Variance gradients and uncertainty budgets for nonlinear measurement functions with independent inputs

2013 Meas. Sci. Technol. 24 025002

(http://iopscience.iop.org/0957-0233/24/2/025002)

IOP Publishing Measurement Science and Technology

Meas. Sci. Technol. **24** (2013) 025002 (16pp) doi:10.1088/0957-0233/24/2/025002

Variance gradients and uncertainty budgets for nonlinear measurement functions with independent inputs

Mark Campanelli[1], **Raghu Kacker and Rüdiger Kessel**

National Institute of Standards and Technology, 100 Bureau Drive, Gaithersburg, MD 20899, USA

E-mail: mark.campanelli@nrel.gov

Received 18 July 2012, in final form 4 December 2012
Published 9 January 2013
Online at stacks.iop.org/MST/24/025002

Abstract

A novel variance-based measure for global sensitivity analysis, termed a variance gradient (VG), is presented for constructing uncertainty budgets under the *Guide to the Expression of Uncertainty in Measurement* (GUM) framework for nonlinear measurement functions with independent inputs. The motivation behind VGs is the desire of metrologists to understand which inputs' variance reductions would most effectively reduce the variance of the measurand. VGs are particularly useful when the application of the first supplement to the GUM is indicated because of the inadequacy of measurement function linearization. However, VGs reduce to a commonly understood variance decomposition in the case of a linear(ized) measurement function with independent inputs for which the original GUM readily applies. The usefulness of VGs is illustrated by application to an example from the first supplement to the GUM, as well as to the benchmark Ishigami function. A comparison of VGs to other available sensitivity measures is made.

Keywords: variance gradient, global sensitivity analysis, variance-based sensitivity measure, sensitivity index, importance measure, uncertainty budget, nonlinear measurement function, GUM, Monte Carlo method, Ishigami function

1. Introduction

This paper introduces a novel variance-based measure for global sensitivity analysis, termed a variance gradient (VG), for the construction of uncertainty budgets involving scalar-valued nonlinear measurement functions with independent inputs. Measurement functions represent an explicit functional relationship between the measurand (the output quantity) and one or more input quantities to a measurement model [1, sections 2.48 and 2.49]. VGs have been developed with regard to measurement problems in which the first supplement to the *Guide to the Expression of Uncertainty in Measurement* (GUM) [2, 3] would be applied because measurement function linearization, or higher order approximation, is inadequate. A VG incorporates the nonlinear effects of the measurement function over the support of the independent inputs. However, unlike certain other sensitivity measures for nonlinear functions of independent input random variables (RVs), the collection of VGs (one for each input) does not decompose the output RV's variance into an apportionment among the inputs.

As an analysis tool for reducing the uncertainty in the result of a measurement, a VG *quantifies* the relative reduction in the variance of the measurand resulting from a small relative reduction in the variance of a given input. For example, a VG with value 0.25 for a particular input means that one anticipates that a measurand's variance will be reduced by approximately 0.25% if the input's variance is reduced by 1%. Ranking of inputs by VG value makes VGs useful in the construction of an uncertainty budget for the result of a measurement. Furthermore, for a linear(ized) measurement function with independent inputs, the collection of VGs reduces to the commonly understood variance decomposition/apportionment and therefore provides an alternative to this 'uncertainty

[1] Present address: Measurements and Characterization Group, National Center for Photovoltaics, National Renewable Energy Laboratory, 15013 Denver West Parkway, Golden, CO 80401, USA.

contribution' interpretation. For a large class of measurement problems, VGs can be computed in a straightforward manner with a Monte Carlo method that uses first-order partial derivatives of the measurement function. Thus, VGs integrate well with the methodology of the first supplement to the GUM.

1.1. Terminology

Uncertainty quantification often involves both uncertainty analysis and sensitivity analysis of models [4]. The goal in uncertainty analysis is to quantify the uncertainty in the output of a model. The goal in sensitivity analysis is a better understanding of the origination of the uncertainty in the model output, e.g., from input uncertainties or modelling assumptions, and how this uncertainty may be controlled. Uncertainty analyses and sensitivity analyses typically go hand in hand. For measurements, uncertainty budgets provide a quantitative summary of both analyses.

The traditional *local* sensitivity analysis considers how a model output changes with small changes around a single selection of inputs. In models described by a deterministic mathematical function, this is often quantified via first-order partial derivatives evaluated at a single-input point. For uncertainty quantification and other purposes, a broader definition of sensitivity analysis considers a range of model input points. The term *global sensitivity analysis* indicates that the analysis accounts for the sensitivity of the model over a large set of model inputs, especially when the model has nonlinear effects combined with uncertainty in the inputs described by a joint probability distribution [5].

While derivatives can still play a role in *global* sensitivity analyses, other mathematical measures have been developed [5, 6]. Some measures require independent inputs and/or the function to be linear. The nonlinearity of a function may create *interaction effects* between inputs. This occurs when the nonlinear function is not additive[2]. We distinguish these interaction effects from *dependence effects* that arise when the inputs are not independent. When both are present, these two effects can interact.

In variance decomposition settings, the adjective *total* has been used to describe global sensitivity measures that account for all interaction effects, while the adjective *first-order* has been used for measures that do not account for any such interactions [5]. The term *one-at-a-time* has been used to describe a sensitivity measure that fixes all inputs but one to their expected values [6]. Thus, one-at-a-time measures are neither completely local nor entirely global. Also, a first-order measure is not necessarily a one-at-a-time measure.

Saltelli *et al* [7] offer the following definition of sensitivity analysis:

> The study of how uncertainty in the output of a model (numerical or otherwise) can be apportioned to different sources of uncertainty in the model input.

The idea of a decomposition of the output variance into an apportionment among the inputs often leads to the use of the term *contribution analysis* instead of sensitivity analysis. However, in this paper, we propose a sensitivity measure that is *not* constructed to decompose/apportion variance. Moreover, 'contribution analysis' has other common uses, such as when the model output is the sum of several components, each attributed to a different group of inputs. Thus, to avoid confusion, we use the term *sensitivity analysis* instead of contribution analysis.

Considering the above discussion of terminology, we have selected the term VG for the variance-based global sensitivity measure presented here. VGs are global because they consider nonlinear effects over the support of the independent inputs. VG specifically refers to the limit of a particular ratio involving relative changes in output and input variances.

Table 1. Classification of measurement problems under the GUM framework.

		Measurement function	
		Linearizable	Nonlinearizable
Input RVs	Independent	Linearizable and independent	Nonlinearizable and independent
	Dependent	Linearizable and dependent	Nonlinearizable and dependent

1.2. Outline of the paper

Section 2 describes the motivation for VGs with respect to the GUM framework and uncertainty budgets. The mathematical derivation and implementation of VGs follows in section 3. Section 4 provides illustrative examples of VGs, including the construction of an uncertainty budget for a metrology example taken from the first supplement to the GUM, as well as for the Ishigami function, which is a benchmark for comparing sensitivity analysis measures. Section 5 discusses the use of VGs in measurement problems in relation to other available sensitivity measures, as well as the potential extension of VGs to an even larger class of measurement problem settings. The appendix contains technical details arising in the mathematical derivation of the formula used to compute VGs.

2. The GUM framework, uncertainty budgets and problem settings

This section reviews the GUM framework and uncertainty budgets for results of measurements stated under this framework. The familiar reader may wish to skim this material while noting (1) the classification of measurement problems in table 1, (2) the discussion of various sensitivity measures that could appear in uncertainty budgets, and (3) the description of three problem settings for sensitivity analysis that inform the choice of sensitivity measure(s).

[2] An *additive function* has the form $g(x_1, \ldots, x_N) = g_1(x_1) + \cdots + g_N(x_N)$. In particular, all additive functions are linear in the transformed variables $g_1(x_1), \ldots, g_N(x_N)$ and all linear functions are additive. A *linear function* has the form $g(x_1, \ldots, x_N) = c_0 + c_1 x_1 + \cdots + c_N x_N$, where $c_0, c_1, \ldots, c_N$ are constants. (Sometimes the term *affine* is used instead of linear when $c_0 \neq 0$.) An example of a nonlinear additive function is $g(x_1, x_2) = 1 - \pi\, e^{x_1} + x_2 + 5x_2^2$. The nonlinear function $g(x_1, x_2, x_3) = x_1 x_2 - 2x_3$ is not additive and admits a second-order interaction effect between x_1 and x_2.

