EUROPEAN ORGANIZATION FOR NUCLEAR RESEARCH

CERN–EP/2000–026
1 February 2000

Uncertainties due to imperfect knowledge of systematic effects: general considerations and approximate formulae

G. D'Agostini[1)] and M. Raso[2)]

Abstract

Starting from considerations about meaning and subsequent use of asymmetric uncertainty intervals of experimental results, we review the issue of uncertainty propagation. We show that, using a probabilistic approach (the so-called Bayesian approach), all sources of uncertainty can be included in a logically consistent way. Practical formulae for the first moments of the probability distribution are derived up to second-order approximations.

[1)] Università "La Sapienza" and Sezione INFN di Roma 1, Rome, Italy, and CERN, Geneva, Switzerland
Email: giulio.dagostini@roma1.infn.it; URL: http://www-zeus.roma1.infn.it/~agostini
[2)] Sezione INFN di Roma 1, Rome, Italy (currently at Parco Scientifico e Tecnologico d'Abruzzo, L'Aquila, Italy)
Email: mirko.raso@roma1.infn.it; URL: http://www-zeus.roma1.infn.it/~raso

1 Introduction

The combination in quadrature of uncertainties due to systematic effects has become quite standard practice in physics. It is also common practice to add these uncertainties in quadrature to those from random effects. Usually the two kinds of uncertainties are given separately, and the systematic-effect uncertainties are listed individually (at least for the most relevant ones) in order to show the potential of further measurements made with the same apparatus. This combination rule has arisen as a kind of pragmatic procedure [1], in analogy to the combination of standard deviations in probability theory, although cannot justifiably be termed within 'conventional' (i.e. non-Bayesian) statistics. The same is true for the use of the covariance matrix to handle correlated uncertainties.

There is less agreement when the uncertainties due to systematic effects are asymmetric and/or they produce asymmetric shifts in the final quantity of interest, due to nonlinear propagation of uncertainty.[1)] As a numerical example of the latter case, take a quantity Y depending on three 'influence parameters' X_1, X_2 and X_3, which could be calibration constants, environment quantities or theoretical parameters. Suppose that, for the reference values of the X's, the analysis procedure gives (in arbitrary units) $Y = 1.000 \pm 0.050$, where the uncertainty associated with the result is that due to random effects. Consider now that by 'varying reasonably the parameters X_i' (the expression is intentionally left vague for the moment) the following deviations from the central values occur: $\Delta Y_{1\pm} = {}^{+0.060}_{-0.090}$, $\Delta Y_{2\pm} = {}^{+0.098}_{-0.147}$, and $\Delta Y_{3\pm} = {}^{+0.104}_{-0.156}$. An often-used practice is to combine in quadrature separately positive and negative deviations, obtaining the following result: $Y = 1.00 \pm 0.05\,(\text{stat.})\,{}^{+0.15}_{-0.23}\,(\text{syst.})$. Now we are faced with the problem that the result of this *ad hoc* procedure has no theoretical justification. Hence the uncertainty content of the statement (i.e. its probabilistic meaning) is unclear and, as a consequence, it is not obvious how to make use of this information in further analyses, even in the simple case in which the data points are uncorrelated.[2)] As a matter of fact, most people remove the asymmetry in further analysis of the results, using something equivalent to the standard deviations to be used in χ^2 fits. This 'standard deviation' is evaluated either by taking the largest value between Δ_+ and Δ_-, or by averaging the two values (some use the arithmetic, others the geometric average). The result is that in both procedures the uncertainty is symmetrized and the result is considered as if it were described, for all practical purposes, by a Gaussian model around the published best estimate.[3)] Our main worry is not that the combined uncertainties will be incorrect (we anticipate that

---

1) As is well known, asymmetric uncertainties arise also from random effects alone, for example in χ^2 fits if the χ^2 around the minimum is not symmetric. In this paper we focus only on asymmetries deriving from systematic effects or non-linearity.

2) A different problem, although also related to systematic effects, is the treatment of overall uncertainty common to all data, or to large subsets of data, in fitting. In fact, it is now understood that the covariance matrix techniques might lead to undesirable effects on the results [2–4]. We intend to review this problem in a separate paper.

3) A more complicated 'prescription' is described by the PDG [5], which we report here for the convenience of the reader: *"When experimenters quote asymmetric errors $(\delta x)^+$ and $(\delta x)^-$ for a measurement x, the error that we use for that measurement in making an average or a fit with other measurements is a continuous function of these three quantities. When the resultant average or fit $\overline{x}$ is less than $x - (\delta x)^-$, we use $(\delta x)^-$; when it is greater than $x + (\delta x)^+$, we use $(\delta x)^+$. In between, the error we use is a linear function of x. Since the errors we use are functions of the result, we iterate to get the final result. Asymmetric output errors are determined from the input errors assuming a linear relation between the input and the output quantities."* This rule does not seem to be applied by others then the PDG. As examples of other ad hoc procedures, see Refs. [6–8].

the arithmetic average of Δ_+ and Δ_- gives indeed the correct uncertainty in most cases of practical interest), but rather that the result itself can be biased with respect to what one could get using consistently the best knowledge concerning the input quantities, as will be shown in this paper.

The purpose of this paper is to review the issue of uncertainty propagation, starting from general considerations and deducing approximate formulae for practical applications. The issue will be analysed in the framework of the so-called Bayesian inference, though Bayes' theorem will never appear in this paper. This just means that, following physics intuition [1], we consider it natural to talk about probability of true values. As a consequence, for this kind of application probability can only have the meaning of degree of belief.[4)] We will show that the rule of combining in quadrature symmetric uncertainties is a natural consequence of the probabilistic approach we follow, assuming that the uncertainties are conceptually properly defined (although the overall result will not depend on their precise values). Formulae to take into account correlations and nonlinearity effects will also be provided.

The paper is structured as follows. First, in Section 2 we illustrate briefly what we mean by probabilistic treatment of measurement uncertainties and why this is meaningful only within the Bayesian approach. In this approach, the most general and logically consistent way to include all kinds of uncertainties is just a straightforward application of probability calculus. However, the application of the general propagation formula (the derivation of which is given in Appendix A) requires the evaluation of integrals which can rapidly become complicated in real-life problems. Therefore, approximate formulae are derived based on linear and quadratic expansion of the output quantity on the input quantities. This is done in Sections 3 and 5, respectively. Several numerical examples are given, the main ones being discussed in Section 6. Section 7 shows that our approximate formulae can also handle the case in which $\pm 1\,\sigma$ variations on an input quantity produce a shift in the same direction on the output quantity. Correlations are also considered. We describe in the text only the linear case (Section 3) and refer to Appendix B for the more complicated formulae which take into account second-order effects. The important issue of how to model uncertainties due to systematic effects is discussed in Section 4. For practical purposes the probability density function (p.d.f.) of the output quantity can be considered approximately Gaussian also in the non-linear cases and also if correlations are present, thanks to the combinatorial effects analogous to those which make the central limit theorem work. In case of important non-Gaussian contributions in the input quantities, or strong non-linearity effects in the propagation, a detailed evaluation of the p.d.f. is needed, usually done using Monte Carlo methods. However, deviations from normality can be checked from skewness and kurtosis, and we give approximate formulae for these quantities in Appendix B. Finally, we show in Appendix C how it is possible, in principle, to get a rough evaluation of the final p.d.f., if this does not differ much from a Gaussian; 'in principle' because we understand that the method described in Appendix C is perhaps more an academic exercise than a real help to practitioners, who most likely will find it more convenient to solve the problem by Monte Carlo integration. Concluding remarks are given in Section 8.

[4)] For a physicist's introduction to subjective probability and Bayesian inference see Ref. [9], or Refs. [10] and [11] for short accounts.

2 Probabilistic treatment of measurement uncertainties

In the Bayesian approach, probability is associated with uncertainty, whether we are interested in the yet to be observed outcome of a measurement, or in the numerical value of the physics quantity. As a result of the experiment, there will be a p.d.f. $f(\mu\,|\,\text{data}, I)$ associated with the numerical value of the true value (generically called μ), conditioned by the observed data and by the status of information (I) concerning measurement and measurand (the conditions data and I will usually be considered implicit). Although we maintain that the proper way of learning from data is to make use of Bayes' theorem, it is easy to show that in most routine[5)] (or, at least, non critical) measurements the usual methods of analysis can be considered as approximations of Bayesian inference (see also Section 2.9 of Ref. [9]). When these kinds of conditions hold, also the Gaussian approximation is usually rather good. Therefore, hereafter we will consider a result of the kind $\mu = \hat{\mu} \pm \sigma_r$, where σ_r is only due to random effects, equivalent to a Gaussian p.d.f. of μ, as usually perceived by physicists [1,9], i.e.

$$\mu = \hat{\mu} \pm \sigma_r \quad \Longleftrightarrow \quad f(\mu_r\,|\,\text{data}, I) = \frac{1}{\sqrt{2\,\pi}\,\sigma_r}\, e^{-\frac{(\mu_r-\hat{\mu})^2}{2\,\sigma_r^2}}\,, \tag{1}$$

where the symbol μ_r is to remind us that only uncertainties due to random effects are considered in Eq. (1). On the other hand, if one sticks strictly to frequentistic ideas, one gets 'results' which have neither a meaning of certain statements about true values, neither the meaning of probabilistic statements. As a consequence, it is not clear what they mean, neither how they should be combined or propagated in a logically consistent way (see more extended discussion in Refs. [13] and [14]).

In the probabilistic framework in which we are moving, the uncertainties due to systematic effects can be easily and consistently included (at least conceptually, although the numerical implementation can present some technical problems). Indeed, there are several ways to proceed, all leading to the same result, though each way can be more or less intuitive or suitable for a particular application (see Section 2.10.3 of Ref. [9]). The closest one to the spirit of probabilistic inference consists in writing explicitly I to depend on other physics quantities, which could be calibration constants, influence parameters (temperature, pressure, etc), theoretical quantities, and so on, plus other pieces of general knowledge not easy to model ($I_\circ$) and which lead the researchers to behave in a given way and to make reasonable assumptions in the many steps of the experimental work. Let us indicate by $I_\circ$ this general knowledge. The physical quantities on which the result can depend will be called *influence quantities* (or parameters) and will be indicated by h_i. The entire set of influence quantities will be indicated by $\mathbf{h} = \{h_1, h_2, \ldots, h_n\}$. In general, the result (1), which takes into account only random effects, is obtained using the set of best estimates[6)] of the parameters ($\mathbf{h}_\circ$), i.e.

$$f(\mu_r\,|\,\text{data}, I) \equiv f(\mu\,|\,\text{data}, \mathbf{h}_\circ, I_\circ). \tag{2}$$

[5)] 'Routine' in the sense of Ref. [12], which applies also for most measurements in frontier research.

