Generalized Lambda Distribution for the Expression of Measurement Uncertainty

Domenico Alessandro Lampasi, *Member, IEEE*, Fabio Di Nicola, *Student Member, IEEE*, and Luca Podestà

***Abstract*—The generalized lambda distribution family fits the probability distributions of a wide variety of data sets, including the most important distributions encountered in the measurement applications (normal, uniform, Student's t, U-shaped, exponential). This paper illustrates how the four parameters needed for such distribution can be exploited in the expression of measurement uncertainty and to extend the information related to a measurement. The obtained representation allows an immediate calculation of coverage intervals and is particularly useful to support the techniques commonly applied in the estimation of the combined uncertainty. Moreover, in order to include the classical measurement information, a novel parameterization of the distribution is proposed.**

***Index Terms*—Calibration, digital signal processing, measurement uncertainty, Monte Carlo method, probability, virtual instruments.**

I. Introduction

THE GUIDE to the Expression of Uncertainty in Measurement (GUM) [1] was an effective effort toward the standardization of the methods for the evaluation and the representation of measurement uncertainty. Although the importance of the GUM is universally recognized, limitations and improvements were clearly individuated by many authors [2], [3]. The main concerns are related to nonlinear and nonconventional algorithms [4], [5], as, for example, the virtual instruments [6]–[8]. Generally, effective solutions are provided by the numerical techniques based on Monte Carlo simulation (MCS), now supposed to be included in a GUM supplement [9]. However, even if proper tools for the uncertainty analysis are available, they risk being ignored or underestimated in the industrial practice. Therefore, their application should be practical, automated, and generalized independently of the measurement complexity. For example, these characteristics would be useful in the modeling of a phenomenon by means of a probability density function (pdf) and in the random generation according to such pdf, as required by the MCS.

On the other hand, the concept itself of uncertainty $u(x)$, associated with an estimation $\widehat{x}$, as standard deviation $\sigma(x)$ could be reductive in many practical situations, without further information about the manner in which the values are distributed. Measurements may help the decision process only if they provide suitable intervals of confidence $U_p(x)$, or, in a more recent vision, the coverage intervals $I_p(x)$ [9], corresponding to the specified levels of confidence p. In this sense, $U_p(x)$ or $I_p(x)$ may be considered as the most important information for a metrologist [1]. The coverage factors k_p [numbers of $\sigma(x)$ or $u(x)$ needed to obtain $U_p(x)$] strongly depend on the actual sample distribution and are inapplicable for asymmetric pdfs. The exact calculation of $U_p(x)$ is possible also without introducing the coverage factors [10], but the input pdfs must be known. Since the expression of the uncertainty, just by indicating $u(x)$ or $U_p(x)$ for a particular p, is incomplete, the uncertainty should ideally be expressed by the pdf and not only by single parameters related to its $\sigma(x)$: Only pdfs, through expanded uncertainty $U_p(x)$ for every p, allow aware decisions. It implies the implementation of methods, as the MCS, to propagate the distributions and not only the uncertainties [11].

We will show that the model nowadays used to represent the uncertainty exploits two (estimated mean and standard deviation) or, in most refined cases, five parameters or information (for example, including also p or k_p and the two extremes of the interval determined by p). It usually forces us to assume some constraints on the measurand: For example, that the pdf is normal or, at least, symmetric. The aim of this paper is to demonstrate that a four-parameter representation is able to completely characterize the results of a measurement, including the information contained in the classical parameters, which is sufficient in the simplest cases. These new parameters are based on the generalized lambda distribution (GλD) [12]. Their introduction is practical and suitable for automatic and software procedures, as required by the industrial standards. Besides, they allow to directly derive $U_p(x)$ and $I_p(x)$. Next sections present the GλD theory and techniques, along with practical algorithms to estimate the GλD parameters and to exploit them in the uncertainty evaluation and expression. Finally, Sections IV and V report some examples concerning the use of the proposed algorithms in the simple measurement applications. Both type A and type B uncertainties will be considered, even though the latter normally gives the main contribution.

II. GλD

In this paper, the expanded uncertainty $U_p(x) = k_p u(x)$will be considered a continue function of p. For asymmetrical distributions, the representation $\widehat{x} \pm U_p(x)$ should be replaced by introducing the endpoints x_{low} and x_{high} of $I_p(x)$ [9]:

$$I_p(x) \triangleq [x_{\mathrm{low}}, x_{\mathrm{high}}] = [\widehat{x} - U_{\mathrm{low}}, \widehat{x} + U_{\mathrm{high}}] \qquad (1)$$

where different endpoints were considered for $U_p(x)$ too. From this point of view, $I_{\overline{p}}(x)$ for a given $\overline{p}$ coincides with the

Manuscript received June 15, 2005; revised April 20, 2006.

The authors are with the Department of Electrical Engineering, University of Rome "La Sapienza." Via Eudossiana, 18-00184 Rome, Italy (e-mail: alessandro.lampasi@uniroma1.it).

Digital Object Identifier 10.1109/TIM.2006.876408

0018-9456/$20.00 © 2006 IEEE

percentile function $Q(p)$ evaluated in $\overline{p}$ and $1-\overline{p}$, which is commonly considered as the inverse of the cumulative distribution function (cdf) $G(x)=\int_{-\infty}^{x} g(\xi)d\xi$, because it gives the value of x, such that $G(x)=p$ [13].