2.1. The GUM framework

The GUM [2], its first supplement (GUM-S1) [3] and its second supplement (GUM-S2) [8] have provided an internationally accepted approach to the evaluation and expression of measurement uncertainty [9]. According to the GUM, no *result of a measurement* is complete without a statement quantifying the uncertainty of measurement [2, section B.2.11, note 2]. The GUM-S2 extends both the GUM and the GUM-S1 methodologies to vector-valued output quantities. Despite certain differences between the GUM and the GUM-S1 [3, section 5.11.4], we collectively term the GUM, GUM-S1 and GUM-S2 methodologies as the GUM framework. (This is not to be confused with the GUM-S1's reference to the GUM's methodology as the GUM uncertainty framework [3, section 3.18].) Furthermore, we consider the main difference between the GUM and the GUM-S1 to be their methodologies for the *evaluation* of uncertainty, but not for the *expression* of uncertainty.

Underlying the GUM framework is the notion that a *particular quantity* [2, section B.2.1] to be measured, or *measurand* [2, section B.2.9], has a *state-of-knowledge probability distribution* associated with it that 'represents belief probabilities about the possible values of a quantity based on all available information' [10]; see also [2, sections 3.3.5, 4.1.6 and G.6.6] and [3, section 5.11.2]. In addition to measurement uncertainty, this state of knowledge generally includes *definitional uncertainty* arising from an incompletely defined measurand; see [1, section 2.27] and [2, section D.3.4].

Both the GUM and the GUM-S1 provide guidance concerning *epistemic* uncertainty in scalar, real-valued measurands. To quote the GUM:

> This guide is primarily concerned with the expression of uncertainty in the measurement of a well-defined physical quantity—the measurand—that can be characterized by an essentially unique value. If the phenomenon of interest can be represented only as a distribution of values or is dependent on one or more parameters, such as time, then the measurands required for its description are the set of quantities describing that distribution or that dependence [2, section 1.2].

Thus, situations in which there is inherent random variability, or *aleatoric* uncertainty, in the measurand require the quantification of epistemic uncertainty in one or more parameters used to describe the variability. Furthermore, we distinguish the aleatoric uncertainty in a measurement process generating *indications* [1, section 4.1] from the resulting epistemic uncertainty in a measurand characterized by an essentially unique value.

In a complete *result of a measurement* [2, section B.2.11], both the single value selected for the measurand and the associated uncertainty are quantified by mathematically modelling the measurand as a RV, also see [1, section 2.9]. The single value selected for a measurand is commonly called an *estimate* of the measurand. Such language often entails a broader meaning for the term estimate than is specified in the GUM [2, section C.2.26], and the term *measured quantity value* [1, section 2.10] is more exact, if less common. The *uncertainty* is defined in the GUM as 'a parameter, associated with the result of a measurement, that characterizes the dispersion of the values that could reasonably be attributed to the measurand' [2, section B.2.18].

Under the GUM framework, the estimate of a measurand is typically the expected value of the state-of-knowledge probability distribution of the measurand (a measure of location) [2, section 4.1.1, note 3] and [11, 12]. The *(combined) standard uncertainty* is the standard deviation of this distribution (a measure of scale/spread) [2, sections 2.3.1 and 2.3.4]. The GUM presumes that both the expected value and the standard deviation are well defined for the state-of-knowledge probability distribution of the measurand. Unlike the moments of a distribution, *coverage intervals* always exist [1, section 2.36; 3, sections 3.12 and 4.11]. The GUM-S1 methodology is capable of determining coverage intervals, even when the expected value and/or standard deviation do not exist [3, section 5.1.1, note 2].

The *International Vocabulary of Metrology (VIM)* defines *measurement model* as a 'mathematical relation among all quantities known to be involved in a measurement' [1, section 2.48][3]. A measurement model is derived from established scientific *principles of measurement* [2, section B.2.6], as well as the understanding of the measurement process of metrologists. The inputs to the measurement model determine the value of the measurand and include, but are not limited to, *influence factors* [2, section B.2.10], corrections for recognized *systematic errors* [2, sections B.2.22 and B.2.23] and standard reference data/materials. Some inputs may themselves be results of measurements. Ultimately, the measurement model dictates how the state-of-knowledge joint probability distribution of the inputs translates into a state-of-knowledge probability distribution of the measurand.

The measurement model can take the form of a *measurement equation* [10], which defines a mathematical relationship, possibly implicit [1, section 2.48, note 1], between the measurand and one or more inputs upon which the measurand depends[4]. *Measurement functions* [1, section 2.49], in which the measurand is an explicit mathematical function of the inputs, occur in many measurement problems. Measurement functions appear throughout the GUM framework [2, section 4.1.1], although the GUM exclusively uses the term *measurement model*. In certain cases, a measurement equation *implicitly* defines a corresponding measurement function.

For the application of this work, we assume the existence of a real-valued measurement function with real-valued inputs and that the state-of-knowledge (marginal) probability distributions of the inputs and resulting measurand have

[3] As noted in [10, footnotes 11 and 12], a measurement model should not be confused with a *statistical model* of a measurement process. Statistical models relate observed data (or indications) to statistical parameters. Recently, *observation equations* have been suggested as statistical models for metrological applications, to which Bayesian inference methods readily apply [13].

[4] The term *measurement equation* was used in [14] with a meaning similar to the aforementioned *observation equation* rather than the meaning indicated here.

sufficiently well-defined moments. A *de facto* understanding throughout this work (and the GUM framework) is that the expected value and the standard deviation together summarize an underlying state-of-knowledge probability distribution [10].

A fundamental *linearizability criterion* underlies the application of the GUM. A *linearizable measurement function* should be linear, or its linearization about the estimates of the inputs must be *adequate*, i.e. be mathematically tractable, practical to compute and give a sufficient approximation over the uncertainty region of the inputs. In this setting, the complete characterization of the state-of-knowledge joint probability distribution of the inputs is not necessary, because the expected values, standard deviations and covariances (or, equivalently, correlation coefficients) of the inputs are sufficient to evaluate the expected value and the standard deviation of the measurand [10]. (For a linear function, there are no interaction effects, and the covariances sufficiently capture the dependence effects.)

In particular, the expected value of a *linear* function of jointly distributed inputs is precisely the value of the function at the expected values of the inputs, and the (combined) standard uncertainty in the measurand can be computed using the *propagation of uncertainties formula* [2, sections 3.3.6 and 5][5]. Under certain additional conditions [3, section 5.7.2], the GUM methodology gives reliable evaluations of coverage intervals for the measurand, which correspond to *expanded* standard uncertainties [10].

The Monte Carlo method of the GUM-S1 offers an alternative to the GUM methodology for evaluating the expected value, standard deviation and coverage intervals of the measurand. The GUM-S1 methodology involves constructing an empirical output distribution for the measurand by repeated numerical evaluation of the measurement function at samples drawn from the joint distribution of the inputs. Thus, the GUM-S1 finds particular use when the linearization of the measurement function is inadequate for application of the GUM, the computational cost of evaluating the measurement function is not prohibitive and the joint distribution of the inputs can numerically be sampled; see [3, section 5.4.3] and [10, 15–17][6].

Enough samples must be taken to achieve sufficient convergence of the Monte Carlo method; see [3, section 7.9] and [18, 19]. The resulting numerical approximation of the state-of-knowledge probability distribution of the measurand [3, sections 5.9 and 7.5] can be used to estimate the expected value and the standard deviation of the true state-of-knowledge probability distribution [3, sections 5.9 and 7.6]. Propagation of distributions via Monte Carlo also offers a method for estimating coverage intervals for the measurand with fewer restrictive conditions than the GUM, [3, sections 5.9 and 7.7] and [10, 15, 16, 18]. For computation of coverage intervals, more samples must generally be taken to sufficiently resolve the probability in the tails of a distribution, as opposed to sufficiently estimating the expected value and standard deviation of the distribution [3, section 7.9.4, note 7]. Various sampling strategies may accelerate convergence [20].

The GUM framework has generally improved both the practice of metrology and the meaningful exchange of results of measurements, notwithstanding certain recognized issues and limitations with both the GUM and the GUM-S1 [10, 21, 22]. For example, the ongoing research is addressing the GUM framework's precise relation to frequentist and Bayesian statistical methodologies [10, 13, 15, 16, 23–29], as well as the difficulties in determining when linearization of the measurement function is adequate [30], which is also addressed in the GUM-S1 [3, section 8]. For a comprehensive review of the evolution of the expression of uncertainty in measurement from the error analysis to the GUM framework, the reader is referred to [10]. A short summary of the GUM-S1, as compared to the GUM, is given in [31].

2.2. Uncertainty budgets and sensitivity measures

An uncertainty budget provides a valuable summary of a result of a measurement [1, section 2.33]. Construction of an uncertainty budget includes recording the estimates and standard uncertainties of the inputs and measurand. By including sensitivity measures, an uncertainty budget may also quantify relationships between the measurand's uncertainty and the uncertainty of the inputs. Such information typically provides guidance on how a measurement uncertainty might best be reduced [10].

The ease/difficulty of construction of an uncertainty budget depends upon the type of measurement problem at hand. The nature of the input RVs and the measurement function gives rise to the following two important classifications of the measurement problem:

(i) independent versus dependent input RVs,
(ii) linearizable versus nonlinearizable measurement function.

These two classifications are not mutually exclusive, and consideration of the possible combinations gives four possible measurement problems, which are summarized in table 1.

Because the GUM makes no suggestion for the content/format of an uncertainty budget, a task force of the European Co-operation for Accreditation formalized requirements for uncertainty budgets for linear(ized) measurement functions with independent inputs [32]. Such uncertainty budgets quantify an uncertainty contribution (UC) for each (independent) input to the uncertainty in the measurand [10]. An UC for a given input is the product of the input's sensitivity coefficient (SC) and its standard uncertainty (i.e. standard deviation), see [2, section 5.1.3] and [10]. These UCs are squared to become the terms in the propagation of uncertainties formula whose sum gives the variance of the measurand. Thus, the squared UCs can be used to make a full decomposition of the variance of the measurand into an apportionment among the inputs [10]. Dividing the UCs by the combined standard uncertainty (i.e. standard deviation) of the measurand gives the unit-less σ*-normalized derivatives* described in [5]. Squaring the σ-normalized derivatives allows each input's contribution to the measurand's variance to be

[5] For reasons given in [10], we prefer the name *propagation of uncertainties formula* over the GUM framework's *law of propagation of uncertainty*.

[6] If the input RVs are independent, then numerically sampling from each input's marginal distribution becomes sufficient.

registered on a 0–100 per cent scale, and the sum of the squares of the σ-normalized derivatives is 1.

Kessel *et al* [33] developed *coefficients of contribution*, with one coefficient attributed to each input, that allow the construction of uncertainty budgets for problems where the measurement function is linear(ized) and dependence effects are characterized via known (linear) correlations between inputs. While these dimensionless measures sum to 1 and decompose the measurand's variance into an apportionment among the inputs, the correlation structure can cause a given coefficient of contribution to be negative or have an absolute value greater than 1. In the case of independent inputs, the coefficients of contribution reduce to the above-mentioned variance decomposition/apportionment derivable from the σ-normalized derivatives of the inputs. Moreover, they have a natural interpretation as the squares of the (linear) correlations between the measurand and the respective inputs [33].

For uncertainty budgets, annex B of GUM-S1 briefly discusses the use of so-called nonlinear sensitivity coefficients (NLSCs) in place of the SCs described above. NLSCs are derived from what we term here *nonlinear uncertainty contributions* (NLUCs). NLSCs and NLUCs are described more thoroughly in [34, section 5.8] and [35, section B.3]. A NLUC for a given input is computed as the standard deviation of the measurand that results from holding all other inputs fixed at their expected values. As with UCs, NLUCs are not normalized. The corresponding NLSC is derived by dividing the NLUC by the given input's standard deviation. NLSCs can be viewed as a *one-at-a-time* generalization of SCs to nonlinear measurement equations that reduce to the absolute values of the SCs for linear(ized) measurement functions [3, section B.1]. An input independence assumption apparently underlies the one-at-a-time treatment used in NLUCs (and the corresponding NLSCs), because each input is allowed to vary independently of the other inputs' fixed values in a variance computation. Because of potential interaction effects between inputs that are missed by NLUCs, the sum of the squares of the NLUCs generally does not equal the measurand's variance. The discussion of NLUCs for uncertainty budgets in GUM-S1 [3, section B] ends with the caveat that

> . . . in cases for which (a valid implementation of) the propagation of distributions is more appropriate, an uncertainty budget should be regarded as a qualitative tool [3, section B.2].

Without specific regard to metrological applications, progress has been made in global sensitivity analysis for a general class of models given by real-valued *nonlinear* functions of *independent* inputs. Based upon the earlier work of I M Sobol' and others [36], Saltelli *et al* [5] have advanced *variance-based sensitivity indices* (SIs) for such models. In this combinatorial approach, the collection of SIs sums to 1 and completely decomposes the output's variance into an apportionment (0–100 per cent) among all combinations of the independent inputs, including inputs taken individually. The full variance decomposition distinguishes all combinations of independent inputs producing all orders of nonlinear interaction effect *between* these inputs. Here, *order* refers to the number of independent inputs that appear (i.e. interact) in a given term of the decomposition. A brief outline of SIs is presented next, based upon the more comprehensive development in [5].

Suppose the following function g of N independent input RVs

$$Y = g(X_1, \ldots, X_N)$$

is square-integrable (so that the output RV Y has finite variance). In metrology applications, g would be a measurement function. SIs are based upon a *particular* high-dimensional model representation (HDMR) of g with the following form:

$$g = g_0 + \sum_{n_1=1}^{N} g_{n_1} + \sum_{n_1=1}^{N-1} \sum_{n_2=n_1+1}^{N} g_{(n_1,n_2)} + \cdots + g_{(1,2,\ldots,N)},$$

in which g_0 is a constant function equal to the expectation of Y and all other terms are square-integrable with zero expectation. For independent inputs, this HDMR (with 2^N terms) is the unique choice with such properties. Each combination of inputs, taken in groups of size 1 to N, occurs as the input to exactly one of the functions in the summation. For example, $g_{(1,4,5)}$ would be a function of only the inputs X_1, X_4 and X_5 that captures the interaction effects between these inputs up to and including the third-order (or three-way) interaction.

Corresponding variance decomposition of the output Y is derived from the HDMR as

$$\mathsf{V}(Y) = \sum_{n_1=1}^{N} \mathsf{V}_{n_1} + \sum_{n_1=1}^{N-1} \sum_{n_2=n_1+1}^{N} \mathsf{V}_{(n_1,n_2)} + \cdots + \mathsf{V}_{(1,2,\ldots,N)},$$

where $\mathsf{V}(\cdot)$ is the variance operator. (This involves subtracting all lower order variance contributions from the variance of any given term in the HDMR; for details, see [5].) Dividing all terms by $\mathsf{V}(Y) > 0$ gives the following relationship among the (non-negative) SIs:

$$1 = \sum_{n_1=1}^{N} S_{n_1} + \sum_{n_1=1}^{N-1} \sum_{n_2=n_1+1}^{N} S_{(n_1,n_2)} + \cdots + S_{(1,2,\ldots,N)}.$$

This combinatorial approach comprehensively accounts for all interaction effects between independent inputs and completely distinguishes the interaction effects by order. However, the computational expense of the complete decomposition becomes prohibitive as the number of inputs increases, because the number of indices $(2^N - 1)$ grows exponentially with N.

To reduce computational expense, partial decompositions are typically employed. The *first-order sensitivity indices* (FOSIs) are $S_1, \ldots, S_N$. The collection of *total sensitivity indices* (TSIs) is denoted $S_{\mathrm{T}1}, \ldots, S_{\mathrm{T}N}$. A TSI for an input is the sum of all SIs that involve that input. For example, for a function of three inputs ($N = 3$),

$$\begin{aligned} S_{\mathrm{T}1} &= S_1 + S_{(1,2)} + S_{(1,3)} + S_{(1,2,3)}, \\ S_{\mathrm{T}2} &= S_2 + S_{(1,2)} + S_{(2,3)} + S_{(1,2,3)}, \\ S_{\mathrm{T}3} &= S_3 + S_{(1,3)} + S_{(2,3)} + S_{(1,2,3)}. \end{aligned}$$

The FOSIs and TSIs provide a description of how the variance of each input contributes to the variance of the output when considered individually and including all interaction effects,

respectively. Comparing the TSI to the FOSI for a given input provides information about the importance of interaction effects for that input.

Additive functions have no interaction effects, so the FOSI and the TSI are equal for any given input. Thus, the collection of FOSIs for an additive function is sufficient to give a full decomposition of the variance of the output into an apportionment among the independent inputs. Furthermore, FOSIs give the same variance decomposition/apportionment mentioned earlier involving the squares of σ-normalized derivatives for linear(ized) measurement functions.

Mathematically, the HDMR allows both FOSIs and TSIs to be represented in terms of variances of conditional expectations. Specifically, for FOSIs, one can show that

$$S_n = \frac{\mathsf{V}(\mathsf{E}(Y|X_n))}{\mathsf{V}(Y)},$$

where in the numerator the expectation operator $\mathsf{E}(\cdot)$ is conditional upon a fixed value of X_n and the variance is taken over the distribution of X_n. The *law of total variance* (or *variance decomposition formula*) is the identity

$$\mathsf{V}(Y) = \mathsf{V}(\mathsf{E}(Y|X_n)) + \mathsf{E}(\mathsf{V}(Y|X_n)).$$

Applying this identity to the expression for S_n gives, equivalently,

$$S_n = \frac{\mathsf{V}(Y) - \mathsf{E}(\mathsf{V}(Y|X_n))}{\mathsf{V}(Y)} = \mathsf{E}\left(\frac{\mathsf{V}(Y) - \mathsf{V}(Y|X_n)}{\mathsf{V}(Y)}\right). \quad (1)$$

Thus, an FOSI can be interpreted as the expected relative reduction in the variance of the output Y that can be achieved by fixing the input X_n. Higher-order SIs have analogous interpretations in which multiple inputs are fixed together so that the various interaction effects are completely distinguished. For TSIs, one can similarly show that

$$S_{\mathrm{T}n} = 1 - \frac{\mathsf{V}(\mathsf{E}(Y|\mathbf{X}_{\sim n}))}{\mathsf{V}(Y)},$$

where $\mathbf{X}_{\sim n}$ denotes that all inputs *except* X_n are fixed in the conditional expectation and the variance is taken over the distribution of all inputs *except* X_n.

These representations indicate that FOSIs and TSIs may be estimated using a Monte Carlo method [5, p 164]. However, the nesting of variance and expectation operations typically leads to computationally intensive algorithms. For computationally expensive functions and/or large numbers of inputs, Saltelli *et al* suggest that the input sensitivities should be pre-screened using a less-expensive method, such as an *elementary effect test*, followed by a variance-based sensitivity analysis for only those inputs deemed significant [5].

Both input dependences and measurement function nonlinearities complicate the quantification of the relationship between the input uncertainties and the output uncertainty. In particular, characterizing the input dependences in a nonlinear setting may require an understanding of the dependence effects beyond linear correlation, which may require working directly with (arbitrary) joint distributions [17].

In some applications, a joint distribution of the inputs can be transformed into a distribution with an independence structure amenable to sensitivity analyses requiring independent inputs. Examples include the principal component analysis of multivariate normal distributions and the Rosenblatt transformation [37], which is a change-of-variables that transforms any absolutely continuous joint distribution into a uniform distribution with independent inputs. Alternatively, a joint distribution of the inputs could be decomposed as the output of one or more functions of independent inputs and the sensitivity analysis conducted with respect to the independent inputs. An important consideration in such schemes is how to interpret the sensitivity analysis results with respect to the original inputs to the measurement function. We know about no practical direct methods that have been developed for handling problems that are *both* nonlinearizable and have dependent inputs, and this represents an open area for research [6].

2.3. Problem settings for sensitivity analysis

Saltelli *et al* emphasize that careful consideration of the problem *setting* is critical in the selection of appropriate sensitivity measures [5]. Three settings that they identify, which have relevance to metrology, are the following[7].

(i) *Factor prioritization*: one seeks to rank each input by the expectation of how much the output variance would be reduced if the uncertainty in an input were completely eliminated. The expectation is computed over an input's probability distribution representing the potential true value for the input. Furthermore, groups of inputs may be fixed simultaneously.

(ii) *Factor fixing*: one seeks to determine which inputs have very little effect on the output variance over their ranges of uncertainty, and thus could be fixed anywhere in their ranges of uncertainty without significant effect.

(iii) *Variance cutting*: one seeks to (efficiently) reduce the output variance below a given target by reducing the variance of one or more inputs.

For nonlinear functions with independent inputs, FOSIs have found use for factor prioritization in which interaction effects between inputs are not considered. TSIs are also useful for factor prioritization in which only a single input is to be fixed. Additional higher order SIs can be computed to better inform the factor prioritization of groups of inputs. TSIs are particularly useful measures for factor fixing because they comprehensively account for interaction effects between inputs. Furthermore, these settings are not mutually exclusive [5].

With regard to the above three settings for global sensitivity analysis, the VGs introduced in this paper are measures most closely related to factor prioritization and variance cutting. Like SIs, but unlike many one-at-a-time measures, VGs capture interaction effects between inputs. However, VGs capture additional nonlinear effects within a single input and have a different quantitative meaning than SIs. These properties can make VGs advantageous over SIs in the prioritization of inputs for variance reduction purposes.

VGs can be used in concert with other measures, including SIs for variance decomposition/apportionment and TSIs or

[7] Saltelli *et al* [5] use the term *factors* instead of inputs.

elementary effect tests for factor fixing. For linear(ized) measurement functions with independent inputs, VGs reduce to the common variance decomposition/apportionment involving σ-normalized derivatives in the propagation of uncertainties formula, thus providing an alternative to the 'uncertainty contribution' and 'output–input correlation' interpretations discussed above.

VGs generate a slightly different factor prioritization than described above, because VGs do not involve fixation of inputs to presumed true values. In many metrology applications, variance cutting is more easily accomplished via small reductions in the variance of an input than by a complete elimination of uncertainty. This consideration is related to an important problem setting for sensitivity analysis in metrology, namely the method used for evaluating the uncertainty in an input to a measurement function.

The GUM defines *type A evaluation (of uncertainty)* as a 'method of evaluation of uncertainty by the statistical analysis of series of observations' [2, sections 2.3.2 and 4.2] and *type B evaluation (of uncertainty)* as a 'method of evaluation of uncertainty by means other than the statistical analysis of series of observations' [2, sections 2.33 and 4.3]. Both evaluations ultimately correspond to underlying state-of-knowledge distributions [2, section 4.1.6]. An input to a measurement function may itself be the output of a different measurement function with inputs whose uncertainty evaluations included both type A and type B. Thus, the delineation between type A and type B evaluation is not always clear, and the GUM-S1 essentially does away with this distinction [3, sections 5.11.4(a) and 6].

We prefer the viewpoint proposed in [28], in which type A and type B evaluations are incorporated into a consistent Bayesian framework. Here, the type A evaluation uses a likelihood function and indication(s) from a measurement process with Bayes' rule to update a prior state-of-knowledge distribution to a posterior state-of-knowledge distribution. The prior chosen may be non-informative. (This type A evaluation appears in GUM-S1 [3, section 6.4.9].) The type B evaluation uses only a prior state-of-knowledge distribution, typically based upon expert judgement or the application of the principle of maximum entropy [3, 29].

Regardless of one's viewpoint regarding type A and type B evaluations, VGs are readily computed for inputs whose state-of-knowledge probability distributions have finite variance. However, the discrete nature of taking additional indications to reduce an input's variance suggests extending the fundamental idea behind the VGs introduced in this paper to sensitivity measures specific to the type A evaluation. This issue is discussed further in section 5.

3. Variance gradients: derivation and implementation

This section introduces VGs, providing the limit-based definition and a derivation of a limit-free formula for computing VGs that uses first-order partial derivatives of the measurement function. An interpretation of VGs in the case of linear(ized) measurement functions is provided, and a method for usefully combining VGs is given. The specific application of VGs to uncertainty budgets within the GUM framework is described. A technique for computing VGs using a Monte Carlo method is presented and discussed, and an iterative scheme for using VGs in variance cutting is suggested.

3.1. Notation

Wherever practical, the notational conventions of GUM-S1 are followed in this paper. For example, a normally distributed (Gaussian) RV with mean μ and strictly positive variance σ^2 is denoted $\mathsf{N}(\mu, \sigma^2)$, and a continuous, uniformly distributed (rectangular) RV with support $[a, b]$, $a < b$, is denoted $\mathsf{R}(a, b)$. $\mathsf{E}(\cdot)$ and $\mathsf{V}(\cdot)$ denote the expectation and variance operators, respectively. However, during the mathematical derivations $\mu_Y = \mathsf{E}(Y)$ and $\sigma_Y = \sqrt{\mathsf{V}(Y)}$ are used instead of y and $u(y)$, respectively, the latter notations (from the GUM framework) being reserved for stated results of measurements. Also, g is used to represent the measurement function, while the more standard f is used to denote a probability density function [38].

3.2. Definition and derivation

Consider the measurement function

$$Y = g(X_1, \ldots, X_N) = g(\mathbf{X}), \tag{2}$$

where $\mathbf{X} = (X_1, \ldots, X_N)$ is a vector of real-valued independent input RVs with respective (finite) expected values $\mu_{X_1}, \ldots, \mu_{X_N}$ and respective (finite) strictly positive variances $\sigma^2_{X_1}, \ldots, \sigma^2_{X_N}$. Throughout this work, assume that the measurement function g is differentiable, Borel-measurable (in the sense of measure-theoretic probability) and produces a real-valued output RV Y with (finite) expected value μ_Y and (finite) strictly positive variance σ^2_Y.

The VG of Y with respect to X_n, denoted G_{X_n} for any $n = 1, \ldots, N$, will be defined below as a real number that indicates how a small relative change in the variance of the input X_n proportionally translates into a relative change in the variance of the output Y. Consistent with the GUM framework, expected values are considered to be estimates of the inputs and measurand. Variance is used here as the measure of uncertainty, because of its mathematically desirable properties and its direct relation to the standard uncertainty in the GUM framework.

For the purpose of defining the VG G_{X_n} for any $n = 1, \ldots, N$, consider the following definition of $\widetilde{X}_n(\rho)$, which is an adjustment to X_n that reduces X_n's relative variance by the (signed) factor ρ while retaining the same expected value as X_n:

$$\widetilde{X}_n(\rho) \stackrel{\text{def}}{=} \sqrt{1-\rho}\,\left(X_n - \mu_{X_n}\right) + \mu_{X_n}, \tag{3}$$

where $-\infty < \rho < 1$. Note that $\widetilde{X}_n(0) = X_n$ for any $n = 1, \ldots, N$. Thus, the relative change in X_n's variance is

$$\frac{\sigma^2_{\widetilde{X}_n(\rho)} - \sigma^2_{X_n}}{\sigma^2_{X_n}} = \frac{(1-\rho)\sigma^2_{X_n} - \sigma^2_{X_n}}{\sigma^2_{X_n}} = -\rho,$$

which indicates a relative reduction in variance if $0 < \rho < 1$ and a relative increase in variance if $-\infty < \rho < 0$.

Furthermore, consider the corresponding adjustment to the measurement function output given by

$$\begin{aligned}\widetilde{Y}_n(\rho) &\stackrel{\text{def}}{=} g(X_1, \ldots, X_{n-1}, \widetilde{X}_n(\rho), X_{n+1}, \ldots, X_N)\\ &= g(\widetilde{\mathbf{X}}_n(\rho)),\end{aligned}$$

where $\widetilde{\mathbf{X}}_n(\rho) = (X_1, \ldots, X_{n-1}, \widetilde{X}_n(\rho), X_{n+1}, \ldots, X_N)$, so that $\widetilde{\mathbf{X}}_n(0) = \mathbf{X}$. $\widetilde{Y}_n(\rho)$ is assumed to have a (finite) expected value $\mu_{\widetilde{Y}_n(\rho)}$ and (finite) variance $\sigma^2_{\widetilde{Y}_n(\rho)}$ for sufficiently small ρ. Note that $\widetilde{Y}_n(0) = Y$ for any $n = 1, \ldots, N$ and that $\mu_{\widetilde{Y}_n(\rho)} \neq \mu_Y$ in general when $\rho \neq 0$.

The VG G_{X_n} is specifically defined as the ratio of the relative variance change in Y to the relative variance change in X_n, in the limit as the relative variance change in X_n tends to zero, assuming this limit exists. The input independence assumption allows separate consideration of the relative variance change in each input's marginal distribution.

For any $n = 1, \ldots, N$, G_{X_n} may be computed as follows:

$$\begin{aligned}G_{X_n} &\stackrel{\text{def}}{=} \lim_{\rho\to 0} \frac{\left(\sigma^2_{\widetilde{Y}_n(\rho)} - \sigma^2_Y\right)/\sigma^2_Y}{\left(\sigma^2_{\widetilde{X}_n(\rho)} - \sigma^2_{X_n}\right)/\sigma^2_{X_n}}\\ &= -\sigma_Y^{-2} \lim_{\rho\to 0} \frac{\mathsf{V}(\widetilde{Y}_n(\rho)) - \mathsf{V}(\widetilde{Y}_n(0))}{\rho}. \qquad (4)\end{aligned}$$

Rewriting the limit as a derivative evaluated at zero gives

$$\begin{aligned}G_{X_n} &= -\sigma_Y^{-2} \frac{\mathrm{d}}{\mathrm{d}z}\mathsf{V}(\widetilde{Y}_n(z))\bigg|_{z=0}\\ &= -\sigma_Y^{-2} \frac{\mathrm{d}}{\mathrm{d}z}[\mathsf{E}((\widetilde{Y}_n(z))^2) - (\mathsf{E}(\widetilde{Y}_n(z)))^2]|_{z=0}\\ &= -\sigma_Y^{-2} \frac{\mathrm{d}}{\mathrm{d}z}[\mathsf{E}((g(\widetilde{\mathbf{X}}_n(z)))^2) - (\mathsf{E}(g(\widetilde{\mathbf{X}}_n(z))))^2]|_{z=0}.\end{aligned}$$

If conditions permit taking derivatives inside the expectation operators, then

$$\begin{aligned}G_{X_n} &= -\sigma_Y^{-2}\left[\mathsf{E}\left(2g(\widetilde{\mathbf{X}}_n(z))\frac{\partial}{\partial z}g(\widetilde{\mathbf{X}}_n(z))\right)\right.\\ &\quad\left. - 2\mathsf{E}(g(\widetilde{\mathbf{X}}_n(z)))\mathsf{E}\left(\frac{\partial}{\partial z}g(\widetilde{\mathbf{X}}_n(z))\right)\right]\bigg|_{z=0}\\ &= -\sigma_Y^{-2}\left[\mathsf{E}\left(2g(\widetilde{\mathbf{X}}_n(z))\left.\frac{\partial g}{\partial X_n}\right|_{\widetilde{\mathbf{X}}_n(z)}\frac{\partial}{\partial z}\widetilde{X}_n(z)\right)\right.\\ &\quad\left. - 2\mathsf{E}(g(\widetilde{\mathbf{X}}_n(z)))\mathsf{E}\left(\left.\frac{\partial g}{\partial X_n}\right|_{\widetilde{\mathbf{X}}_n(z)}\frac{\partial}{\partial z}\widetilde{X}_n(z)\right)\right]\bigg|_{z=0}\\ &= -\sigma_Y^{-2}\left[\mathsf{E}\left(2g(\widetilde{\mathbf{X}}_n(z))\left.\frac{\partial g}{\partial X_n}\right|_{\widetilde{\mathbf{X}}_n(z)}\frac{X_n - \mu_{X_n}}{-2\sqrt{1-z}}\right)\right.\\ &\quad\left. - 2\mathsf{E}(g(\widetilde{\mathbf{X}}_n(z)))\mathsf{E}\left(\left.\frac{\partial g}{\partial X_n}\right|_{\widetilde{\mathbf{X}}_n(z)}\frac{X_n - \mu_{X_n}}{-2\sqrt{1-z}}\right)\right]\bigg|_{z=0}\\ &= -\sigma_Y^{-2}\left[\mathsf{E}\left(2g(\widetilde{\mathbf{X}}_n(0))\left.\frac{\partial g}{\partial X_n}\right|_{\widetilde{\mathbf{X}}_n(0)}\frac{X_n - \mu_{X_n}}{-2\sqrt{1-0}}\right)\right.\\ &\quad\left. - 2\mathsf{E}(g(\widetilde{\mathbf{X}}_n(0)))\mathsf{E}\left(\left.\frac{\partial g}{\partial X_n}\right|_{\widetilde{\mathbf{X}}_n(0)}\frac{X_n - \mu_{X_n}}{-2\sqrt{1-0}}\right)\right]\\ &= \sigma_Y^{-2}\left[\mathsf{E}\left(g(\mathbf{X})\left.\frac{\partial g}{\partial X_n}\right|_{\mathbf{X}}(X_n - \mu_{X_n})\right)\right.\\ &\quad\left. - \mathsf{E}(g(\mathbf{X}))\mathsf{E}\left(\left.\frac{\partial g}{\partial X_n}\right|_{\mathbf{X}}(X_n - \mu_{X_n})\right)\right].\end{aligned}$$

Simplifying the notation gives

$$\begin{aligned}&G_{X_n}\\ &= \frac{\mathsf{E}\left(g(\mathbf{X})\frac{\partial}{\partial X_n}g(\mathbf{X})(X_n - \mu_{X_n})\right) - \mu_Y\mathsf{E}\left(\frac{\partial}{\partial X_n}g(\mathbf{X})(X_n - \mu_{X_n})\right)}{\sigma_Y^2}\\ &= \frac{\mathsf{E}\left(g(\mathbf{X})\frac{\partial}{\partial X_n}g(\mathbf{X})(X_n - \mu_{X_n}) - \mu_Y\frac{\partial}{\partial X_n}g(\mathbf{X})(X_n - \mu_{X_n})\right)}{\sigma_Y^2},\end{aligned}$$

which can be written more compactly by factoring, using two alternative notations, as

$$\begin{aligned}G_{X_n} &= \frac{\mathsf{E}\left((g(\mathbf{X}) - \mu_Y)\frac{\partial}{\partial X_n}g(\mathbf{X})(X_n - \mu_{X_n})\right)}{\sigma_Y^2}\\ &= \frac{\mathsf{E}\left((Y - \mu_Y)\frac{\partial Y}{\partial X_n}(X_n - \mu_{X_n})\right)}{\sigma_Y^2}. \qquad (5)\end{aligned}$$

This last form may be particularly suitable for computation by a Monte Carlo method. Note that g must be differentiable and that partial derivatives of g must be computed (analytically or numerically). Also, a VG can have any sign and the sum of the collection of VGs for all the inputs is not necessarily 1.

The derivation of (5) required taking a derivative with respect to the parameter z inside two expectation operators, namely $\mathsf{E}(\widetilde{Y}_n(z))$ and $\mathsf{E}((\widetilde{Y}_n(z))^2)$, where z lies in an open interval containing zero in which $\widetilde{Y}_n(z)$ has finite expectation and variance. In general, expectation operators are Lebesgue integrals of measurable functions over finite measure spaces [39, 40]. Under these conditions, taking a derivative with respect to a parameter inside an expectation operator can be justified by exhibiting a non-negative RV with finite expectation that dominates the absolute value of the partial derivative of the parametrized RV [39, section 9.2] or [40, corollary 5.9]. Such a dominating RV is not a function of the parameter and dominates over an interval of interest for the parameter. Thus, a sufficient condition for the validity of (5) for computing G_{X_n} is the existence of dominating RVs W_1 and W_2, each with finite expectation, such that

$$\left|\frac{\partial}{\partial z}\widetilde{Y}_n(z)\right| \leqslant W_1 \quad \text{and} \quad \left|\frac{\partial}{\partial z}(\widetilde{Y}_n(z))^2\right| \leqslant W_2,$$

for all z in an open interval containing zero. Typically, dominating RVs can be established for a given measurement function (see the appendix). However, the reader is cautioned that the existence of the expectation in (5) is a necessary, although possibly not sufficient, condition for a VG to be computable by (5).

3.3. The linear special case

Consider the following special case of a linear measurement function (with independent inputs):

$$Y = g(\mathbf{X}) = c_0 + c_1X_1 + \cdots + c_NX_N,$$

where $c_0, \ldots, c_N$ are constants with $c_1, \ldots, c_N$ each nonzero. In this case, the computation of G_{X_n} using (5) reduces to a

common formula as follows:

$$G_{X_n} = \frac{\mathsf{E}\big((Y-\mu_Y)\frac{\partial Y}{\partial X_n}(X_n-\mu_{X_n})\big)}{\sigma_Y^2}$$
$$= \frac{\mathsf{E}([(c_0+c_1X_1+\cdots+c_NX_N)-(c_0+c_1\mu_{X_1}+\cdots+c_N\mu_{X_N})]c_n(X_n-\mu_{X_n}))}{\sigma_Y^2}$$
$$= \big[\mathsf{E}(c_1(X_1-\mu_{X_1})c_n(X_n-\mu_{X_n}))+\cdots+\mathsf{E}\big(c_n^2(X_n-\mu_{X_n})^2\big)$$
$$+\cdots+\mathsf{E}(c_N(X_N-\mu_{X_N})c_n(X_n-\mu_{X_n}))\big]\big/\sigma_Y^2$$
$$= \frac{c_n^2\mathsf{V}(X_n)}{\sigma_Y^2} = \frac{c_n^2\sigma_{X_n}^2}{\sigma_Y^2}$$
$$= \left(c_n\frac{\sigma_{X_n}}{\sigma_Y}\right)^2.$$

In the above derivation, independent inputs make the expectation of most terms in the numerator vanish, e.g., for $i \neq j$, $\mathsf{E}(c_i(X_i-\mu_{X_i})c_j(X_j-\mu_{X_j})) = c_i\,c_j(\mathsf{E}(X_i-\mu_{X_i})\mathsf{E}(X_j-\mu_{X_j})) = c_i\,c_j(0\cdot 0) = 0$.

In this linear special case, $0 < G_{X_n} < 1$ for all $n = 1,\ldots,N$, with $\sum_{n=1}^N G_{X_n} = 1$. Furthermore, for any input X_n, the constant $c_n = \frac{\partial Y}{\partial X_n}$ is the *sensitivity coefficient* (SC), $c_n\,\sigma_{X_n}$ is the *uncertainty contribution* (UC) and $c_n\frac{\sigma_{X_n}}{\sigma_Y}$ is the *σ-normalized derivative*. The above derivation shows that the VGs for a *linear* measurement function with independent inputs are precisely the squares of the σ-normalized derivatives. Thus, these VGs give the common decomposition of the output variance into an apportionment among the inputs. Furthermore, if we denote the VGs for a *linearized* measurement function derived from a given nonlinear measurement function as $G_{X_n}^{\mathsf{L}}$, $n = 1,\ldots,N$, then $\sum_{n=1}^N G_{X_n}^{\mathsf{L}} = 1$ follows immediately.

3.4. Combined variance gradients

The variance in the output may be reduced by simultaneously reducing the variance in more than one input. Suppose $(X_{n_1},\ldots,X_{n_j},\ldots,X_{n_J})$ is a nonempty sublist of input variables, with $j = 1,\ldots,J \leqslant N$ and $1 \leqslant n_1 < \cdots < n_j \cdots < n_J \leqslant N$. If $(X_{n_1},\ldots,X_{n_j},\ldots,X_{n_J})$ are simultaneously replaced with their reduced variance counterparts $(\widetilde{X}_{n_1}(\rho),\ldots,\widetilde{X}_{n_j}(\rho),\ldots,\widetilde{X}_{n_J}(\rho))$, then *the combined VG* (CVG), denoted $G_{(X_{n_1},\ldots,X_{n_j},\ldots,X_{n_J})}$, can be defined similarly to the single-input G_{X_n} and computed to give

$$G_{(X_{n_1},\ldots,X_{n_j},\ldots,X_{n_J})} \stackrel{\text{def}}{=} -\sigma_Y^{-2}\lim_{\rho\to 0}\frac{\sigma^2_{\widetilde{Y}_{(n_1,\ldots,n_j,\ldots,n_J)}(\rho)}-\sigma_Y^2}{\rho}$$
$$= \frac{\mathsf{E}\big((Y-\mu_Y)\sum_{j=1}^J\frac{\partial Y}{\partial X_{n_j}}(X_{n_j}-\mu_{X_{n_j}})\big)}{\sigma_Y^2},$$

and taking the expectations through the summations and rearranging gives

$$G_{(X_{n_1},\ldots,X_{n_j},\ldots,X_{n_J})} = \sum_{j=1}^J\frac{\mathsf{E}\big((Y-\mu_Y)\frac{\partial Y}{\partial X_{n_j}}(X_{n_j}-\mu_{X_{n_j}})\big)}{\sigma_Y^2}$$
$$= \sum_{j=1}^J G_{n_j}.$$

Table 2. An uncertainty budget with sensitivity measures.

Quantity[a]	Estimate	Standard uncertainty	VG[b]
X_1	x_1	$u(x_1)$	G_{X_1}
$\vdots$	$\vdots$	$\vdots$	$\vdots$
X_N	x_N	$u(x_N)$	G_{X_N}
Y	y	$u(y)$	$\sum_{n=1}^N G_{X_n}$

[a] A separate column may indicate the quantities' units.
[b] VG: variance gradient.

Thus, CVGs have a linearity property in that they are simple sums of the corresponding single-variable VGs. Note that the relative reduction in variance for each of the relevant inputs is assumed to be identical in the limit as this relative reduction goes to zero[8].

3.5. Application to uncertainty budgets

With regard to the GUM framework, VGs are readily incorporated into uncertainty budgets. As shown in table 2, VGs can be tabulated alongside the estimates and standard uncertainties of the inputs and measurand. Recall that a VG approximates the relative variance reduction in the measurand relative to a small relative reduction in the variance of a given input. This quantification enables a natural ordering/ranking of the inputs by value of the VGs (cf the factor prioritization setting). A VG can be negative, so the absolute value of the VG measures the size of the anticipated change while a positive (negative) sign of the VG indicates that the input variance should be decreased (increased) to decrease the measurand variance.

CVGs may be computed for groups of inputs via simple addition of the corresponding single input VGs. The last row in the fourth column of table 2 shows the sum of all the VGs, which quantifies the total relative reduction in the variance of the measurand that can be anticipated for a small relative reduction in the variance of all the inputs simultaneously. Recall that if the measurement function is linear, then $\sum_{n=1}^N G_{X_n} = 1$ and G_{X_n} are the squares of the σ-normalized derivatives, i.e. $G_{X_n} = \left(c_n\frac{\sigma_{X_n}}{\sigma_Y}\right)^2$ for all $n = 1,\ldots,N$, with $c_n = \frac{\partial Y}{\partial X_n}$ being constant.

3.6. Computation of VGs

A VG G_{X_n} can be estimated using (5) with a Monte Carlo method. First, an input sample from $\mathbf{X}$ is used to generate a sample from $Y = g(\mathbf{X})$. (Conveniently, this first step is already a standard procedure to estimate μ_Y and σ_Y^2 using the Monte Carlo method in GUM-S1 [3, section 7]. However, other methods to compute μ_Y and σ_Y^2 are permissible.) Second, this

[8] This technique can be extended by assigning relative (non-negative) weights to the variance reductions in each of the input variables, reflecting, say, the relative (inverse) costs for reducing the variance of each input. The resulting CVG is a linear combination of the single-variable VGs with corresponding non-negative weights. For example, if reducing the variance of X_2 were half as expensive as reducing the variance in X_1, so that the variance of X_2 is to be reduced twice as much as the variance of X_1, then $\widetilde{X}_1(\rho) = \sqrt{1-\rho}(X_1-\mu_{X_1})+\mu_{X_1}$ and $\widetilde{X}_2(\rho) = \sqrt{1-2\rho}\,(X_2-\mu_{X_2})+\mu_{X_2}$ give (in the limit as $\rho\to 0$) $G_{(X_1,X_2)} = G_{X_1}+2G_{X_2}$.

sample is used to generate samples from $Y - \mu_Y = g(\mathbf{X}) - \mu_Y$, $\frac{\partial Y}{\partial X_n} = \frac{\partial}{\partial X_n} g(\mathbf{X})$ and $X_n - \mu_{X_n}$, which are then used to estimate the expected value in (5). Finally, dividing by σ_Y^2 gives an estimate of G_{X_n}.

As discussed in [3], a sufficiently large sample size must be used when estimating μ_Y and σ_Y^2 with a Monte Carlo method. For a given input X_n, μ_{X_n} is typically known from its marginal distribution or a previous estimation, and in such cases need not to be estimated from the $\mathbf{X}$ sample. Importantly, any sampling uncertainty in the estimation of μ_{X_n}, μ_Y and σ_Y^2 propagates into a Monte Carlo estimation of G_{X_n}. Depending upon the problem at hand, additional sampling beyond that which was required to estimate μ_Y and σ_Y^2 may be required to estimate G_{X_n} sufficiently. As with the estimation of μ_Y and σ_Y^2, adaptive, bootstrap, or other procedures may be employed to ensure sufficient convergence of the Monte Carlo method [3, 6, 18, 19]. Various sampling strategies may accelerate convergence [20].

Because VGs are *global* sensitivity measures, the partial derivatives of the measurement function g in (5) must be evaluated over the entire support of the independent inputs. This exceeds the requirements of the GUM, in which partial derivatives need only be evaluated at the expected values of the inputs [2, sections 5.1.3 and 5.1.4]. For many measurement functions, a computer algebra system (CAS) [41, 42] can quickly find closed form expressions for partial derivatives, which reduces computational mistakes and the workload of metrologists. For complicated measurement functions such as those defined by a computer program, numerical partial differentiation of g may be required, which adds numerical uncertainty to the estimation of VGs. Applicable methods include finite-difference, complex-step, or automatic differentiation. Also, the numerical approximation of a partial derivative at a single input point may require multiple evaluations of g, adding to the overall computational expense. Guidance concerning algebraic/numerical partial differentiation in metrology applications is given in [34, 35, 43].

The Monte Carlo method is not the only possible method for computing the expected value in (5). For example, the expectation may be computable in a closed form or via numerical quadrature, possibly with the help of a CAS. If the expectation in (5) does not exist, then one may still attempt to compute the VG by resorting to the fundamental limit definition (4), assuming that this limit exists. Note that sampling uncertainty typically makes Monte Carlo methods poor at estimating the indeterminate form in the limit definition (4), i.e. by fixing a small nonzero value for ρ.

3.7. Application to variance cutting

Because VGs are based upon infinitesimally small relative changes in the input variances, one may question their applicability in a variance cutting setting in which one or more input variances must be substantially reduced in order to produce a measurement result with acceptable uncertainty. A proposed application of VGs to this setting would employ an iterative optimization scheme analogous to a steepest-decent algorithm. First, a target reduction in the variance of the measurand is established. Next, guided by the VGs of the inputs, the variances of inputs (one at a time or in groups) are iteratively changed in small amounts until the desired output variance is reached or no further significant variance reduction occurs. Note that the VGs are re-computed at each iteration and the expected values of the inputs are fixed in this basic scheme.

The above optimization scheme would benefit from automation in software and will not be developed further here. However, lack of automation does not preclude consideration of what–if scenarios in which VGs guide the manual adjustment of the inputs' variances in order to reduce the measurand variance below a given threshold. Thus, VGs can have both diagnostic and prognostic applications to the result of a measurement.

4. Examples

This section demonstrates the usefulness of VGs to measurement problems with nonlinear measurement functions and independent inputs. An illustrative metrological example is selected from the GUM-S1, for which VGs are compared with several other sensitivity measures. An additional benchmark problem example further distinguishes VGs from FOSIs and TSIs. Recall the notational conventions described in subsection 3.1.

4.1. GUM-S1 section 9.3 example

The example from GUM-S1 section 9.3 is analysed here [3, section 9.3]. In this example, a nominal mass is calibrated by balancing against a reference mass, with consideration of the buoyancy effects of air. The nonlinear measurement function for the deviation δm of the unknown mass from its nominal mass m_{nom} is ultimately given by

$$\delta m = (m_{\text{R,c}} + \delta m_{\text{R,c}})\left[1 + (\rho_{\text{a}} - \rho_{\text{a}_0})\left(\rho_{\text{W}}^{-1} - \rho_{\text{R}}^{-1}\right)\right] - m_{\text{nom}}, \tag{6}$$

where the conventional density of air, i.e. $\rho_{\text{a}_0} = 1.2\ \text{kg m}^{-3}$, and $m_{\text{nom}} = 100$ g are treated as exact constants. As described further in [3, section 9.3], (i) $m_{\text{R,c}}$ is the conventional mass of the reference mass, (ii) $\delta m_{\text{R,c}}$ is the conventional mass of the additional reference mass added to achieve balance with the unknown mass, (iii) ρ_{a} is the actual density of air, (iv) ρ_{W} is the density of the unknown mass and (v) ρ_{R} is the density of the reference mass. The distributions of the five input RVs and the associated first partial derivatives of the measurement function are summarized in table 3.

A CAS [42] was used to compute exactly the estimate (the expected value, $\mu_{\delta m}$) and standard uncertainty (from the variance, $\sigma_{\delta m}^2$) of the measurand δm, which agreed with the results from the Monte Carlo simulation in [3, section 9.3]. Using these two exact values in (5), 10^6 Monte Carlo simulations were used to compute the VGs appearing in the resulting uncertainty budget, given in table 4. Note that 10^6 Monte Carlo simulations are larger than the 0.72×10^6 simulations used in [3, section 9.3], and were verified to give

Table 3. Inputs for subsection 4.1 example.

Input RV ($*$)	Units	σ_*/μ_*	$\frac{\partial \delta m}{\partial *}$
$m_{\mathrm{R,c}} \sim \mathsf{N}(\mu_{m_{\mathrm{R,c}}} = 100\,000.000, \sigma^2_{m_{\mathrm{R,c}}} = (0.050)^2)$	mg	5.0×10^{-7}	$1 + (\rho_{\mathrm{a}} - \rho_{\mathrm{a_0}})\left(\rho_{\mathrm{W}}^{-1} - \rho_{\mathrm{R}}^{-1}\right)$
$\delta m_{\mathrm{R,c}} \sim \mathsf{N}(\mu_{\delta m_{\mathrm{R,c}}} = 1.234, \sigma^2_{\delta m_{\mathrm{R,c}}} = (0.020)^2)$	mg	0.016	$1 + (\rho_{\mathrm{a}} - \rho_{\mathrm{a_0}})\left(\rho_{\mathrm{W}}^{-1} - \rho_{\mathrm{R}}^{-1}\right)$
$\rho_{\mathrm{a}} \sim \mathsf{R}(a_{\rho_{\mathrm{a}}} = 1.10, b_{\rho_{\mathrm{a}}} = 1.30)$	kg m^{-3}	0.048	$(m_{\mathrm{R,c}} + \delta m_{\mathrm{R,c}})\left(\rho_{\mathrm{W}}^{-1} - \rho_{\mathrm{R}}^{-1}\right)$
$\rho_{\mathrm{W}} \sim \mathsf{R}(a_{\rho_{\mathrm{W}}} = 7 \times 10^3, b_{\rho_{\mathrm{W}}} = 9 \times 10^3)$	kg m^{-3}	0.072	$-(m_{\mathrm{R,c}} + \delta m_{\mathrm{R,c}})(\rho_{\mathrm{a}} - \rho_{\mathrm{a_0}})/\rho_{\mathrm{W}}^2$
$\rho_{\mathrm{R}} \sim \mathsf{R}(a_{\rho_{\mathrm{R}}} = 7.95 \times 10^3, b_{\rho_{\mathrm{R}}} = 8.05 \times 10^3)$	kg m^{-3}	0.0036	$(m_{\mathrm{R,c}} + \delta m_{\mathrm{R,c}})(\rho_{\mathrm{a}} - \rho_{\mathrm{a_0}})/\rho_{\mathrm{R}}^2$

Table 4. Uncertainty budget for subsection 4.1 example.

Quantity ($*$)	Units	Est.[a] (μ)	Std. Unc.[a] (σ)	VG[b] (G_*)	Linearized VG[c] (G_*^{L})
$m_{\mathrm{R,c}}$	mg	100 000.000	0.050	0.4	0.862
$\delta m_{\mathrm{R,c}}$	mg	1.234	0.020	0.07	0.138
ρ_{a}	kg m^{-3}	1.20	$\frac{0.10}{\sqrt{3}}$	0.5	0
ρ_{W}	kg m^{-3}	8×10^3	$\frac{1000}{\sqrt{3}}$	0.5	0
ρ_{R}	kg m^{-3}	8.00×10^3	$\frac{50}{\sqrt{3}}$	0.001	0
δm	mg	1.234	0.075 49	$\sum G_* = 1.5$	–
δm^{L}	mg	1.234	0.053 85	–	$\sum G_*^{\mathsf{L}} = 1.000$

[a] Input estimates and standard uncertainties considered to be exact values. Output estimates and standard uncertainties computed exactly with a CAS and reported to four significant digits.
[b] VGs computed to one significant digit using 10^6 Monte Carlo simulations of (5) for the measurement function (6).
[c] VGs for the linearized measurement function (7) computed exactly as squares of the σ-normalized derivatives equal to the coefficients of contribution of [33], and reported to three significant digits.

at least one significant digit [3, section 8]. Wherever possible, the Monte Carlo estimates of the VGs were verified against exact computations of the VGs with (5) using the CAS.

The VGs in the uncertainty budget given in table 4 provide a very good indication to metrologists as which inputs' variance should be reduced to produce the largest reduction in the measurand's variance. For example, the VGs tell metrologists to expect about a 0.5% reduction in the measurand variance if the variance of either ρ_{a} or ρ_{W} is reduced by 1%. To verify these indications, the variance of each of the inputs in table 3 was reduced by 10% (one input at a time with expectations unchanged), and the ratio of the relative reduction in the measurand's variance to the relative reduction in each input's variance was exactly computed. The resulting ratios agreed with the VGs reported in table 4 to at least one significant digit, with the exception of the input ρ_{R} with a small VG (ratio = 0.003 versus $G_{\rho_{\mathrm{R}}} = 0.001$)[9].

As shown in table 5, the VGs and TSIs for this example all agree to at least one significant digit. This result is interesting given that these measures do not share the same interpretation. The FOSIs reported here agree well with values reported in [6]. The TSIs reported here agree well with the strong two-way interaction effect between ρ_{a} and ρ_{W} ($S_{(\rho_{\mathrm{a}},\rho_{\mathrm{W}})} \approx 0.489$) and the weak two-way interaction effect between ρ_{a} and ρ_{R} ($S_{(\rho_{\mathrm{a}},\rho_{\mathrm{R}})} \approx 0.004$) reported in [6].

[9] Other measurement functions may require a variance reduction smaller than 10% to obtain such good agreement with the VGs.

The FOSIs (for the nonlinear measurement function) give a similar ranking as the VGs for the linearization of the measurement function about the input estimates given by

$$\delta m^{\mathsf{L}} = m_{\mathrm{R,c}} + \delta m_{\mathrm{R,c}} - m_{\mathrm{nom}}. \tag{7}$$

Note that the standard uncertainty of the measurand is underestimated by the linearized measurement function (7), and, as expected, the linearized VGs equal the squares of the respective σ-normalized derivatives and sum to 1. (Compare the last two rows of table 4.)

Table 5 also indicates that the FOSIs, linearized VGs, UCs and NLUCs do not adequately prioritize the inputs for reducing the measurand's variance. The linearized VGs, NLUCs and UCs in table 5 only discern between $m_{\mathrm{R,c}}$ and $\delta m_{\mathrm{R,c}}$, while indicating that the measurand's variance is totally insensitive to the remaining three inputs. In particular, $G_{\rho_{\mathrm{a}}} = G_{\rho_{\mathrm{W}}} = 0.5$, whereas $G^{\mathsf{L}}_{\rho_{\mathrm{a}}} = G^{\mathsf{L}}_{\rho_{\mathrm{W}}} = 0$. The FOSIs do only slightly better.

4.2. