncertainty. For the data set we chose as illustration, the two alternatives produced statistically indistinguishable results.

We applied the same models and techniques to blend three independent estimates of the distance from Earth to galaxy M87, as an alternative to the procedure that The Event Horizon Telescope Collaboration [41] used for the same purpose, having obtained a result that is statistically indistinguishable from its counterpart in the original study.

The same techniques may be used to model the uncertainties associated with the standard atomic weights that are expressed as intervals whose endpoints are not equidistant from the conventional atomic weight: for example, the standard atomic weight of lithium is [6.938, 6.997], and its conventional atomic weight is 6.94 (www.ciaaw.org, retrieved June 26, 2019).

In all cases, great caution is recommended in verifying that the model adopted to describe the asymmetries can indeed reproduce the measured value, and the endpoints of the reported coverage interval, closely enough for the intended purpose. This verification should be done both numerically and graphically, similarly to what we have done in figure 1.

Acknowledgments

The authors are very grateful to their NIST colleagues Jolene Splett and Thomas Lafarge for reading a draft of this contribution critically and meticulously, and for offering many suggestions and corrections that improved it greatly. The article
also was improved in consequence of comments and suggestions generously offered by Juris Meja (NRC Canada) and by two anonymous referees.

**ORCID iDs**

Antonio Possolo [https://orcid.org/0000-0002-8691-4190](https://orcid.org/0000-0002-8691-4190)

Christos Merkatas [https://orcid.org/0000-0001-7880-7443](https://orcid.org/0000-0001-7880-7443)

Olha Bodnar [https://orcid.org/0000-0003-1359-3311](https://orcid.org/0000-0003-1359-3311)

**References**

[1] Abel G J 2015 Fanplot: an R package for visualising sequential distributions R J 7 15–23

[2] Anders E and Grevesse N 1989 Abundances of the elements: meteoritic and solar Geochim. Cosmochim. Acta 53 197–214

[3] Antonina A N, Shazili N A M, Kamaruzzaman B Y, Ong M C, Abel G J 2015 Fanplot: an R package for visualising sequential distributions R J 7 15–23

[4] Audi G, Kondev F G, Wang M, Huang W J and Naimi S 2017–7

[5] Azzalini A 2018 The R package sn: the skew-normal and related distributions such as the skew-t (version 1.5-2) Università di Padova, Italy http://azzalini.stat.unipd.it/SN

[6] Azzalini A and Capitanio A 2014 The Skew-Normal and Related Families (Cambridge: Cambridge University Press)

[7] Barlow R 2003 Asymmetric errors PHYSSTAT2003: Statistical Problems in Particle Physics, Astrophysics and Cosmology (Menlo Park, CA, 8–11 September 2003) (SLAC National Accelerator Laboratory) pp 250–5 (www.slac.stanford.edu/ econf/C030908/proceedings.html)

[8] Barlow R 2004 Asymmetric statistical errors (arXiv: physics/0406120)

[9] Betancourt M 2018 A conceptual introduction to Hamiltonian Monte Carlo (arXiv:1701.02434v2)

[10] Bird S, Harris W E, Blakeslee J P and Flynn C 2010 The inner halo of M87: a first direct view of the red-giant population Astron. Astrophys. 524 A71

[11] Blakeslee J P, Jordán A, Mei S, Côté P, Ferrarese L, Infante L, Peng E W, Tonry J L and West M J 2009 The ACS Fornax Cluster Survey V. Measurement and recalibration of surface brightness fluctuations and a precise value of the Fornax–Virgo relative distance Astrophys. J. 694 556–72

[12] Patrignani C et al (Particle Data Group) 2016 Review of particle physics Chin. Phys. C 40 100001

[13] Cantiello M, Blakeslee J P, Ferrarese L, Côté P, Roediger J C, Raimondo G, Peng E W, Gwyn S, Durrell P R and Cuillandre J-C 2018 The Next Generation Virgo Cluster Survey (NGVS). XVIII Measurement and calibration of surface brightness fluctuation distances for bright galaxies in Virgo (and beyond) Astrophys. J. 856 126

[14] Casella G and Berger R L 2002 Statistical Inference 2nd edn (Pacific Grove, CA: Duxbury)

[15] Clements M P 2004 Evaluating the Bank of England density forecasts of inflation Econ. J. 114 844–66

[16] Crowther P A et al 2016 The R136 star cluster disected with Hubble Space Telescope/STIS. I. Far-ultraviolet spectroscopic census and the origin of He II λ1640 in young star clusters Mon. Not. R. Astron. Soc. 458 624–59

[17] DerSimonian R and Laird N 1986 Meta-analysis in clinical trials Control. Clin. Trials 7 177–88

[18] Duane S, Kennedy A D, Pendleton B J and Roweth D 1987 Hybrid Monte Carlo Phys. Lett. B 195 216–22

[19] Gavryliuk Y M, Gangapushpe A M, Kazakov V V, Kuzminov V V, Panasenko S I and Ratkevich S S 2013 Indications of 2ν2K capture in 78Kr Phys. Rev. C 87 035501

[20] Hodges J et al 2019 Recommendation of a consensus value for the ozone absorption cross-section at 253.65 nm based on literature review Metrologia 53 034001

[21] Hoffman M D and Gelman A 2014 The No-U-Turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo J. Mach. Learn. Res. 15 1593–623

[22] IceCube Collaboration 2018 Neutrino emission from the direction of the blazar TXS 0506 + 056 prior to the IceCube–170922A alert Science 361 147–51

[23] Johnson N L, Kotz S and Balakrishnan N 1995 Continuous Univariate Distributions vol 2, 2nd edn (New York: Wiley)

[24] Joint Committee for Guides in Metrology 2008 Evaluation of Measurement Data—Supplement 1 to the “Guide to the Expression of Uncertainty in Measurement”—Propagation of Distributions using a Monte Carlo Method (Sèvres: International Bureau of Weights and Measures) (www.bipm.org/en/publications/guides/gum.html BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP and OIML JCGM 101:2008)

[25] Joint Committee for Guides in Metrology 2009 Evaluation of Measurement Data—an Introduction to the “Guide to the Expression of Uncertainty in Measurement” and Related Documents (Sèvres: International Bureau of Weights and Measures) (www.bipm.org/en/publications/guides/vim.html BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP and OIML JCGM 104:2009)

[26] Joint Committee for Guides in Metrology (JCGM) 2008 Evaluation of Measurement Data—Guide to the Expression of Uncertainty in Measurement (Sèvres: International Bureau of Weights and Measures) (www.bipm.org/en/publications/guides/gum.html BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP and OIML JCGM 100:2008; GUM 1995 with minor corrections)

[27] Koepeke A, Lafarge T, Possolo A and Toman B 2017 Consensus building for interlaboratory studies, key comparisons and meta-analysis Metrologia 54 S34–S62

[28] Lafarge T and Possolo A 2015 The NIST uncertainty machine NCSLI Meas. J. Meas. Sci. 10 20–7

[29] Monetary Policy Committee 2016 Inflation Report—August 2016 Bank of England, London, UK (www.bankofengland.co.uk)

[30] Morgan M G and Henrion M 1992 Uncertainty—a Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis (New York: Cambridge University Press)

[31] Neal R M 2011 MCMC using Hamiltonian dynamics Handbook of Markov Chain Monte Carlo ed S Brooks et al (Boca Raton, FL: Chapman and Hall) ch 5, pp 113–62

[32] Nelson M A, Bedner M, Lang B E, Toman B and Lippa K A 2015 Metropolitan approaches to organic chemical purity: primary reference materials for vitamin D metabolites Anal. Bioanalytical Chem. 407 8557–69

[33] Nelson M A, Waters J F, Toman B, Lang B E, Rück A, Breitkuck O, Oskircher M, Windust A and Lippa K A 2018 A new realization of SI for organic chemical measurement: NIST PS1 primary standard for quantitative NMR (Benzoic Acid) Anal. Chem. 90 10510–7

[34] Possolo A 2010 Copulas for uncertainty analysis Metrologia 47 262–71

[35] Possolo A and Iyer H K 2017 Concepts and tools for the evaluation of measurement uncertainty Rev. Sci. Instrum. 88 011301
[36] Possolo A, Toman B and Estler T 2009 Contribution to a conversation about the Supplement 1 to the GUM Metrologia 46 L1–L7

[37] R Core Team 2018 R: a Language and Environment for Statistical Computing (Vienna: R Foundation for Statistical Computing) (www.R-project.org/)

[38] Stan Development Team 2018 RStan: the R interface to Stan 2018 http://mc-stan.org/ R package version 2.18.2

[39] Stephenson A G 2002 evd: Extreme value distributions R News 2 31–2

[40] The Event Horizon Telescope Collaboration 2019 First M87 Event Horizon Telescope results, I. The shadow of the supermassive black hole Astrophys. J. Lett. 875 L1

[41] The Event Horizon Telescope Collaboration 2019 First M87 Event Horizon Telescope results, VI. The shadow and mass of the central black hole Astrophys. J. Lett. 875 L6

[42] Thompson M and Ellison S L R 2011 Dark uncertainty Accreditation Qual. Assur. 16 483–7

[43] Wallis K F 2014 The two-piece normal, binormal, or double Gaussian distribution: its origin and rediscoveries Stat. Sci. 29 106–12

[44] Yoshino K, Freeman D E, Esmond J R and Parkinson W H 1988 Absolute absorption cross-section measurements of ozone in the wavelength region 238–335 nm and the temperature dependence Planet. Space Sci. 36 395–8