[6)] Note that sometimes the result is meant, instead, for a *nominal* set of parameters, which are not necessarily the ones in which the experimentalists believe mostly. In principle, it is possible to derive the approximate formulae considering nominal values, as has been done by one of the authors in Ref. [9]. However, when second-order effects are taken into account the formalism becomes more complicated and, therefore, we prefer to start from expected values and standard deviations on the influence quantities. The reader should be aware that the use of nominal values different from expected values produce shifts in the results which need to be corrected [9], similar to those produced by nonlinearity effects which will be discussed later in this paper.

The most general inference on μ will depend, instead, on all possible values of $\mathbf{h}$, and the resulting p.d.f. will be $f(\mu \,|\, \mathrm{data}, I_\circ)$. Probability theory teaches us how to get rid of the uncertainty about the exact value of the influence parameters. Describing the uncertainty about the influence parameters with the joint p.d.f. $f(\mathbf{h} \,|\, I_\circ)$, we obtain that the probabilistic result which takes into account systematic uncertainties is given by

$$f(\mu \,|\, \mathrm{data}, I_\circ) = \int f(\mu \,|\, \mathrm{data}, \mathbf{h}, I_\circ)\, f(\mathbf{h} \,|\, I_\circ)\, \mathrm{d}\mathbf{h}\,. \tag{3}$$

(We use the symbol $f(\cdot)$ for all p.d.f.'s, and implicitly consider the integrals done over the range of definition of the variables.) If the influence parameters are perfectly known, i.e. $f(\mathbf{h} \,|\, I_\circ) = \Pi_i \delta(h_i - h_{\circ_i})$, we reobtain Eq. (2), and hence Eq. (1), i.e. the uncertainty is that due to the random effects alone. Hereafter we shall consider implicit the general condition $I_\circ$.

As an example, let us consider the result of a single measurement yielding the observed value $X = x$, and in which the most relevant systematic effect is a not exactly known offset Z, the uncertainty about which is described by a Gaussian p.d.f. around zero and standard deviation σ_z. We have (see Ref. [10] for further details):

$$\begin{aligned} f(\mu \,|\, x, z) &= \frac{1}{\sqrt{2\,\pi}\,\sigma_r}\, e^{-\frac{(\mu - x - z)^2}{2\,\sigma_r^2}}\,, & (4)\\ f(\mu \,|\, x) &= \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\,\pi}\,\sigma_r}\, e^{-\frac{(\mu - x - z)^2}{2\,\sigma_r^2}}\, \frac{1}{\sqrt{2\,\pi}\,\sigma_z}\, e^{-\frac{z^2}{2\,\sigma_z^2}}\, \mathrm{d}z & (5)\\ &= \frac{1}{\sqrt{2\,\pi}\,\sqrt{\sigma_r^2 + \sigma_z^2}}\, e^{-\frac{(\mu - x)^2}{2\,(\sigma_r^2 + \sigma_z^2)}}\,. & (6) \end{aligned}$$

The p.d.f. which describes μ is still centred around the observed value x, but with a standard deviation which is the quadratic combination of σ_r and σ_z. The commonly used combination rule is recovered, but now as a theorem with well-defined conditions, instead of just a 'prescription'.

An alternative way of including systematic effects, very convenient for deriving approximate formulae, consists in considering a function g which relates the true value μ to μ_r and of the influence factors, i.e.

$$\mu = g(\mu_r, \mathbf{h})\,. \tag{7}$$

Therefore the uncertainty about μ is obtained from the propagation of uncertainties about μ_r and $\mathbf{h}$ (see Appendix A):

$$f(\mu) = \int f(\mu_r, \mathbf{h}) \cdot \delta(\mu - g(\mu_r, \mathbf{h}))\, \mathrm{d}\mu_r\, \mathrm{d}\mathbf{h}\,. \tag{8}$$

This formula also has a simple interpretation which makes it convenient for Monte Carlo evaluation:[7)] the infinitesimal probability element $f(\mu)\,\mathrm{d}\mu$ depends on 'how many' elements $\mathrm{d}\mu_r\, \mathrm{d}\mathbf{h}$ contribute to it, each element weighted with the p.d.f. calculated in $\{\mu_r, \mathbf{h}\}$.

[7)] For example, this is the basic reasoning behind the methods used by several authors to evaluate the p.d.f.'s of physical quantities, like the direct CP-violation parameter ϵ'/ϵ [15, 16], or the parameters of the quark mixing matrix [16, 17]: Beliefs on the input quantities (experimental and theoretical quantities) are propagated into the beliefs on ϵ'/ϵ, or on ρ and η, respectively. The result has a clear probabilistic meaning and is, as we shall see, rather insensitive to the exact shape of the input p.d.f.'s. Instead, the so-called 'scanning' [15, 16] or other *ad hoc* procedures (see e.g. Ref. [18]) do not have such an intuitive interpretation and can be misleading, especially when the very conservative 'regions of confidence' produced by these methods are improperly called 95% C.L. regions.

At this point, an interesting observation is that μ_r and h_i have a symmetric role in the propagation of uncertainty, and therefore there is no real need to keep them separate in the formalism. Therefore, following the ISO *Guide* [19], we prefer to speak, generically, of *input quantities*, and to indicate them all by X_i. We indicate the *output quantity* by Y. In many problems of interest also the output quantities might be more than one. Their values are evaluated using the common data or (which is conceptually equivalent) making use of the same instrumentation. In such a case we have to consider correlations among the output quantities even if the input quantities were uncorrelated. Hereafter we will indicate the generic functions $g_i(\mathbf{X})$ with the same symbol as the output variables, and speak about $Y_i = Y_i(\mathbf{X})$.

3 Linear expansion around E[X], role of central limit theorem and numerical implementation of the linear propagation

Having illustrated the general solution to the problem, it is now interesting to obtain approximate formulae which, in many practical cases, save us from making complicated integrals. The case in which the dependence of Y_j on X_i is approximately linear in a range of several standard deviations around their expected value is well known, and leads to the standard propagation formula of variances and covariances. What is less well known is that the use of these formulae is justified only if the numerical values of the physics quantities are associated with random numbers (or uncertain numbers), and the probability is meant as degree of belief [9, 10].

The first-order expansion of $Y_i(\mathbf{X})$ around the expected values of X_i gives

$$Y_j \approx Y_j(\mathrm{E}[\mathbf{X}]) + \sum_i \left.\frac{\partial Y_j}{\partial X_i}\right|_{\mathrm{E}[\mathbf{X}]} (X_i - \mathrm{E}[X_i]) \tag{9}$$

$$\approx k + \sum_i \left.\frac{\partial Y_j}{\partial X_i}\right|_{\mathrm{E}[\mathbf{X}]} X_i \,, \tag{10}$$

where $\mathrm{E}[\cdot]$ stands for expected value and the derivatives are evaluated for $\mathbf{x} = \mathrm{E}[\mathbf{X}]$ (this will be implicit hereafter). The second formula is very convenient to calculate the variance, having put in k all terms which do not contain X_i. Evaluating the expected values from Eq. (9), and variances and covariances from Eq. (10), we get (we have replaced the symbol '$\approx$' by '$=$' to indicate that there are no further approximations than linearization):

$$\mathrm{E}[Y_j] = Y_j(\mathrm{E}[\mathbf{X}])\,, \tag{11}$$

$$\sigma^2(Y_j) = \sum_i \left(\frac{\partial Y_j}{\partial X_i}\right)^2 \sigma_i^2 + \left\{ 2 \sum_{l<m} \left(\frac{\partial Y_j}{\partial X_l}\right) \left(\frac{\partial Y_j}{\partial X_m}\right) \rho_{lm}\, \sigma_l\, \sigma_m \right\}, \tag{12}$$

$$\mathrm{Cov}(Y_j, Y_k) = \sum_i \left(\frac{\partial Y_j}{\partial X_i}\right) \left(\frac{\partial Y_k}{\partial X_i}\right) \sigma_i^2 + \left\{ 2 \sum_{l<m} \left(\frac{\partial Y_j}{\partial X_l}\right) \left(\frac{\partial Y_k}{\partial X_m}\right) \rho_{lm}\, \sigma_l\, \sigma_m \right\}, \tag{13}$$

where σ_i are shorthand for $\sigma(X_i)$ and ρ_{lm} are the correlation coefficients, such that $\rho_{lm}\, \sigma_l\, \sigma_m = \mathrm{Cov}(X_l, X_m)$. The terms within $\{\cdot\}$ vanish if the input quantities are uncorrelated, as it often the case when relevant systematic effects are considered. However, sometimes this is not the case, as when several calibration constants are simultaneously obtained from a fit. Equations (12)–(13) can be written in the more compact form of covariance matrix transformation. However, for the purpose of this paper, we prefer not to use the matrix formalism, in order to separate the contributions due to variances and covariances.

Equations (11)–(13) give only some limited information about the joint p.d.f. of $\mathbf{Y}$, namely only 1st and 2nd moments. However the central limit theorem plays the important role of making the p.d.f. of each Y_j practically Gaussian in most of the cases of interest (see e.g. examples and words of caution in Ref. [9], and discussion in Ref. [21]). The joint p.d.f. can be considered for practical purposes a multivariate Gaussian. Anyway, in case of doubt, it is good practice to check the shape of each marginal p.d.f. (see Appendix C).

In complex real-life cases the derivatives are not performed analytically. Instead, the effects of the input values on the output values are evaluated numerically, often by Monte Carlo techniques. In these cases the derivatives can be estimated numerically by $\pm 1\,\sigma$ variations around the expected values. Calling $\Delta_{\pm_{ji}}$[8] the variation of Y_j due to a variation of X_i of $\pm 1\,\sigma_i$ around $\mathrm{E}[X_i]$, linearity implies that

$$\frac{\partial Y_j}{\partial X_i} \approx \frac{\Delta_{+_{ji}}}{\sigma_i} \approx \frac{\Delta_{-_{ji}}}{\sigma_i}. \tag{14}$$

Since in the linear approximation $\Delta_{+_{ji}}$ and $\Delta_{-_{ji}}$ are practically equal, we call Δ_{ji} either of them (taking the average of the two if there are small differences; the case of large differences, hint of non-linear effects, will be discussed below). We get, finally, the following practical formulae for the elements of the covariance matrix:

$$\begin{aligned} \sigma^2(Y_j) &= \sum_i \Delta_{ji}^2 + \left\{ 2 \sum_{l<m} \rho_{lm}\, \Delta_{jl}\, \Delta_{jm} \right\}, & (15) \\ \mathrm{Cov}(Y_j, Y_k) &= \sum_i \Delta_{ji}\, \Delta_{ki} + \left\{ 2 \sum_{l<m} \rho_{lm}\, \Delta_{jl}\, \Delta_{km} \right\}. & (16) \end{aligned}$$

In the simple case of independent input quantities, Eqs. (15)–(16) reduce to

$$\begin{aligned} \sigma^2(Y_j) &= \sum_i \Delta_{ji}^2 & (17) \\ \mathrm{Cov}(Y_j, Y_k) &= \sum_i \Delta_{ji}\, \Delta_{ki} \left[= \sum_i \mathrm{Cov}_i(Y_j, Y_k) = \sum_i s_{ijk}\, |\Delta_{ji}|\, |\Delta_{ki}| \right], & (18) \end{aligned}$$

where $\mathrm{Cov}_i(Y_j, Y_k)$ stands for the contribution to the covariance from the ith input quantity, and s_{ijk} indicate the product of the signs of the absolute increments of Y_j and Y_k for a variation of X_i ($|\Delta_{ji}|$ have the meaning of standard uncertainty of Y_j due to X_i alone).

At this point, we have to remember that μ_r defined in Sec. 2 is considered as one of the input quantities, and that in the most general case there will be many μ_{r_j}, each associated with one and only one output quantity Y_j. The resulting covariance matrix will be equal to the sum of the covariance matrix of the μ_{r_j} (they can be correlated as they could come from fitting procedures, unfolding, or other statistical techniques) and the covariance matrix due to the systematic effects. Let us write down, as an easy and practical example, the formulae for the case when we have N values μ_{r_j} and the influence

[8] We have used the following notation: $\Delta_+ = Y(\mathrm{E}[X]+\sigma_X) - Y(\mathrm{E}[X])$ and $\Delta_- = Y(\mathrm{E}[X]) - Y(\mathrm{E}[X] - \sigma_X)$. Therefore, for monotonic functions around $\mathrm{E}[X]$ the increments Δ_+ and Δ_- have the same sign.

quantities are uncorrelated:

$$\sigma^2(Y_j) = \sigma^2_{r_j} + \sum_{i>N} \Delta^2_{ji}\,, \tag{19}$$

$$\mathrm{Cov}(Y_j, Y_k) = \mathrm{Cov}(\mu_{r_j}, \mu_{r_j}) + \sum_{i>N} s_{ijk}\, |\Delta_{ji}|\, |\Delta_{ki}|\,, \tag{20}$$

where we have taken into account that the Δ_{ji} associated with μ_{r_i} are given by $\Delta_{ji} = \sigma_i\, \delta_{ij}$, where δ_{ij} is Kronecker symbol. In fact, the derivatives of Y_j with respect to μ_{r_i}, evaluated at the point of best estimate of $\mathbf{X}$, are equal to 1 if $i = j$, and equal to 0 otherwise.

4 Modelling the uncertainty due to systematic effects: ISO type-B uncertainties

At this point it is important to define somewhat better how the several ingredients appearing in the previous formulae should be evaluated. In fact, the results of the above formulae have a defined probabilistic meaning only if the various Δ's are obtained as variations of the output quantities for $1\,\sigma$ variations of the input quantities, and not, generically, as reasonable variations, or, prudentially, as 'conservative variations'. Now we are confronted with the problem that in the evaluation of uncertainties due to imperfect knowledge of systematic effects, the case in which the input uncertainties are evaluated from standard statistical procedures which provide standard deviations in an automatic way is rare. These latter cases would be those in which we feel comfortable. More often, *"for estimate x_i of an input quantity X_i that has not been obtained from repeated observations, the ...standard uncertainty ...is evaluated by scientific judgement based on all the available information on the possible variability of X_i. The pool of information may include: previous data; experience with or general knowledge of the behaviour and properties of relevant materials and instruments; manufacturer's specifications; data provided in calibration and other certificates; uncertainties assigned to reference data taken from handbooks."* (ISO Guide [19]). This is along the spirit that *"the evaluation of uncertainty is neither a routine task nor a purely mathematical one; it depends on detailed knowledge of the nature of the measurand and of the measurement"* [19].

Following the recommendation of the BIPM recommendations [20], the ISO Guide calls this kind of uncertainty type B, in contrast to type-A uncertainties obtained, to say it shortly, by the dispersion of readings (see also Section 6.1.2 of Ref. [9]). The evaluation of type-B uncertainties implies the adoption of the *"viewpoint ...that probability is a measure of the degree of belief that an event will occur"* [19]. In practice it requires a realistic and honest modelling of the case. The most common models are discussed in the ISO Guide itself: For example, if one is practically sure that an input value is in a certain interval, and all values inside the interval appear similarly likely, the proper model for the uncertainty is a uniform distribution. Other times the edges of the interval seem still to be really extreme values for the quantity; but one tends to believe more in central values, and the belief decreases roughly linearly from the centre to the edges. In this case, a more suitable distribution is a symmetric triangular distribution. Alternatively, if the belief decreases towards the edges, but the maximum belief does not coincide with (approximately) the centre of the interval, it is preferable to use an asymmetric triangular distribution. Finally, if the interval seems simply to be just a highly probable one (e.g. 90%, 95% or 99%), but also far away values are believed to be possible, one can use a Gaussian model with a standard deviation which is a suitable fraction of the uncertainty interval.

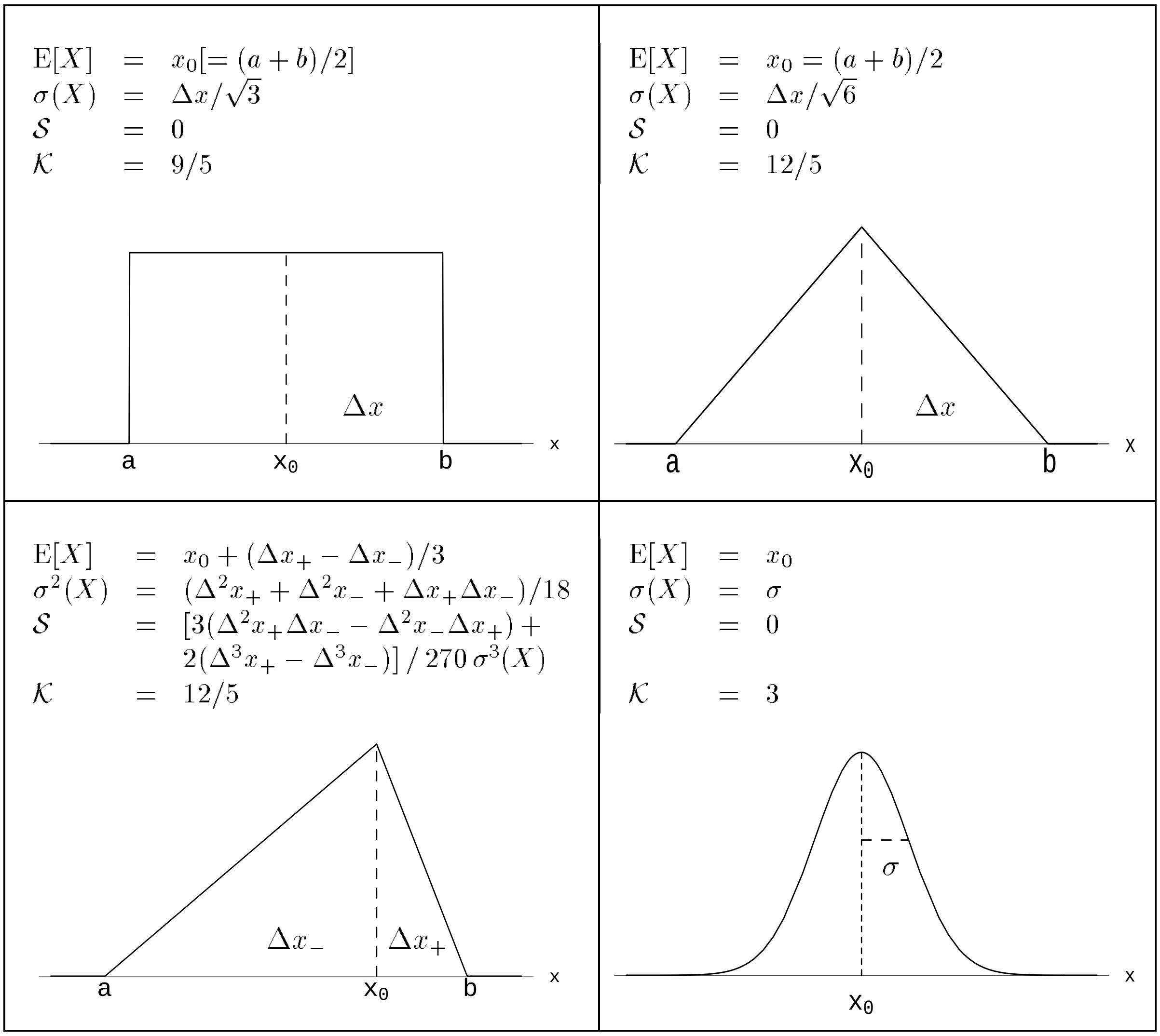

Figure 1: Typical models to assess type-B uncertainties: uniform distribution, symmetric triangular distribution, asymmetric triangular distribution, and Gaussian distribution. The expressions of the most relevant statistical parameters are reported ($\mathcal{S}$ stands for skewness, $\mathcal{K}$ for kurtosis).

Figure 1 shows the most common models to handle type-B uncertainties, together with their most interesting statistical parameters.

Once the mathematical model has been chosen, the p.d.f. of the output quantity can be evaluated using Eq. (8), or their first moments can be obtained using approximated formulae. It is important to realize that neither the choice of model, nor the value of the standard deviation are very critical, as discussed in Refs. [9] and [19] (see also Section 4.6 of Ref. [21]). In fact, the central limit theorem makes the result Gaussian independently of the initial distribution, if none of the non-Gaussian components dominates. Moreover, some unavoidable over- or under-estimates, compensate, if one makes the effort of assessing model and standard deviation in an honest way. Finally, in agreement with the ISO Guide, we think that it is bad practice to overestimate intentionally type-B uncertainties (see Section 10.2 of Ref. [9]).

As a numerical example, let us consider the standard deviations of input quantities believed to be, with certainty or with high probability, in the interval between -1 and $+1$.

$$\text{Uniform:} \quad \sigma(X) = \frac{1}{\sqrt{3}} \approx 0.58\,. \tag{21}$$

$$\text{Symmetric triangular:} \quad \sigma(X) = \frac{1}{\sqrt{6}} \approx 0.41\,. \tag{22}$$

$$\text{Asymmetric triangular peaked at } 1/2: \quad \sigma(X) = \sqrt{\frac{13}{72}} \approx 0.42\,. \tag{23}$$

$$\text{Gaussian, 90\% probability interval:} \quad \sigma(X) = \frac{1}{1.64} \approx 0.61\,. \tag{24}$$

$$\text{Gaussian, 95\% probability interval:} \quad \sigma(X) = \frac{1}{1.96} \approx 0.51\,. \tag{25}$$

We see that, for practical purposes, the differences between the σ's are irrelevant. Nevertheless, in order to avoid a bias of the overall uncertainty, one should try to model each component according to the best knowledge of the physics case, rather than by choosing systematically the model which gives the most conservative uncertainty.[9] Note that in the case of asymmetric triangular distribution, the expected value of X is neither the centre of the interval, nor the peak of the distribution. In this case we have $\mathrm{E}[X] = 1/6 \approx 0.17$. If one uses, incorrectly, the peak value, one introduces a bias which is $\approx 80\%$ of a standard deviation. As an example, Fig. 2 shows the resulting uncertainty on the quantity $Y = X_1 + X_2$, where the X_i are independent and their uncertainty is described by identical asymmetrical triangular distributions. The combined result is obtained analytically using

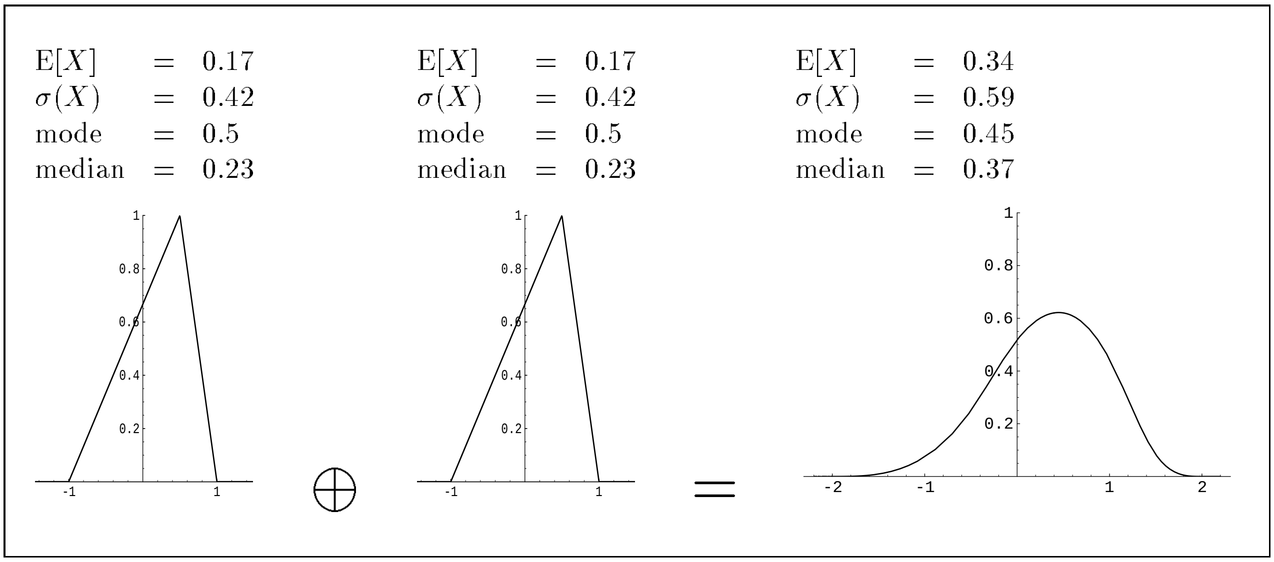

Figure 2: Probability density function resulting from the sum of two quantities, each described by an asymmetric triangular p.d.f. with $x_0 = 0.5$, $\Delta x_+ = 0.5$ and $\Delta x_- = 1.5$.

[9] The so called coherent bet, a recognised normative tool to assess probability, can help a lot to elicitate model and parameters of type-B uncertainty. For a concise introduction on its concept see Ref. [10]. An extensive discussion about its role to force people to assess uncertainty in non-standardised situations, see Ref. [12].

Eq. (8). One can see how good the Gaussian approximation already is and how biased a result can be, if the best estimate of the sum is performed using mode or median, and if the final uncertainty is evaluated with ad hoc rules of the kind shown in the introduction.

5 Small deviations from linearity

Let us consider now nonlinearity effects, which are mostly responsible for the published asymmetric uncertainties due to systematics. Nonlinearity makes in fact $\Delta_{+_{ji}}$ and $\Delta_{-_{ji}}$ differ considerably. We treat here only second-order effects. Figure 3 shows an example of the transformation of some important p.d.f.'s, all characterized by $\mathrm{E}[X] = 0$ and $\sigma(X) = 1$, while Fig. 4 shows the convolution of the original and of the transformed quantities of Fig. 3. One can see that indeed the p.d.f. of the sum of of both the original and the transformed quantities can be described by a Gaussian for the practical purposes of interest in uncertainty evaluations.

In order to simplify the formulae, let us consider first the case of only one input quantity and one output quantity (see Appendix B for the general case). Taking the second-order expansion, we have

$$Y = Y(\mathrm{E}[X]) + \frac{\partial Y}{\partial X}\,(X - \mathrm{E}[X]) + \frac{1}{2}\frac{\partial^2 Y}{\partial X^2}\,(X - \mathrm{E}[X])^2\,. \tag{26}$$

Expected value and variance of Y are then

$$\begin{aligned}
\mathrm{E}[Y] &= Y(\mathrm{E}[X]) + \frac{1}{2}\,\frac{\partial^2 Y}{\partial X^2}\,\sigma^2(X)\,, & (27)\\
\sigma^2(Y) &= \left(\frac{\partial Y}{\partial X}\right)^2 \sigma^2(X) + \frac{\partial Y}{\partial X}\,\frac{\partial^2 Y}{\partial X^2}\,\mathrm{E}\left[(X - \mathrm{E}[X])^3\right] \\
&\quad + \frac{1}{4}\left(\frac{\partial^2 Y}{\partial X^2}\right)^2 \left[\mathrm{E}[(X - \mathrm{E}[X])^4] - \sigma^4(X)\right]\,. & (28)
\end{aligned}$$

These formulae can be transformed into more practical ones if the derivatives are replaced by their numerical evaluations from the $\pm 1\,\sigma$ of X around $\mathrm{E}[X]$, which produce variations $\Delta_\pm$ in Y. The approximate derivatives evaluated in $E[X]$ are

$$\begin{aligned}
\frac{\partial Y}{\partial X} &\approx \frac{1}{2}\left(\frac{\Delta_+}{\sigma(X)} + \frac{\Delta_-}{\sigma(X)}\right) = \frac{\Delta_+ + \Delta_-}{2\,\sigma(X)}\,, & (29)\\
\frac{\partial^2 Y}{\partial X^2} &\approx \frac{1}{\sigma(X)}\left(\frac{\Delta_+}{\sigma(X)} - \frac{\Delta_-}{\sigma(X)}\right) = \frac{\Delta_+ - \Delta_-}{\sigma^2(X)}\,. & (30)
\end{aligned}$$

The formula of the variance, Eq. (28), can be simplified using skewness ($\mathcal{S}$) and kurtosis ($\mathcal{K}$), defined as $\mathcal{S} = \mathrm{E}\left[(X - \mathrm{E}[X])^3\right]/\sigma^3(X)$ and $\mathcal{K} = \mathrm{E}\left[(X - \mathrm{E}[X])^4\right]/\sigma^4(X)$, respectively. We get finally

$$\begin{aligned}
\mathrm{E}[Y] &= Y(\mathrm{E}[X]) + \delta\,, & (31)\\
\sigma^2(Y) &= \overline{\Delta}^2 + 2\,\overline{\Delta}\cdot\delta\cdot S(x) + \delta^2\cdot\left[\mathcal{K}(X) - 1\right]\,, & (32)
\end{aligned}$$

where δ is the semi-difference of the two shifts [$\delta = (\Delta_+ - \Delta_-)/2$] and $\overline{\Delta}$ is their average [$\overline{\Delta} = (\Delta_+ + \Delta_-)/2$]. The interpretation of Eq. (31) is simple and corresponds to a procedure that some might have already guessed: Asymmetric uncertainties produce a

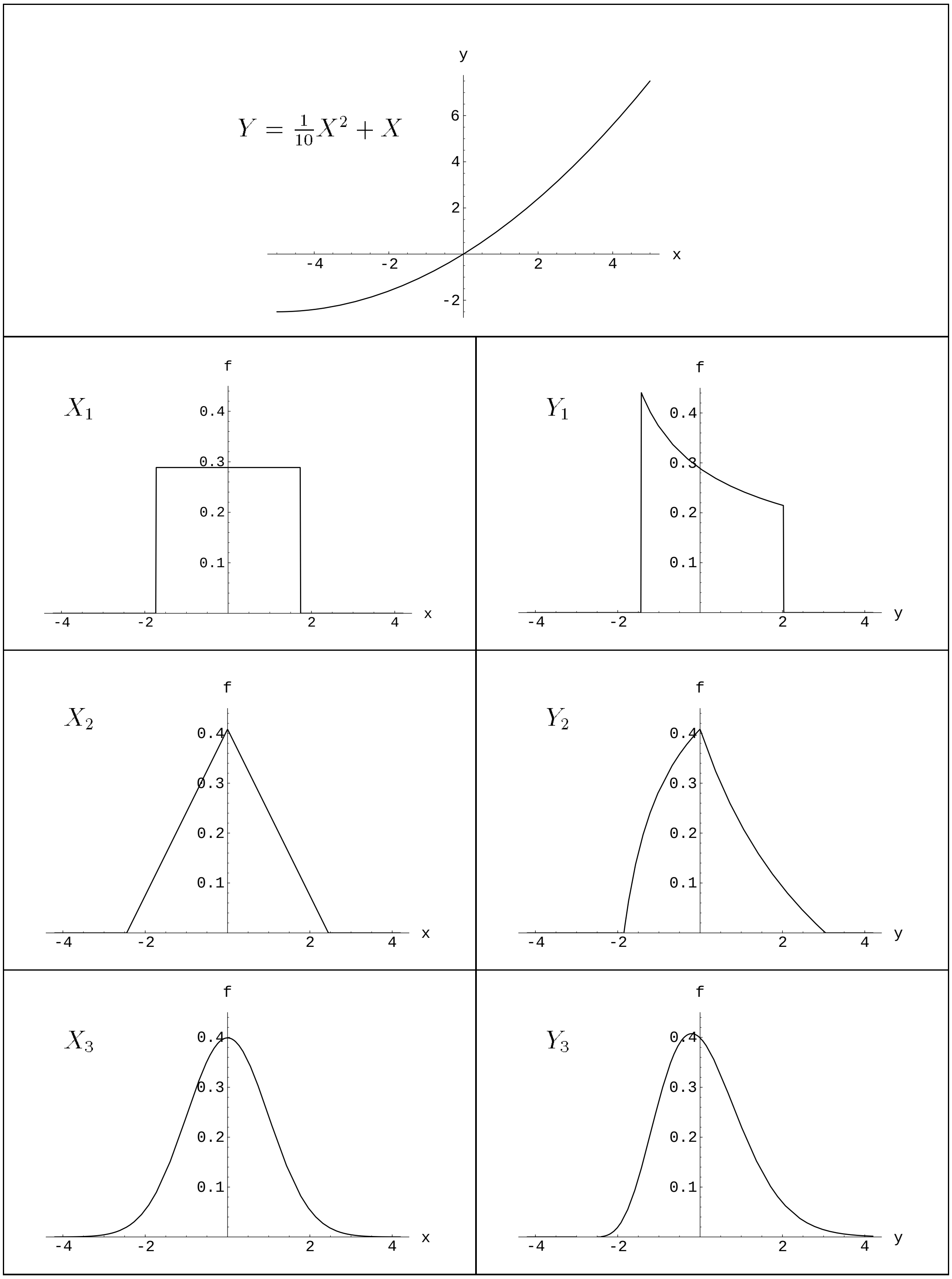

Figure 3: Propagation of a uniform, a triangular and a Gaussian distribution under a nonlinear transformation. The p.d.f.'s of X_i have been evaluated analytically using Eq. (8).

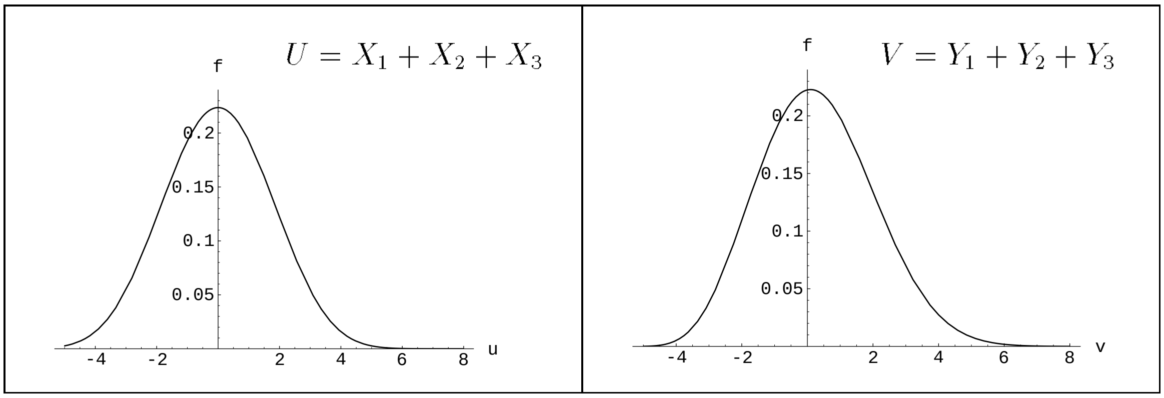

Figure 4: Probability density functions of the sum of the quantities X_i and of their nonlinear transformations Y_i defined in Fig. 3.

shift in the best estimate of the quantities. In the case that the dependence between Y and X is linear, δ is ≈ 0 and we recover the result given in Section 3. Note also that the second term of Eq. (32) disappears if the distribution describing the uncertainty on X is symmetric around $\mathrm{E}[X]$, and that the third term plays a second-order role, since the difference between Δ_+ and Δ_- is usually smaller than their sum, and $\mathcal{K}(X)$ is around 2 or 3 for the distributions of interest (see Fig. 1).

The extension to several independent input quantities is straightforward, as one only needs to add together the individual contributions to expected value and the variance. Considering the most common case in which the second and third terms of the r.h.s. of Eq. (32) are negligible[10)], we obtain the following simple practical formulae:

$$\mathrm{E}[Y] \approx Y(\mathrm{E}[\mathbf{X}]) + \sum_i \delta_i \,, \tag{33}$$

$$\sigma^2(Y) \approx \sum_i \overline{\Delta}_i^2 \,. \tag{34}$$

Averaging positive and negative deviations is indeed a good practice, but the shift of the central value should not be neglected. For the separation of input quantities into μ_{r_i} and influence factor, see Eqs. (19)–(20). The formulae for the more general case of several output quantities and of correlations among input quantities will be considered in Appendix B.

6 Numerical examples

Let us go back to the numerical example of the introduction. Those numbers were obtained from a quadratic dependence of Y on the influence quantities, each having a slightly different functional form and a different model to describe its uncertainty. Including also μ_r as X_0, we can write the dependence of Y on X_i in the following explicit form:

$$Y = \sum_{i=0}^{3} \alpha_i X_i + \beta_i X_i^2 \,, \tag{35}$$

where α_i and β_i are given in Table 1, in which also the uncertainty model is indicated. As stated in the introduction, the expression 'reasonable variation of the parameters' was

[10)] For symmetric distributions the skewness is zero, while the kurtosis is around 3 for the distributions of interest and enters with δ^2.

Table 1: Parameters of the input quantities used in the numerical example of the text. X_0 is identified with the value μ_r, obtained when X_{1-3} are equal to their expected values.

Interpretation 1: 'reasonable variations' = $\pm 1\,\sigma$ for all X_i							
Input/Output	Distribution	E[X]	$\sigma(X)$	α	β	Δ_-	Δ_+
$X_0(\equiv \mu_r)$	Gaussian	1	0.05	1	0	+0.050	+0.050
X_1	Gaussian	0	0.3	0.25	-0.167	+0.090	+0.060
X_2	Triangular $[-1,1]$	0	0.41	0.30	-0.147	+0.147	+0.098
X_3	Uniform $[-1,1]$	0	0.58	0.225	-0.078	+0.156	+0.104
Y	$\approx$ Gaussian	0.93	0.20				

Interpretation 2: 'reasonable variations' = $\pm 1\,\sigma$ for μ_r and X_1; $\pm\Delta x$ for others							
Input/Output	Distribution	E[X]	$\sigma(X)$	α	β	Δ_- (rescaled at $1\,\sigma$)	Δ_+
$X_0(\equiv \mu_r)$	Gaussian	1	0.05	1	0	+0.050	+0.050
X_1	Gaussian	0	0.3	0.25	-0.167	+0.090	+0.060
X_2	Triangular $[-1,1]$	0	0.41	0.123	-0.0245	+0.054	+0.046
X_3	Uniform $[-1,1]$	0	0.58	0.130	-0.026	+0.084	+0.066
Y	$\approx$ Gaussian	0.97	0.13				

intentionally left vague. We consider the two cases in which the variations of non-Gaussian quantities correspond to $\pm 1\,\sigma$ or to $\pm$ half-interval, respectively ('interpretation 1' and 'interpretation 2' in Table 1). The details of the first evaluation are (see Appendix B for the second case and for the values of central moments of higher order)

$$\mathrm{E}[Y] = 1.00 + \sum_i \delta_i = 1.00 + (-0.015 - 0.026 - 0.0245) = 0.9345\,, \qquad (36)$$

$$\sigma^2(Y) = \sigma_r^2(Y) + \sigma_{sys}^2(Y) = (0.05)^2 + (0.1983)^2 = (0.2046)^2\,, \qquad (37)$$

a result which can be summarized as $Y = 0.93(0.20)$, or $Y = 0.93 \pm 0.20$, although the latter expression might be misleading, since it is traditionally used as a 68% probability interval, which is exactly true only if the p.d.f. is perfectly Gaussian[11)] (see also comments in Ref [19]). The result (36)–(37) is in excellent agreement with $\mathrm{E}[Y] = 0.9344$ and $\sigma(Y) = 0.2046$ obtained directly from the p.d.f. of Y estimated by Monte Carlo with 10^6 extractions. In contrast, the result obtained combining separately positive and negative deviations in quadrature (see introduction) shows a bias which amounts to 35% of σ.

11) As stated above and shown with several figures, in most of the practical cases the Gaussian approximation is a good one, even if the models describing the uncertainty of the input quantities are not Gaussian or if there are some nonlinearity effects. In summarizing the result with a couple of numbers, our preference goes to expected value and standard deviation, because these are the parameters which matter in further propagation of uncertainty, and this is the most common use of scientific results. In this respect we agree with the recommendations of the ISO Guide [19]. In the case that the final p.d.f. differs considerably from a Gaussian, is indeed a good practice to provide also mode and median of the distribution, as well as probability intervals of interests (see e.g. Refs. [21] and [22]). However, it is clear that if, for decision problems, one wishes to assess in the most precise way probability intervals, the exact form of the p.d.f. is required. For complex problems this evaluation is performed by numerical or Monte Carlo techniques. We give in Appendix B approximate formulae of skewness and kurtosis, which give an idea of the deviation of the p.d.f. from a Gaussian. Appendix C shows also how to get an idea of the shape of the p.d.f., assuming that it is not 'too' different from a Gaussian.

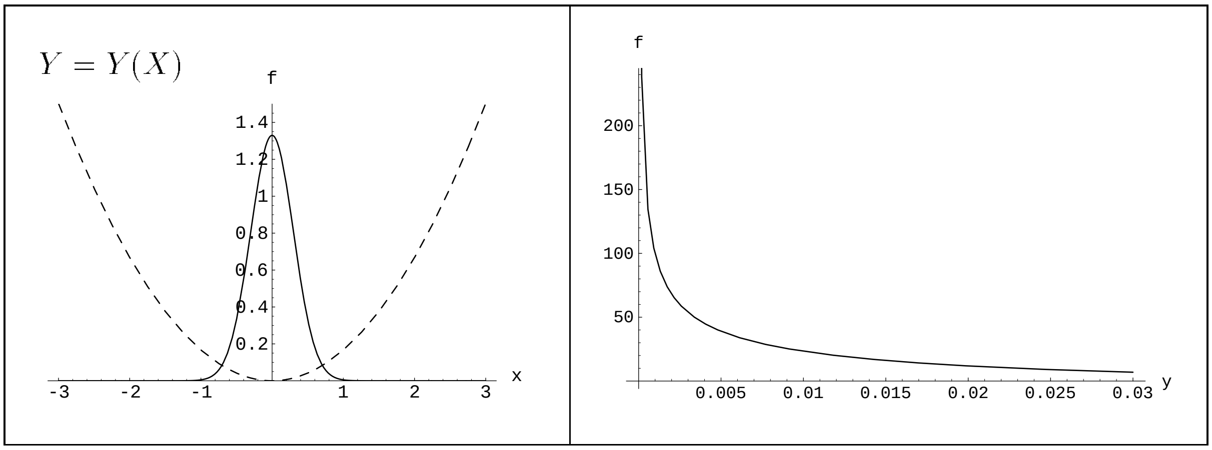

Figure 5: Example of non-monotonic relation between input and output quantity. The left plot show the parabolic dependence of Y on X (dashed line) and the Gaussian p.d.f. of X (solid line). The right plot shows the p.d.f. of Y.

7 The non-monotonic case

Sometimes a variation of $\pm 1\sigma$ of an influence parameter might produce values of Y which are both above or both below the value obtained with the reference value. Using the notation of this paper, Δ_+ and Δ_- have opposite signs in that case. This result indicates that the function is not monotonic, and this situation has to be treated with some care. In fact, although the formulae derived in this paper do not depend on whether the functions are monotonic or not, the transformed distribution can be very different from those of Fig. 3 and can bring a large non-Gaussian contribution to the overall distribution. As an example, let us consider Fig. 5, which describes an input quantity normally distributed around 0 with $\sigma = 0.3$, a parabolic dependence of Y on X given by $Y = 0.167\,X^2$ (i.e. like X_1 of Table 1, but with the $\alpha = 0$ and β reversed in sign, just for graphical convenience). The $\pm 1\,\sigma$ variations are $\Delta_+ = +0.015$ and $\Delta_- = -0.015$, but certainly one would not quote 0 as the expected value of Y, nor 0.015 its standard deviation. E$[X]$ being at the minimum of the distribution, the p.d.f. of X ends sharply at zero, and is very asymmetric. In fact it is easy to recognize in $f(y)$ a scale transformation of the χ^2 with one degree of freedom, namely $Y = 0.015 \times \chi^2_1$. Expected values and standard deviation are then $\mathrm{E}[Y] = 0.015$ and $\sigma(Y) = 0.015 \times \sqrt{2} = 0.021$. We can compare the result with what we get from Eqs. (31)–(32):

$$\left\{ \begin{array}{rcl} \overline{\Delta} & = & 0 \\ \delta & = & 0.015 \end{array} \right. \Longrightarrow \left\{ \begin{array}{rcl} \mathrm{E}[Y] & = & 0 + 0.015 = 0.015 \\ \sigma^2(Y) & = & 0 + 0 + 0.015^2 \times 2 = (0.021)^2 \end{array} \right. \tag{38}$$

The result is exactly the same, as it should be, since in this example the function is parabolic and, therefore, there are no approximations in Eqs. (31)–(32). We see that in this case only the quadratic terms appear. Similarly, it would be wrong to consider the best estimate of Y as equal 0, with an uncertainty equal to the deviation: The result would have a standard deviation smaller by $\sqrt{2}$, and the best estimate would have a bias of -140% of the reported standard deviation.

8 Conclusions

The issue of uncertainty propagation has been reviewed using the probabilistic approach nowadays called Bayesian. Following physicists' intuition, this approach allows probabilistic reasoning to be applied to not exactly-known values of physical quantities. Therefore, probabilistic statements then have a clear meaning and can be propagated to the values of related quantities using probability calculus. In particular, uncertainties due to systematic effects or theoretical inputs can be included in a conceptually easy way. Indeed, there is no qualitative distinction between uncertainty due to random and systematic (or theoretical) effects, and in this respect we agree fully with the ISO Guide [19].

Approximate formulae have been derived to calculate the first moments of output p.d.f., taking into account up to second-order effects in the uncertainty propagation. Non-linear effects are responsible, together with the less frequent case of asymmetric uncertainties on input quantities, of results given with asymmetric uncertainties, if uncertainties are evaluated using *ad hoc* procedures. We have shown that in most cases of practical interest, a combinatorial effect similar to the central limit theorem symmetrizes automatically the overall uncertainty. It follows that expected value and standard deviation of the distribution should be the correct way of presenting the result, as also pointed out by metrological authorities [19]. In the case that the final distribution is considerably non-Gaussian, other position and width indicators of the distribution (and whenever possible the p.d.f. itself) should also be provided, but expected value and standard deviation should always be given, as they are what mostly matters in further propagation of uncertainty. We have also shown that asymmetric uncertainty on input quantities and/or non-linear effects produces shifts in the expected value of the output quantity with respect to the value calculated for the best estimate of the input quantities. *If these shifts are not applied, the result can be biased.*

As far as modelling the uncertainty due to systematic effects is concerned, we have found particularly helpful the concept of type-B uncertainty, according to the ISO terminology [19]. We have shown that the results are rather stable against reasonable variations of models and values of the parameters. On the other hand, sticking to the position that there is no way of modelling uncertainty due to systematic effects would lead to impossibility of providing an overall uncertainty, or to providing this uncertainty using arbitrary prescriptions which do not have a clear meaning. In other words, we prefer beliefs assessed by people we trust, rather than empty prescriptions, in accordance with the scheme: beliefs in, beliefs out; nothing in, nothing out (see Ref. [10] for a discussion about 'belief' and 'arbitrariness'). This position is based on the assumption that *"it is scientific only to say what is more likely and what it is less likely"* [25].

We would like to conclude with some remarks about when to stop making systematic checks and adding contributions to the overall uncertainty. As eloquently said [26], *"one could correlate the result with the phase of the Moon or the position of Jupiter, and find most likely no significant effect, with some uncertainty; but certainly we don't want to take care of this uncertainty."* Only contributions which are in principle relevant should be considered in the uncertainty evaluation. Even if the effect is 'statistically significant', one should try to understand if it can physically influence the result, before including it in the analysis. There is certainly some subjectivity in deciding what is relevant and what is not, but this is consistent with the spirit that *"the quality and the utility of the uncertainty quoted for the result of a measurement therefore ultimately depend on the understanding, critical analysis, and integrity of those who contribute to the assignment of its value"* [19].

Appendix A – Derivation of Eq. (8) and comparison with Eq. (3)

Equation (8) is one of the most convenient ways to formulate the problem of p.d.f. propagation, especially for physicists, who are familiar with the Dirac delta. Better known formulae which make use of the Jacobian are easily recovered by making use the properties of the delta. The same is true for the convolution formula to obtain the p.d.f. of the sum of two independent quantities. In this latter case, the derivation is straightforward: $f(y) = \int f(x_1)\cdot f(x_2)\cdot \delta(y-x_1-x_2)\,\mathrm{d}x_1\mathrm{d}x_2 = \int f(x_1)\cdot f(y-x_1)\,\mathrm{d}x_1$. Equation (8) is also important because it justifies the Monte Carlo estimation of the p.d.f. and clarifies the role of the various ingredients which enter the game (see comments in footnote 7).

Although we think that Eq. (8) does not need to be proved, as it can be considered itself the formulation of the problem, we give here a formal derivation based on the properties of the characteristic function. The characteristic function associated with X is defined as (see e.g. [23]):

$$\phi_X(t) \equiv \mathrm{E}\left[e^{itX}\right] = \int e^{itx} f(x)\,\mathrm{d}x\,, \tag{39}$$

from which the p.d.f. can be reobtained as

$$f(x) = \frac{1}{2\pi}\int e^{-itx}\phi_X(t)\,\mathrm{d}t. \tag{40}$$

If we have a function $Y = g(X)$, the p.d.f. of Y can be obtained from the following characteristic function [23]:

$$\phi_Y(t) = \int e^{itg(x)} f(x)\,\mathrm{d}x\,. \tag{41}$$

This property can be extended to a variable Y depending on many variables, i.e.

$$\phi_Y(t) = \int e^{itg(\mathbf{x})} f(\mathbf{x})\,\mathrm{d}\mathbf{x}\,. \tag{42}$$

It follows

$$f(y) = \frac{1}{2\pi}\int e^{-ity}\phi_Y(t)\,\mathrm{d}t = \int e^{-ity}\,\mathrm{d}t\int e^{itg(x)} f(\mathbf{x})\,\mathrm{d}\mathbf{x} = \int e^{-it(y-g(x))} f(\mathbf{x})\,\mathrm{d}\mathbf{x}\mathrm{d}t. \tag{43}$$

Noting that $\int e^{it(x-y)}\,\mathrm{d}t = 2\pi\,\delta(x-y)$ we get finally

$$f(y) = \int f(\mathbf{x})\,\delta(y-g(\mathbf{x}))\,\mathrm{d}\mathbf{x}\,, \tag{44}$$

which is equivalent to Eq. (8).

Once we have proved Eq. (8), there is no need to prove also that this is equivalent to Eq. (3), as the latter comes from a general theorem of probability theory. We show their equivalence in a simple case, using the example of the offset uncertainty given in Section 2. Using the notation of Eq. (8), the correspondence of symbols is $Y \equiv \mu$, $h_1 \equiv z$ (with $z_\circ = 0$) and "data" $\equiv x$. The p.d.f.'s of our input quantities are

$$f(\mu_r\,|\,x)\,[\equiv f(\mu\,|\,x, z_\circ)] \;=\; \frac{1}{\sqrt{2\,\pi}\,\sigma_r}\,e^{-\frac{(\mu_r-x)^2}{2\,\sigma_r^2}}\,, \tag{45}$$

$$f(z) \;=\; \frac{1}{\sqrt{2\,\pi}\,\sigma_z}\,e^{-\frac{z^2}{2\,\sigma_z^2}}\,, \tag{46}$$

while the function that relates output to input quantities is $\mu = \mu_r + z$. Applying Eq. (8), and writing explicitly the integration limits, we have

$$
\begin{aligned}
f(\mu \,|\, x) &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\,\pi}\,\sigma_r}\, e^{-\frac{(\mu_r - x)^2}{2\,\sigma_r^2}}\, \frac{1}{\sqrt{2\,\pi}\,\sigma_z}\, e^{-\frac{z^2}{2\,\sigma_z^2}}\, \delta(\mu - \mu_r - z)\, \mathrm{d}\mu_r \mathrm{d}z \qquad (47)\\
&= \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\,\pi}\,\sigma_r}\, e^{-\frac{(\mu - z - x)^2}{2\,\sigma_r^2}}\, \frac{1}{\sqrt{2\,\pi}\,\sigma_z}\, e^{-\frac{z^2}{2\,\sigma_z^2}}\, \mathrm{d}z\,, \qquad (48)
\end{aligned}
$$

which is exactly Eq. (5).

Appendix B – Useful formulae up to second-order approximation

In the general case of several input quantities and several output quantities the second-order expansion is given by

$$
\begin{aligned}
Y_j &= Y_j(\mathrm{E}[\mathbf{X}]) + \sum_i \frac{\partial Y_j}{\partial X_i}\,(X_i - \mathrm{E}[X_i]) \\
&\quad + \frac{1}{2}\sum_{l,m} \frac{\partial^2 Y_j}{\partial X_l \partial X_m}\,(X_l - \mathrm{E}[X_l])\,(X_m - \mathrm{E}[X_m])\,. \qquad (49)
\end{aligned}
$$

Expected value, variances and covariances of Y_j are then

$$
\begin{aligned}
\mathrm{E}[Y_j] &= Y_j(\mathrm{E}[\mathbf{X}]) + \frac{1}{2}\sum_{l,m} \frac{\partial^2 Y_j}{\partial X_l \partial X_m}\,\mathrm{Cov}(X_l, X_m)\,, \qquad (50)\\
\sigma^2(Y_j) &= \sum_{l,m} \left(\frac{\partial Y_j}{\partial X_l}\right)\left(\frac{\partial Y_j}{\partial X_m}\right)\mathrm{Cov}(X_l, X_m) \\
&\quad + \sum_{l,m,n} \frac{\partial Y_j}{\partial X_l}\,\frac{\partial^2 Y_j}{\partial X_m \partial X_n}\,\mathrm{E}[\widetilde{X_l}\widetilde{X_m}\widetilde{X_n}] \\
&\quad + \frac{1}{4}\sum_{h,l,m,n} \left(\frac{\partial^2 Y_i}{\partial X_h \partial X_l}\right)\left(\frac{\partial^2 Y_i}{\partial X_m \partial X_n}\right) \\
&\quad \times \left(\mathrm{E}[\widetilde{X_h}\widetilde{X_l}\widetilde{X_m}\widetilde{X_n}] - \mathrm{Cov}(X_h, X_l)\,\mathrm{Cov}(X_m, X_n)\right) \qquad (51)\\
\mathrm{Cov}(Y_j, Y_k) &= \sum_{l,m} \frac{\partial Y_j}{\partial X_l}\frac{\partial Y_k}{\partial X_m}\,\mathrm{Cov}(X_l, X_m) \\
&\quad + \frac{1}{2}\sum_{l,m,n} \left(\frac{\partial Y_k}{\partial X_l}\frac{\partial^2 Y_j}{\partial X_m \partial X_n} + \frac{\partial Y_j}{\partial X_l}\frac{\partial^2 Y_k}{\partial X_m \partial X_n}\right)\mathrm{E}\left[\widetilde{X_l}\widetilde{X_m}\widetilde{X_n}\right] \\
&\quad + \frac{1}{4}\sum_{h,l,m,n} \left(\frac{\partial^2 Y_j}{\partial X_l \partial X_h}\right)\left(\frac{\partial^2 Y_k}{\partial X_m \partial X_n}\right) \\
&\quad \times \left(\mathrm{E}\left[\widetilde{X_h}\widetilde{X_l}\widetilde{X_m}\widetilde{X_n}\right] - \mathrm{Cov}(X_h, X_l)\,\mathrm{Cov}(X_m, X_n)\right), \qquad (52)
\end{aligned}
$$

where $\widetilde{X_i} = X_i - [X_i]$. The expressions involving expected values of products of more than two $\widetilde{X_i}$'s can be evaluated by iterating the procedure described here, as they are functions of X_i. An interesting case for the applications is, using the same notation as at the end of Section 3, that in which we have N input quantities of the kind μ_{r_i}, and all the remaining ones have the meaning of uncorrelated influence quantities. Using the derivatives evaluated numerically and the definition of skewness and kurtosis, we can write the previous formulae as

$$\begin{aligned}
\sigma^2(Y_j) &= \sigma^2_{r_j} + \sum_{i>N} \overline{\Delta}^2_{ji} \\
&\quad + 2\sum_{i>N} \overline{\Delta}_{ji} \cdot \delta_{ji} \cdot S(X_i) + \sum_{i>N} \delta^2_{ji} \cdot [\mathcal{K}(X_i) - 1] \,, \qquad (53)\\
\mathrm{Cov}(Y_j, Y_k) &= \mathrm{Cov}(\mu_{r_j}, \mu_{r_k}) + \sum_{i>N} \overline{\Delta}_{ji} \overline{\Delta}_{ki} \\
&\quad + \sum_{i>N} (\overline{\Delta}_{ji}\, \delta_{ki} + \overline{\Delta}_{ki}\, \delta_{ji})\, \mathcal{S}(X_i) + \sum_{i>N} \delta_{ji}\delta_{ki}\, [\mathcal{K}(X_i) - 1] \,, \qquad (54)
\end{aligned}$$

where $\overline{\Delta}_{ji} = (\Delta_{+ji} + \Delta_{-ji})/2$ and $\delta_{ji} = (\Delta_{+ji} - \Delta_{-ji})/2$ and $\Delta_{\pm ji}$ are defined according to the convention of footnote 8. As explained in the text, the terms depending on skewness and kurtosis are negligible in most practical cases.

It might be useful to have also the approximate expressions of skewness and kurtosis of the output quantities Y_j, to have a rough idea of how much their p.d.f.'s differ from the Gaussian. Since the formulae become quite awful, we only consider the most common case of uncorrelated input quantities.

$$\begin{aligned}
\mathcal{S}(Y_j) &= \Bigg[3\sum_i \delta_{ji}\, \overline{\Delta}^2_{ji} \cdot [\mathcal{K}(X_i) - 1] + \sum_i \delta^3_{ji} \cdot [\mathcal{E}(X_i) - 3\,\mathcal{K}(X_i) + 2] \\
&\quad + \sum_i \overline{\Delta}^3_{ji}\, \mathcal{S}(X_i) + 6\sum_i \delta^2_{ji}\, \overline{\Delta}^2_{ji} \cdot [\mathcal{P}(X_i) - 2\,\mathcal{S}(X_i)] \Bigg] \Big/ \sigma^3(Y_j) \,, \qquad (55)\\
\mathcal{K}(Y_j) &= \Bigg[\sum_i \mathcal{K}(X_i)\, \overline{\Delta}^4_{ji} + 6\sum_{l>m} \overline{\Delta}^2_{jl}\, \overline{\Delta}^2_{jm} + 6\sum_{l\neq m} \overline{\Delta}^2_{jl}\, \delta^2_{jm} \cdot [\mathcal{K}(X_m) - 1] \\
&\quad + 6\sum_i \overline{\Delta}^2_{ji}\, \delta^2_{ji} \cdot [\mathcal{E}(X_i) - 2\,\mathcal{K}(X_i) + 1] + 6\sum_{l>m} \delta^2_{jl}\, \delta^2_{jm} \cdot [\mathcal{K}(X_l) - 1] \cdot [\mathcal{K}(X_m) - 1] \\
&\quad + \sum_i \delta^4_i \cdot [\mathcal{O}(X_i) + 6\,\mathcal{K}(X_i) - 4\,\mathcal{E}(X_i) - 3] + 4\sum_i \delta_{ji}\, \overline{\Delta}^3_{ji} \cdot [\mathcal{P}(X_i) - \mathcal{S}(X_i)] \\
&\quad + 4\sum_i \overline{\Delta}_{ji}\, \delta^3_{ji} \cdot [\mathcal{H}(X_i) - 3\,\mathcal{P}(X_i) + 3\,\mathcal{S}(X_i)] \\
&\quad + 24\sum_{l>m} \overline{\Delta}_{jl}\, \overline{\Delta}_{jm}\, \delta_{jl}\, \delta_{jm}\, \mathcal{S}(X_l)\, \mathcal{S}(X_m) \Bigg] \Big/ \sigma^4(Y_j) \,, \qquad (56)
\end{aligned}$$

where

$$\begin{aligned}
\mathcal{P}(X) &= \mathrm{E}\left[(X - \mathrm{E}[X])^5\right] / \sigma^5 \,, \qquad & \mathcal{E}(X) &= \mathrm{E}\left[(X - \mathrm{E}[X])^6\right] / \sigma^6 \,, \\
\mathcal{H}(X) &= \mathrm{E}\left[(X - \mathrm{E}[X])^7\right] / \sigma^7 \,, \qquad & \mathcal{O}(X) &= \mathrm{E}\left[(X - \mathrm{E}[X])^8\right] / \sigma^8
\end{aligned}$$

are higher order scaled central moments, analogues of skewness and kurtosis. Table 2 gives these higher order moments for most relevant p.d.f.'s used to assess type-B uncertainty.

Table 2: The rescaled central moments of fifth to eighth order for the distribution of Fig. 1. The parameters α, β and γ are given in Eqs. (57)–(59).

Distribution	$\mathcal{P}(X)$	$\mathcal{E}(X)$	$\mathcal{H}(X)$	$\mathcal{O}(X)$
Gaussian	0	15	0	105
Uniform	0	27/7	0	9
Triangular	0	54/7	0	144/5
Asymmetric triangular	β/σ^5	$2/7 \times (31 - 27\,\alpha)$	γ/σ^7	$16/5 \times (13 - 27\,\alpha)$

The moments of the asymmetric triangular distribution are given in terms of the following parameters:

$$
\begin{aligned}
\alpha &= \Delta x_-^2\,\Delta x_+^2(\Delta x_- + \Delta_+)^2 / \, (\Delta x_-^2 + \Delta x_- \Delta_+ + \Delta x_+^2)^3\,, && (57)\\
\beta &= -(\Delta x_-^5 + 5/2\,\Delta x_-^4\,\Delta x_+ + \Delta x_-^3\,\Delta x_+^2 - && (58)\\
&\qquad \Delta x_-^2\,\Delta x_+^3 - 5/2\,\Delta_-\,\Delta_+^4 - \Delta_-^5\,)/425.2\,, \\
\gamma &= -(2\,\Delta x_-^7 + 7\,\Delta x_-^6\,\Delta x_+ + 9\,\Delta x_-^5\,\Delta x_+^2 + 5\,\Delta x_-^4\,\Delta x_+^3 - && (59)\\
&\qquad 5\,\Delta x_-^3\,\Delta x_+^4 - 9\,\Delta x_-^2\,\Delta x_+^5 - 7\,\Delta_-\,\Delta_+^6 - 2\,\Delta_-^7\,)/2915.0\,.
\end{aligned}
$$

As a numerical example, continuing on from Section 6, we compare the results obtained with the approximate formulae with those obtained calculating the moments directly from the p.d.f. estimated by Monte Carlo. The overall results are given in Table 3. The agreement is well above that needed for practical purposes.

Table 3: Comparison between the moments evaluated using approximate formulae and those obtained by Monte Carlo, based on the examples of Table 1.

	'Interpretation 1'		'Interpretation 2'	
	Monte Carlo	Approx. formulae	Monte Carlo	Approx. formulae
E[Y]	0.9344	0.9345	0.9722	0.9720
σ(Y)	0.2046	0.2046	0.1297	0.1295
$\mathcal{S}$(Y)	−0.370	−0.372	−0.318	−0.321
$\mathcal{K}$(Y)	2.857	2.859	3.076	3.082

Appendix C – Approximate evaluation of the p.d.f. from the first four central moments

We have seen that the propagation of uncertainty can be solved either exactly, at the cost of having to compute complicated multidimensional integrals, or by using approximate formulae which only give the first moments of the distribution. The probabilistic interpretation of the latter case is based on the assumption that the final p.d.f. is approximately Gaussian, as is often so in most cases of practical interest. We would like to show here an intermediate solution to the problem, which allows the shape of the p.d.f. of each Y_i to be estimated starting from the approximate evaluation of the first four central moments (see Appendix B), and solving only one integral in only one variable.

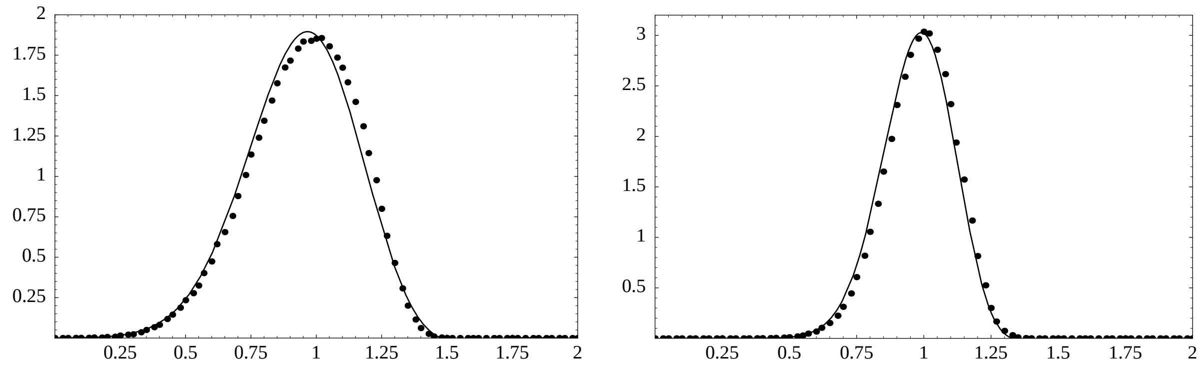

Figure 6: The distribution given by the complete uncertainty propagation (dots) is compared with the approximate probability density function for interpretations 1 (left side) and 2 (right side).

Take the generic output variable Y. Its characteristic function $\phi_Y(t)$ can be expressed as

$$\phi_Y(t) = \exp\left(\sum_{n=1}^{\infty} \frac{\chi_n}{n!}(i\,t)^n\right)\,, \tag{60}$$

where χ_n is the nth 'semi-invariant' [24]. Semi-invariants can be expressed in terms of central moments, the first four being

$$\begin{aligned} \chi_1 &= \mathrm{E}(X)\,, \\ \chi_2 &= \sigma^2\,, \\ \chi_3 &= \sigma^3\,\mathcal{S}(X)\,, \\ \chi_4 &= \sigma^4\,[\mathcal{K}(X)-3]\,. \end{aligned}$$

An approximate evaluation of the p.d.f. can be obtained under the assumption that the first four moments, for which we have obtained approximate expressions, are the most relevant ones to characterize the p.d.f.. This implies that the p.d.f.'s of interest do not differ 'too much' from a Gaussian. In practice this means that they are unimodal, have no discontinuity, and that the probability mass is concentrated around a few standard deviations from the expected value. Assuming, then, that all semi-invariants χ_n with $n > 4$ are the same as for a Gaussian distribution, namely zero, the characteristic function becomes

$$\phi_Y(t) = \exp\left(i\,t\,\mathrm{E}[Y] - \frac{1}{2}\,t^2\,\sigma^2(Y) - \frac{i}{6}\,t^3\,\sigma^3\,\mathcal{S}(Y) + \frac{1}{24}\,t^4\,\sigma^4\,[\mathcal{K}(Y)-3]\right)\,. \tag{61}$$

Inserting Eq. (61) in the anti-transformation formula [see Eq. (43)] and taking only the real part of the solution, since the p.d.f. is real, we get

$$f(y) = \frac{1}{2\pi}\int \exp\left(-\frac{1}{2}\sigma^2 t^2 + \frac{1}{24}\,t^4\sigma^4\,[\mathcal{K}(Y)-3]\right)\,\cos\left(t\,[\mathrm{E}[Y]-y] - \frac{1}{6}t^3\,\sigma^3\,\mathcal{S}(Y)\right)\,\mathrm{d}t\,. \tag{62}$$

Applied to our examples of Section 6 and using the values of Table 3, we get the p.d.f.'s drawn with the continuous curves of Fig. 6 [the integral (62) has been solved with *Mathematica*]. For comparison the dotted curves show the estimations of p.d.f.'s obtained by

Monte Carlo. The agreement as it can been judged by eye is excellent. However, since the method is approximate, there are slight problems of normalization and positiveness. But these problems affect the tails of the distribution and are not really relevant if one is interested in having an idea of the shape under the assumptions of the approximation. Note also the divergent term in Eq. (62) for $\mathcal{K} > 3$. But in practical cases the kurtosis is never much larger than this value. In fact one starts usually from distributions which have $\mathcal{K}(X) \leq 3$ (see Fig. 1) and, thanks to the central limit theorem, there is a natural tendency to have $\mathcal{K}(Y) \approx 3$. Therefore, in case of values of kurtosis slightly larger that 3, a good approximation is to limit it at 3. This approximation has been indeed applied to the 'interpretation 2' of example of Table 3, and the resulting p.d.f. is still in excellent agreement with the Monte Carlo evaluation (see right hand plot of Fig. 6).

References

[1] G. D'Agostini, *Bayesian reasoning versus conventional statistics in high energy physics*, Proc. XVIII International Workshop on Maximum Entropy and Bayesian Methods, Garching, Germany, July 1998, eds. V. Dose, W. von der Linden, R. Fischer, and R. Preuss, (Kluwer Academic Publishers, Dordrecht, 1999), pp. 157–170, `physics/9811046`.

[2] G. D'Agostini, *On the use of the covariance matrix to fit correlated data*, Nucl. Instrum. and Methods in Phys. Res. **A346** (1994) 306.

[3] D. Seibert, *Undesirable effects on covariance matrix techniques for error analysis*, Phys. Rev. **D49** (1994) 6240.

[4] S. Chiba and D.L. Smith, *Impact of data transformations on least-square solutions and their significance in data analysis and evaluation*, J. Nucl. Sc. Tech. **31** (1994) 770, and references therein.

[5] Particle Data Group, C. Caso et al., *Review of particle physics*, Eur. Phys. J. **C3** (1998) 1 (`http://pdg.lbl.gov`).

[6] C. Pascaud and F. Zomer, *QCD analysis from the proton structure function F_2 measurement: issues on fitting, statistical and systematic errors*, LAL 95-05, June 1995.

[7] S. Alekhin, *Extraction of parton distributions and α_s from DIS data within the Bayesian treatment of systematic errors*, Eur. Phys. J. **C10** (1999) 395.

[8] M. Botje, *A QCD analysis of HERA and fixed target structure function data* , ZEUS Note 98-062 DESY-99-038, December 1999, `hep-ph/9912439`.

[9] G. D'Agostini, *Probabilistic reasoning in HEP - principles and applications*, Report CERN 99–03, July 1999, also available at the author's URL.

[10] G. D'Agostini, *Teaching statistics in the physics curriculum. Clarifying and unifying role of subjective probability*, Am. J. Phys. **67** (1999) 1260.

[11] P. Astone and G. D'Agostini, *Inferring the intensity of Poisson processes at the limit of detector sensitivity (with a case study on gravitational wave burst search)*, CERN–EP/99–126, August 1999, `hep-ex/9909047`.

[12] G. D'Agostini, *Overcoming priors anxiety*, to be published in the monographic issue of the Revista de la Real Academia de Ciencias on Bayesian Methods in the Sciences, ed. J.M. Bernardo, `physics/9906048`, June 1999.

[13] G. D'Agostini, *Confidence limits : what is the problem? Is there the solution?*, Contribution to the Workshop on "Confidence Limits", CERN 17–18/1/2000, `hep-ex/0002055`

[14] G. D'Agostini, *Quantum mechanics and interpretation of probability (with comments on confidence intervals)*, Contribution to the Workshop on "Confidence Limits", CERN 17–18/1/2000, paper in preparation.

[15] S. Bertolini, J.O. Eeg, M. Fabbrichesi and E.I. Lashin, *ϵ'/ϵ at the $O(P^4)$ in the chiral expansion*, *Nucl. Phys.* **B514** (1998) 93;
A.J. Buras, M. Gorbahn, S. Jäger, M. Jamin, M.E. Lautenbacher and L. Silvestrini, *Standard Model confronting new results for ϵ'/ϵ*, TUM-HEP-347/99, `hep-ph/9904408`;
T. Hambye, G.O. Kohler, E.A. Paschos and P.H. Soldanet, *Analysis of ϵ'/ϵ in the $1/N_c$ expansion*, DO–TH–99–10, LNF–99/016(P), June 1999, `hep-ph/9906434`;
A.A. Bel'kov, G. Bohm, A.V. Lanyov and A.A. Moshkin, *Phenomenological analysis of ϵ'/ϵ within an effective chiral lagrangian approach at $O(p^6)$*, JINR E2–99–236, July 1999, `hep-ph/9907335`.

[16] M. Ciuchini, E. Franco, L. Giusti, V. Lubicz and G. Martinelli, *Combined analysis of the unitarity triangle and CP violation in the Standard Model*, BUHEP–99–24, RM3–TH/99–9 and ROME 99/1267, October 1999, `hep-ph/9910236`.

[17] F. Parodi, P. Roudeau and A. Stocchi, *Constraints on the parameters of the CKM matrix by End 1998*, LAL 99-03, March 1999, `hep-ex/9903063`

[18] S. Plaszczynski and M.-H. Schune, *Overall determination of the CKM matrix*, LAL 99-03, November 1999, `hep-ph/9911280`.

[19] International Organization for Standardization (ISO), *Guide to the expression of uncertainty in measurement*, Geneva, Switzerland, 1993.

[20] R. Kaarls, *BIPM proc.-Verb. Com. Int. Poids et mesures* **49** (1981), A1-A2; P. Giacomo, Metrologia **17** (1981) 73 (draft of English version; for the official BIPM translation see Ref. [19]).

[21] G. D'Agostini and G. Degrassi, *Constraints on the Higgs boson mass from direct searches and precision measurements*, Eur. Phys. J. **C10** (1999) 663.

[22] G. D'Agostini, *Sceptical combination of experimental results: general considerations and applications to ϵ'/ϵ*, CERN–EP/99-139, October 1999, `hep-ex/9910036`.

[23] M.G. Kendal and A. Stuart, *The advanced theory of statistics*, 1943 (C. Griffin & Co., 1969).

[24] H. Cramer *Mathematical Methods of Statistics.*, Princeton University Press, 1946.

[25] R. Feynman, *The character of the physical law*, 1965 (The MIT press, 1967).

[26] H. Wahl, private communication.