The individuation of the pdf fitting experimental data is a basic statistical operation. The parametric estimation consists in the estimation of a limited number of parameters: Assuming that the data were produced by a particular family of distributions, such parameters are sufficient to establish the pdf. For example, the mean μ and σ completely characterize a normal distribution. Very used statistical parameters are the moments $\boldsymbol{m}(x)=[m_1(x),m_2(x),m_3(x),m_4(x)]$. The first two moments coincide with $\mu(x)$ and $\sigma^2(x)$; $\widehat{m}_3(x)$ and $\widehat{m}_4(x)$ are called the skewness and kurtosis, respectively. For a given data set $x_1,x_2,\ldots,x_N$, the sample moments are

$$\widehat{m}_1(x)=\widehat{\mu}(x)=\widehat{x}=\sum_{n=1}^{N}\frac{x_n}{N} \tag{2}$$

$$\widehat{m}_2(x)=\widehat{\sigma}^2(x)=s^2(x)=\sum_{n=1}^{N}\frac{(x_n-\widehat{x})^2}{N-1} \tag{3}$$

$$\widehat{m}_i(x)=\sum_{n=1}^{N}\frac{(x_n-\widehat{x})^i}{N\widehat{\sigma}^{i(x)}} \qquad i>2. \tag{4}$$

The nonparametric estimation tries to determine the pdf without restricting the analysis to a particular family and to include as many distributions as possible. The most used method relates the samples to the coefficients of the sum of particular functions, for instance, the orthogonal ones or, usually, the histograms:

$$h(x)=\frac{1}{N}\sum_{n=1}^{N}h_\Delta(x-x_n). \tag{5}$$

In normal applications, $h_\Delta(x)$ is a Δ-width rectangle but tends to the Dirac function $\delta(\cdot)$ for $N\to\infty$.

In this paper, an intermediate approach is considered: It is based on the classification of all the interesting distributions in a general family. For example, the Pearson curves include all the pdfs satisfying a differential equation [14]. The one-parameter lambda family, suggested in 1947 and developed by Tukey in 1962, approximates several common symmetric distributions [15]. The four-parameter GλD introduced by Ramberg and Schmeiser (RS-parameterization) in 1972 is described by the percentile [12]

$$Q_{\lambda_1,\lambda_2,\lambda_3,\lambda_4}(p)\triangleq\lambda_1+\frac{p^{\lambda_3}-(1-p)^{\lambda_4}}{\lambda_2}, \qquad 0\le p\le 1 \tag{6}$$

where λ_1, λ_2, λ_3, and λ_4 are the λ-parameters of the distribution. Its pdf can be derived from [12]

$$g_{\lambda_1,\lambda_2,\lambda_3,\lambda_4}(x)=\frac{\lambda_2}{\lambda_3 p^{\lambda_3-1}+\lambda_4(1-p)^{\lambda_4-1}}, \qquad x=Q(p). \tag{7}$$

The numerical integration of (7) allows to find $G(x)$.

TABLE I
REGIONS WHERE $g_{\lambda_1,\lambda_2,\lambda_3,\lambda_4}(x)>0$. THEIR CHARACTERISTICS ARE DETERMINED BY THE VALUES OF λ_3 AND λ_4. TAILS EXIST ONLY FOR NEGATIVE λ_3 AND λ_4

λ_3	λ_4	Interval with $g(x)>0$
$\lambda_3>0$	$\lambda_4>0$	$\lambda_1\pm\frac{1}{\lambda_2}$
$\lambda_3>0$	$\lambda_4=0$	$[\lambda_1,\lambda_1+\frac{1}{\lambda_2}]$
$\lambda_3=0$	$\lambda_4>0$	$[\lambda_1-\frac{1}{\lambda_2},\lambda_1]$
$\lambda_3<0$	$\lambda_4<0$	$(-\infty,\infty)$
$\lambda_3<0$	$\lambda_4=0$	$(-\infty,\lambda_1+\frac{1}{\lambda_2}]$
$\lambda_3=0$	$\lambda_4<0$	$[\lambda_1-\frac{1}{\lambda_2},\infty)$

TABLE II
BASIC GλD APPROXIMATIONS OF TYPICAL DISTRIBUTIONS

Distribution	λ_1	λ_2	λ_3	λ_4
Normal ($\mu=0,\sigma=1$)	0	0.2	0.1	λ_3
Uniform in $[a,b]$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$	1	λ_3
Student's t ($\nu=10$)	0	0.02	0.01	λ_2
U–shaped	μ	1	0.5	λ_3
Exponential ($\mu=1$)	0.01	-10^{-3}	-10^{-6}	λ_2
Cauchy ($\mu=0$)	0	1	-1	λ_3
Logistic ($\mu=0,\sigma=1$)	0	-10^{-4}	λ_2	λ_2
Weibull ($\alpha=1,\beta=5$)	1	1	0.2	0.1

The parameters λ_1 and λ_2 determine the location and scale of the distribution:

$$Q_{\overline{\lambda_1},\overline{\lambda_2},\lambda_3,\lambda_4}(p)=\overline{\lambda_1}+\overline{\lambda_2}Q_{0,1,\lambda_3,\lambda_4}(p) \tag{8}$$

where $Q_{0,1,\lambda_3,\lambda_4}(p)$ is the standardized GλD [12]. The parameters λ_3 and λ_4 capture the shape characteristics; in fact, they univocally fix m_3 and m_4. In particular, as summarized in Table I, λ_3 and λ_4 establish the points where $g(x)>0$ [12] and, therefore, the measurand admissible values: For example, tails may exist only if λ_3 or λ_4 are negative. The pdf $g_{\lambda_1,\lambda_2,\lambda_3,\lambda_4}(x)$ is the symmetric image, about the line $x=\lambda_1$, of $g_{\lambda_1,\lambda_2,\lambda_4,\lambda_3}(x)$; this means that a symmetric distribution has $\lambda_3=\lambda_4$.

As briefly shown in Table II, the GλD fits the distributions to a wide variety of data sets; especially, it appears to cover the typical measurement families (normal, uniform, U-shaped, Student's t) [7]–[9]. The quality of the approximation is different for each distribution: For example, the uniform pdf is exactly reproduced, whereas for the normal pdf $g_N(x)$ [12]

$$\sqrt{\int_{-\infty}^{\infty}\left(g_{\lambda_1,\lambda_2,\lambda_3,\lambda_4}(x)-g_N(x)\right)^2dx}\simeq 10^{-3}. \tag{9}$$

The simplicity of $Q(p)$ given by the RS-parameterization has two disadvantages: A limitation in the admissible λ-parameters and a nonimmediate relationship between them and effective moments. In fact, a vector $\boldsymbol{\lambda} = [\lambda_1, \lambda_2, \lambda_3, \lambda_4]$ gives the valid GλDs, only if it respects the conditions $G(\infty) = 1$ and $g(x) \geq 0$. The former is always verified and the latter can be verified through (7). Practically, since λ_1 and λ_2 are unrestricted, the search of the λ-parameters should be limited to valid regions of the (λ_3, λ_4) plane [12]. It is not a peculiarity of the GλD: Actually, every distribution imposes its own constraints (for example, normal distribution implies $\sigma > 0$). However, other parameterizations with different characteristics were introduced in literature [13]. The most successful

$$Q_{\lambda_1,\lambda_2,\lambda_3,\lambda_4}(p) \triangleq \lambda_1 + \frac{\frac{p^{\lambda_3}-1}{\lambda_3} - \frac{(1-p)^{\lambda_4}-1}{\lambda_4}}{\lambda_2} \tag{10}$$

which is known as the FMKL-parameterization [16], was proposed by Freimer *et al.* in 1988 [17] and is well defined over the entire $(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$ domain.

The moments of a GλD can be analytically and univocally derived from the λ-parameters through a vector of functions $\boldsymbol{T}(\cdot) = [T_1(\cdot), T_2(\cdot), T_3(\cdot), T_4(\cdot)]$:

$$\begin{aligned} m_1 &= T_1(\boldsymbol{\lambda}) = \lambda_1 + \frac{A_1(\lambda_3,\lambda_4)}{\lambda_2} \\ m_2 &= T_2(\boldsymbol{\lambda}) = \frac{A_2(\lambda_3,\lambda_4) - A_1^2(\lambda_3,\lambda_4)}{\lambda_2^2} \\ m_{3,4} &= T_{3,4}(\lambda_3,\lambda_4) \end{aligned} \tag{11}$$

where $A_{1,2}(\lambda_3, \lambda_4)$ and $T_{3,4}(\lambda_3, \lambda_4)$ represent functions in λ_3 and λ_4, whose expressions can be found in [12], even if they are differently defined. In particular, our interest is limited to

$$A_1(\lambda_3,\lambda_4) = \frac{\lambda_4 - \lambda_3}{(1+\lambda_3)(1+\lambda_4)} \tag{12}$$

$$A_2(\lambda_3,\lambda_4) = \frac{\lambda_3 + \lambda_4 + 1}{(0.5+\lambda_3)(0.5+\lambda_4)} - 2\int_0^1 z^{\lambda_3}(1-z)^{\lambda_4}dz \tag{13}$$

where the integral corresponds to the beta function [18] evaluated in $(\lambda_3 + 1, \lambda_4 + 1)$. Since in this paper, the four λ-parameters are used to integrate the information provided by $\mu(x)$ and $\sigma(x)$, we propose to transform them in two λ-parameters. It is accomplished by the new $\mu\sigma$-parameterization:

$$Q_{\lambda_\mu,\lambda_\sigma,\lambda_3,\lambda_4}(p) \triangleq \lambda_\mu + \lambda_\sigma \left(\frac{p^{\lambda_3} - (1-p)^{\lambda_4} - A_1(\lambda_3,\lambda_4)}{\sqrt{A_2(\lambda_3,\lambda_4) - A_1^2(\lambda_3,\lambda_4)}} \right). \tag{14}$$

Its usefulness in measurement is given by the properties:

$$\mu(x) = m_1(x) = \lambda_\mu, \qquad \sigma(x) = \sqrt{m_2(x)} = |\lambda_\sigma|. \tag{15}$$

Obviously, $\mu(x)$ and $\sigma(x)$ could be computed from the RS-parameters through (11), but the expression of the uncertainty and the devices' specifications would be less legible. Due to the direct derivation from the RS-parameterization, $\mu\sigma$-parameterization can exploit the information and the tables available for the originating family simply by adapting the first two parameters:

$$\begin{aligned} \lambda_\mu &= \lambda_1 + \frac{A_1(\lambda_3,\lambda_4)}{\lambda_2} \\ \lambda_\sigma &= \frac{\sqrt{A_2(\lambda_3,\lambda_4) - {A_1}^2(\lambda_3,\lambda_4)}}{\lambda_2}. \end{aligned} \tag{16}$$

III. From Experimental Data to λ-Parameters

The theory of the methods employed to estimate the λ-parameters from the experimental data can be found in texts of the statistics, especially in [12]. Here, we introduce some practical considerations to understand their implementation.

The method of moments (MM) is the initial and still the most popular approach for estimating the λ-parameters, which consists of matching the first four empirical moments (2)–(4) to the λ-moments (11) by solving the equation system:

$$\widehat{m}_i(x) = m_i(x), \qquad i = 1, 2, 3, 4. \tag{17}$$

Some other methods are based on a moment matching: the least-square method [19], "starship" method [20], and controlled-randomization method [16]. These methods are restricted to the regions where moments exist, even if there are the GλD members without the first four moments providing the perfect fits to the data. Due to the limitation on (λ_3, λ_4) values, solutions may not exist for some (m_3, m_4) couples. The impossible region can be reduced by using the auxiliary generalized beta distribution [21].

Recently, a method based on the approximation of the percentile was proposed (percentile method, PM) [22]. We verified the effectiveness of the PM in several measurement applications. In particular, the computational complexity is lower, and the accuracy is greater than the MM. The PM defines four new sample statistics based on the empirical percentiles $\widehat{Q}(p)$ and on a reference point p_0:

$$\begin{aligned} \widehat{\rho}_1 &= \widehat{Q}(0.5) \\ \widehat{\rho}_2 &= \widehat{Q}(1-p_0) - \widehat{Q}(p_0) \\ \widehat{\rho}_3 &= \frac{\widehat{Q}(0.5) - \widehat{Q}(p_0)}{\widehat{Q}(1-p_0) - \widehat{Q}(0.5)} \\ \widehat{\rho}_4 &= \frac{\widehat{Q}(0.75) - \widehat{Q}(0.25)}{\widehat{\rho}_2}. \end{aligned} \tag{18}$$

Such statistics are arbitrary, because there are other possible choices, but they individuate relevant characteristics of the distributions: $\widehat{\rho}_1$ is the sample median, $\widehat{\rho}_2$ is the range between the 10% and the 90% of the values, $\widehat{\rho}_3$ is a left–right tail ratio,

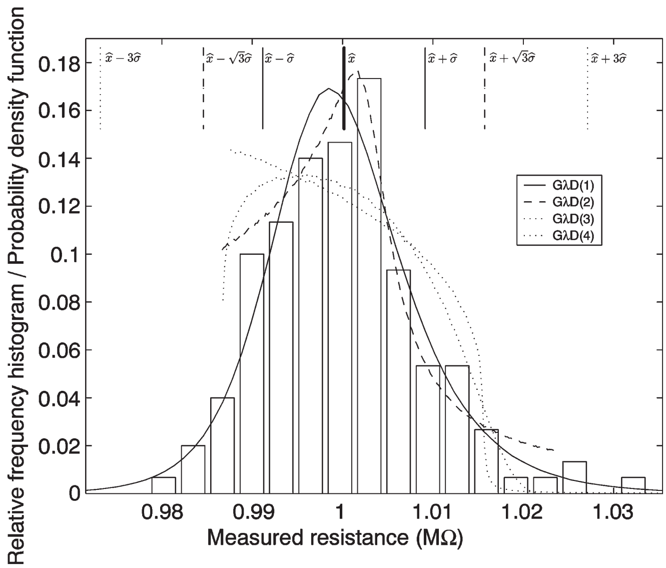

Fig. 1. Relative frequency histogram of the measured resistance R_1 with four possible GλD-based pdf fits, indicated as GλD(i). The experimental statistics are $\widehat{x} = 1.0002$ MΩ and $\widehat{\sigma}(x) = 0.0090$ MΩ, and the median is 0.9995 MΩ. The upper vertical lines represent $\widehat{x}$ and the intervals $\pm\widehat{\sigma}(x)$, $\pm 3\widehat{\sigma}(x)$, and $\pm\sqrt{3}\widehat{\sigma}(x)$.

and $\widehat{\rho}_4$ is an index of the tail weight. The same logic is behind the common choice $0 < p_0 < 0.25$; as a compromise, we used $p_0 = 0.1$ [22]. Accordingly, the new system to be solved is

$$\widehat{\rho}_i = \rho_i, \qquad i = 1, 2, 3, 4 \tag{19}$$

with ρ_i directly defined by (6).

The solutions for the systems (17) and (19) are not available in closed form. Anyway, the numerical solution is pretty straightforward. Since the subsystems for (m_3, m_4) and (ρ_3, ρ_4) involve only λ_3 and λ_4, it is better to first solve these subsystems and then to obtain λ_2 and λ_1. Specific tables for the solution of (17) and (19) (or for an initial guess for numerical algorithms) may be found in literature.

It is important to note that, for the same set of m_i and ρ_i and thus of $\widehat{m}_i$ and $\widehat{\rho}_i$ $(i = 1, 2, 3, 4)$, the different λ-parameters (namely the different distributions) may exist. It is absolutely correct because, otherwise, the moments or other four statistics would be sufficient to characterize all the distributions. Since most of the results obtained for the λ-parameters (generally three distinct curves for the PM) are very different from the actual pdf, a simple observation of the graphs is normally sufficient to establish the best fit. Alternatively, an immediate goodness-of-fit numerical test as (9) can be implemented.

IV. Expression of Uncertainty

As a simple application of the GλD approach, we measured a statistically meaningful set of resistors by means of a 7.5-digit DMM Keithley 2001. Fig. 1 reports the relative frequency histogram of the measured resistance (the nominal value is $R_1 = 1$ MΩ), assuming that the repeatability uncertainty is negligible. To express the distribution of these values, we can indicate $\widehat{x}$, along with a symmetric interval: $\pm\widehat{\sigma}$ or an expanded uncertainty $U_{p\to 1} = \pm 3\widehat{\sigma}$ (under the assumption of a normal pdf) or $U_{p\to 1} = \pm\sqrt{3}\widehat{\sigma}$ (uniform). These parameters are depicted through the upper vertical lines in Fig. 1. Due to the nonsymmetry of the histogram, the confidence interval could be specified as $\widehat{x} - U_{\text{low}} \le x \le \widehat{x} + U_{\text{high}}$, thus using three parameters. Even when the data comply with a well-known distribution (as normal or uniform), it has to be indicated as a further information. Alternatively, the effective U_p could be estimated directly from the experimental data: In this case, $\sigma(x)$ and $u(x)$, which are also necessary to combine the uncertainties, are derivable by hypothesizing the pdf; moreover, p should be indicated along with the expanded uncertainty. In fact, a meaningful example of reporting the measurement results is quoted in [9]:

$$x = \overline{x}, \qquad u(x) = \overline{u}(x), \qquad I_{\overline{p}}(x) = [\overline{x}_{\text{low}}, \overline{x}_{\text{high}}]. \tag{20}$$

It means five parameters (estimated mean $\overline{x}$, standard uncertainty $\overline{u}(x)$, chosen level of confidence $\overline{p}$, relative extremes $\overline{x}_{\text{low}}$ and $\overline{x}_{\text{high}}$). The obtained specification is valid only for the particular $\overline{p}$, and more than one $I_{\overline{p}}$ may exist (even if the shortest one has to be considered) [9]. The same considerations should be taken into account for the resistor specifications reported by the manufacturer (actually, the specified accuracy was larger than the effective one).

Clearly, the experimental data may be fitted by means of a GλD, as shown in Fig. 1. Using the PM, there are four possible solutions indicated as the GλD(i): Two pdfs are quite different, the GλD(2) appears to well fit just the central values. The GλD(1) includes the whole histogram $h(x)$ defined in (5), covering also two reasonable tails and, therefore, was chosen as the best fit. The RS-parameters of the GλD(1) are $\lambda_1 = 0.9983$, $\lambda_2 = -7.6602$, $\lambda_3 = -0.0304$, $\lambda_4 = -0.0430$, whereas the $\mu\sigma$-parameters are $\lambda_\mu = 1.0002$, $\lambda_\sigma = -0.0091$, $\lambda_3 = -0.0304$, and $\lambda_4 = -0.0430$. The resulting shape looks like a normal pdf, but it is slightly asymmetric. This example shows how the experimental distribution is completely expressed by means of four parameters. The amount of data is extremely smaller than required by histograms or by the approximation presented in [9]. From the same four parameters, it is immediate to derive $I_p(x)$ or $U_p(x)$ for every p.

It is interesting to stress that the GλD is able to fit the data, even when they are incompletely given, typically as histograms. An extreme but meaningful example is summarized in Fig. 2: The histogram is reproduced directly from the current transducer gain plot published in [4]. One fit, marked as GλD(1), represents a sound and easy way to report the gain measurement and to accomplish the procedures described in that paper, using only four parameters.

V. Propagation of Uncertainty

The possibility of storing the complete pdf may remarkably improve the analysis of the combined uncertainty. According to the studies in [9], four approaches can be used for this purpose: the well-known law of propagation of uncertainty (LPU), the numerical methods (specifically, the MCS), the LPU including higher order terms, and the analytical methods.

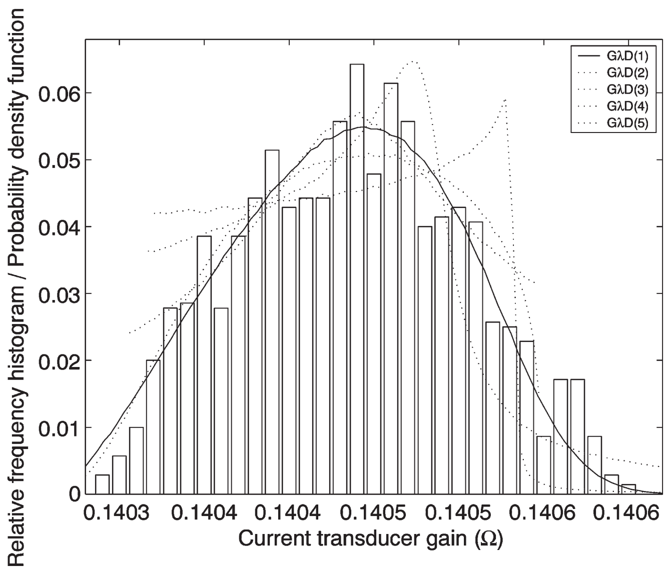

Fig. 2. GλD approach applied to a histogram reproduced directly from literature and representing a nonconventional pdf. The fit is obtained using only the histogram values and not the actual measurement data. $G\lambda D(1)$ represents the best fit to data.

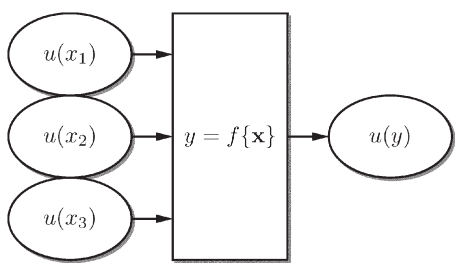

Fig. 3. LPU block diagram: Due to the limited information provided by the inputs, the combined uncertainty may be erroneously evaluated, and the output pdf is completely unknown.

The LPU formula combines the contributions of several quantities $\mathbf{x} = [x_1, x_2, \ldots, x_n]$ to a measurand $y = f\{\mathbf{x}\}$ [1]:

$$\begin{aligned} u_c^2(y) &= m_2(y) \\ &= \sum_{i=1}^{N} \left(\frac{\partial f}{\partial x_i}\right)^2 m_2(x_i) + 2 \sum_{i=1}^{N-1} \sum_{\substack{j=1 \\ j\neq i}}^{N} \frac{\partial f}{\partial x_i}\frac{\partial f}{\partial x_j} \\ &\quad \cdot r(x_i, x_j)\sqrt{m_2(x_i)m_2(x_j)}. \end{aligned} \tag{21}$$

Since λ-parameters contain $m_2(x)$, explicitly as in (15) for $\mu\sigma$-parameterization or implicitly as in (11) for RS-parameterization, (21) is still applicable. Actually, such approach does not provide information about the propagation of distributions, as illustrated in Fig. 3. Due to the difficult estimation of the matrix $r(x_i, x_j)$, sources are usually considered uncorrelated, or simplified approaches are adopted [7]. Furthermore, the LPU applies well under four basic conditions [2], [9]. In particular, the nonlinearity of $f\{\mathbf{x}\}$ has to be insignificant, and the Central Limit Theorem has to apply. Otherwise, a paradoxical result could be produced [8].

The most popular approach to overcome the LPU limitations and to propagate distributions is based on the MCS. Once the uncertainty sources are classified in generalized and predefined groups [8] by modeling their effect on the measurand, random values must be generated according to the pdfs assigned to these sources. The principles of the operation are shown in Fig. 4.

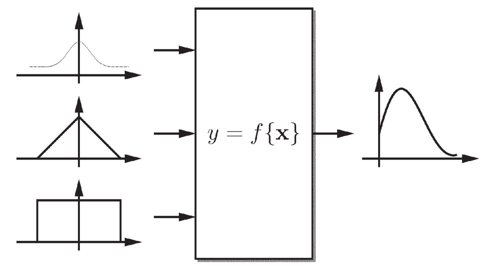

Fig. 4. MCS for propagation of the distributions. Once the measurement is modeled, the estimation of the output pdf is possible, provided that the input pdfs are known and proper pdf-like random generators are available.

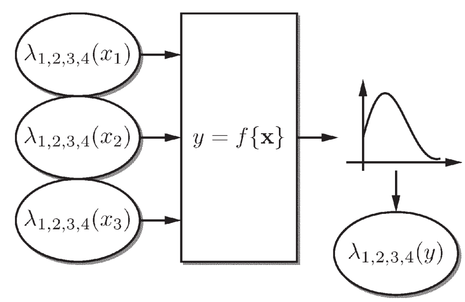

Fig. 5. Using the GλD, propagation of the distributions is accomplished, independently of the method adopted to combine uncertainties, by representing each uncertainty source in a general way through the four parameters. The same representation may be used for the output measurement.

Normally, very few standard pdfs are used as inputs, even due to the lack of manufacturer information. The ideal way to perform a sampling from a selected pdf consists in reducing the problem to a pseudorandom number generation from a uniform distribution, supported by several tools. In fact, if $\mathcal{U}_{a,b}$ is a variable uniformly distributed in the interval $[a, b]$, then $Q(\mathcal{U}_{0,1})$ has the pdf of $Q(p)$. Unluckily, $Q(p)$ is not available for many important distributions, even with a well-known and simple pdf (software routines and libraries are available for some distributions, such as normal and Student's). It is evident that the GλD allows easy generation of random variables from every kind of distribution, because featuring an explicit and accessible $Q(p)$ reduces it to a uniform generation in [0,1]. The quality of the obtained pdfs and the relative statistical properties [6] were experienced to be generally adequate, for example applying (9). Moreover, independently of the distribution, a vector of (uniformly generated) samples, even previously recorded, can be used. Accordingly, as represented in Fig. 5, the GλD-based MCS consists in providing the four λ-parameters for each uncertainty source. Finally, also the output pdf can be expressed through its λ-parameters.

Convolution of distributions (CD) is the basic analytical method to propagate distributions:

$$g(y) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} g(\mathbf{x})\delta\left(y - f\{\mathbf{x}\}\right) dx_N \ldots dx_1. \tag{22}$$

The CD does not introduce any approximation but can be applied just in simple cases [1], [2]. Otherwise, the numerical calculation of (22) is possible using (7). However, it is not suggested, because the results eventually coincide with the MCS.

According to (11), the four moments of the variables $\mathbf{x}$ and y can be derived from their respective λ-parameters $\lambda_{\mathbf{x}}$ and λ_y through the systems:

$$\mathbf{m}(\mathbf{x}) = \mathbf{T}(\lambda_{\mathbf{x}}), \qquad \mathbf{m}(y) = \mathbf{T}(\lambda_y). \tag{23}$$

Anyway, it is impossible to write general formulas for the relationships among $\mathbf{m}(\mathbf{x})$ and $\mathbf{m}(y)$. Analogously to (21), based on a first-order Taylor expansion and valid for the propagation of m_2, the other moments can be as well related by approximating $f\{\mathbf{x}\}$ using a Taylor series arrested to the Kth order:

$$\mathbf{m}(y) \simeq \mathbf{f}^{(K)}\{\mathbf{m}(\mathbf{x})\}. \tag{24}$$

Such system may be extended to any order and resumes the "propagation of moments." For example, $m_{1,2}(y) = f_{1,2}^{(K)}\{\mathbf{m}(\mathbf{x})\}$, as proposed in [23]. A measurement-oriented simplification of (24) for additive models can be found in [24].

The comparison between (23) and (24) suggests the possibility to implement a "propagation of λ-parameters":

$$\lambda_y = \mathbf{T}^{-1}\left(\mathbf{f}^{(K)}\{\mathbf{T}(\lambda_{\mathbf{x}})\}\right). \tag{25}$$

Actually, as we have already observed, (11) cannot be inverted in closed form and usually provides more than one set of solutions. The same circumstances were verified for (25), along with a further inconvenience: It is not easy to carry out a goodness-of-fit test, unless using the pdf obtained by the MCS, namely introducing another method. However, the propagation of the λ-parameters appears to be a relevant concern for future works.

As a special case of (24), (21) can be extended to the third-order terms. For a single input quantity $y = f\{x\}$ [2], independently of the input pdf

$$m_1(y) = f\{m_1\} + \frac{1}{2}f''\{m_1\}m_2 \tag{26}$$

$$\begin{aligned} m_2(y) = {} & f'^2\{m_1\}m_2 + f'\{m_1\}f'''\{m_1\}m_3 \\ & + \frac{1}{4}f''^2\{m_1\}\left(m_4 - m_2^2\right) + \frac{1}{3}f'\{m_1\}f'''\{m_1\}m_4 \end{aligned} \tag{27}$$

where the second-member moments refer to the variable x. However, (26) and (27) can be simplified only through the knowledge of the moments [1], [2], as actually allowed by (11) and (15).

A. Example: Uncertainty of a Voltage Divider

The preceding considerations were applied in a simple case, but useful in several measurement and circuital applications: the pdf of the voltage divider obtained by the series of R_1 and another set of resistors $R_2 = 10$ kΩ, analyzed in Fig. 6 analogously to Fig. 1. The nominal attenuation $a_v = 101$ (gain $g_v = 0.0099$) is suitable for measurements on the typical ac and dc power supplies. Under the hypothesis that $u(R_1)$ and $u(R_2)$ are not correlated, the standard uncertainty of a_v evaluated using the LPU is $u_c(a_v) \simeq 2.5$, even if such hypothesis strictly does not apply, because the resistors operate in the same conditions and were calibrated by the same instruments (in fact, in a voltage divider, the effects of the influence quantities tend to compensate each other).

Fig. 7 presents the pdf and the statistical parameters of a_v, computed through the MCS and using only the λ-parameters of

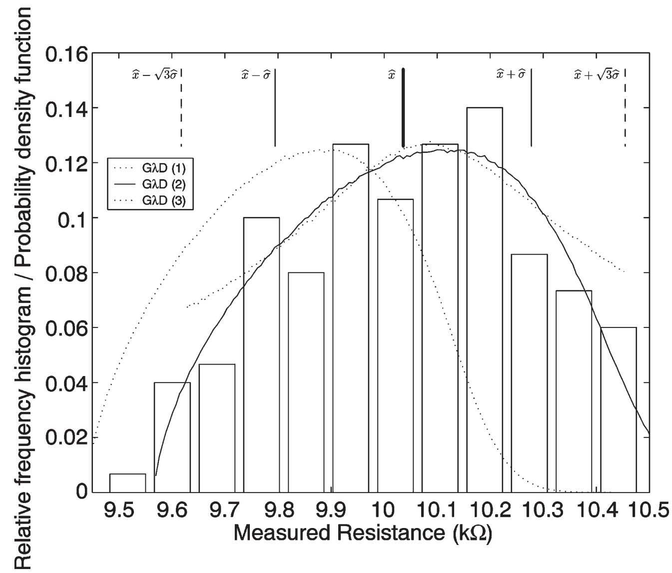

Fig. 6. Relative frequency histogram of the measured resistance R_2 with three GλD fits, indicated as GλD(i). The experimental statistics are $\widehat{x} = 10.0361$ kΩ and $\widehat{\sigma}(x) = 0.2415$ kΩ, and the median is 10.0612 kΩ. The upper vertical lines represent $\widehat{x}$ and the intervals $\pm\widehat{\sigma}(x)$, $\pm 3\widehat{\sigma}(x)$, and $\pm\sqrt{3}\widehat{\sigma}(x)$.

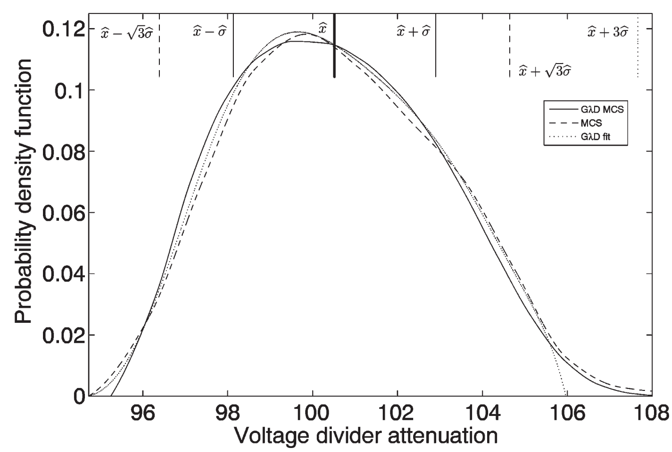

Fig. 7. Comparison between two different evaluations of a_v carried out by the MCS: The solid line was obtained using only the λ-parameters of R_1 and R_2, whereas the dashed line was derived by considering the entire distributions. A GλD fit of a_v obtained at the output of the GλD MCS is shown as a dotted line in the same graph. The statistics of such distribution are $\widehat{x} = 100.5124$ and $\widehat{\sigma}(x) = 2.3816$, and the median is 100.3890. The resulting pdf is clearly asymmetric.

R_1 and R_2. It coincides with the MCS pdf (and to the CD pdf) derived by considering the entire distributions, thus requiring a large amount of data and two random generators based on two different nonstandard pdfs. The dotted line indicates a GλD fit for a_v.

VI. Conclusion

This paper investigated the properties of the GλD and the methods to exploit it to fit the experimental data, even when such data are incomplete. The proposed examples illustrated the capability of this distribution to improve the representation of the measurement results, uncertainties, and coverage intervals. The specifications of the four λ-parameters for a measurement, a device, a process, or a calibration are sufficient to completely individuate their statistical characteristics. The evaluation of the combined uncertainty is simple and effective, using any approach based on the propagation of the distributions, without compromising the classical analysis.

It is interesting to note that, due to the possibility to obtain general, comprehensive, and simulation-ready indications, the λ-parameters may be reported in the standard documentation, but currently, they can be included in the transducer electronic data sheets (TEDS) [25] and then read directly by the measurement devices and used by tools for automated analysis of the uncertainty.

References

[1] BIPM, IEC, IFCC, ISO, IUPAC, IUPAP, and OIML, *Guide to the Expression of Uncertainty in Measurement*, 1995.

[2] M. G. Cox and P. M. Harris, *Software Support for Metrology Best Practice Guide no. 6—Uncertainty Evaluation*. Middlesex, U.K.: NPL, 2004.

[3] A. Ferrero and S. Salicone, "The random-fuzzy variables: A new approach to the expression of uncertainty in measurement," *IEEE Trans. Instrum. Meas.*, vol. 53, no. 5, pp. 1370–1377, Oct. 2004.

[4] A. Ferrero, M. Lazzaroni, and S. Salicone, "A calibration procedure for a digital instrument for electric power quality measurement," *IEEE Trans. Instrum. Meas.*, vol. 51, no. 4, pp. 716–721, Aug. 2002.

[5] G. Betta, C. Liguori, and A. Pietrosanto, "Structured approach to estimate the measurement uncertainty in digital signal elaboration algorithms," *Proc. Inst. Elect. Eng.—Sci. Meas. Technol.*, vol. 146, no. 1, pp. 21–26, Jan. 1999.

[6] E. Ghiani, N. Locci, and C. Muscas, "Auto-evaluation of the uncertainty in virtual instruments," *IEEE Trans. Instrum. Meas.*, vol. 53, no. 3, pp. 673–677, Jun. 2004.

[7] S. Nuccio and C. Spataro, "Approaches to evaluate the virtual instrumentation measurement uncertainties," *IEEE Trans. Instrum. Meas.*, vol. 51, no. 6, pp. 1347–1352, Dec. 2002.

[8] D. A. Lampasi and L. Podestà, "A practical approach to evaluate the measurement uncertainty of virtual instruments," in *Proc. 21st IEEE IMTC*, May 2004, vol. 1, pp. 150–156.

[9] *Guide to the Expression of Uncertainty in Measurement—Supplement 1—Numerical Methods for the Propagation of Distributions (Draft)*, 2004.

[10] M. J. Korczynski and A. Hetman, "Calculation of expanded uncertainties without knowledge of coverage factor," in *Proc. IMEKO TC-4 13th Symp.*, Sep. 2004, vol. 1, pp. 128–133.

[11] M. Cox and P. Harris. 2003. *Up a GUM Tree? Try the Full Monte!*, Middlesex, U.K.: NPL. [Online]. Available: http://www.npl.uk

[12] Z. A. Karian and E. J. Dudewicz, *Fitting Statistical Distributions: The Generalized Lambda Distribution and Generalized Bootstrap Methods*. Boca Raton, FL: CRC, 2000.

[13] W. G. Gilchrist, *Statistical Modelling With Quantile Functions*. Boca Raton, FL: CRC, 2000.

[14] E. W. Weisstein. 2005. *Pearson System. From MathWorld*. [Online]. Available: http://mathworld.wolfram.com/PearsonSystem.html

[15] NIST/SEMANTECH. 2005. *E-Handbook of Statistical Methods*. [Online]. Available: http://www.itl.nist.gov/div898/handbook

[16] A. Lakhany and H. Mausser, "Estimating the parameters of the generalized lambda distribution," *Algo Res. Q.*, vol. 3, no. 3, pp. 47–58, Dec. 2000.

[17] M. Freimer, G. Mudholkar, G. Kollia, and C. Lin, "A study of the generalized Tukey lambda family," *Commun. Stat., Theory Methods*, vol. 17, no. 10, pp. 3547–3567, 1988.

[18] E. W. Weisstein. 2005. *Beta Function. From MathWorld*. [Online]. Available: http://mathworld.wolfram.com/BetaFunction.html

[19] A. Ozturk and R. Dale, "Least squares estimation of the parameters of the generalized lambda distribution," *Technometrics*, vol. 27, no. 1, pp. 81–84, Oct. 1985.

[20] R. King and H. MacGillivray, "A starship estimation method for the generalized lambda distributions," *Aust. N. Z. J. Stat.*, vol. 41, no. 3, pp. 353–374, Oct. 1999.

[21] A. Lakhany and H. Mausser, "The extended generalized lambda distribution system for fitting distributions to data," *Commun. Stat.*, vol. 25, no. 3, pp. 611–642, 1996.

[22] Z. A. Karian and E. J. Dudewicz, "Fitting the generalized lambda distribution to data: A method based on percentiles," *Commun. Stat.*, vol. 28, no. 3, pp. 793–819, 1999.

[23] A. Zanobini, G. Iuculano, and G. Pellegrini, "The evaluation of the combined standard uncertainty for nonlinear models," in *Proc. 19th IEEE IMTC*, May 2002, vol. 1, pp. 359–363.

[24] G. D'Antona, "Expanded uncertainty and coverage factor computation by higher order moments analysis," in *Proc. 21st IEEE IMTC*, May 2004, vol. 1, pp. 234–238.

[25] *IEEE Standard for a Smart Transducer Interface for Sensors and Actuators*, IEEE Std. 1451.2, 1997.

Domenico Alessandro Lampasi (S'99–M'06) received the M.Sc. degree in electronic engineering and the Ph.D. degree in electrical engineering from the University of Rome "La Sapienza," Rome, Italy, in 2002 and 2006, respectively.

He is currently a Contract Researcher with the University of Rome "La Sapienza." His research interests include measurement uncertainty, virtual instrumentation, accurate signal generation, and power quality.

Fabio Di Nicola (S'04) received the M.Sc. degree in electrical engineering from the University of L'Aquila, L'Aquila, Italy, in 2003. He is currently working toward the Ph.D. degree in electrical engineering from the University of Rome "La Sapienza," Rome, Italy.

His current research interests include analysis and optimization of intervals between periodic calibrations for measurement instruments, expression of measurement uncertainty, virtual instrumentation, and performance evaluation of Hi-Fi systems.

Luca Podestà was born in Rimini, Italy, in 1961. He received the M.Sc. degree in electrical engineering from the University of Rome "La Sapienza," Rome, Italy, in 1989.

In 1991, he joined the Department of Electrical Engineering, University of Rome "La Sapienza," Rome, Italy, as a Researcher in electrical and electronic measurements. Since 1995, he has been an Assistant Professor in electrical measurements with the Faculty of Engineering. His main research interests include power quality, electromagnetic compatibility, web-based sensors, and didactic.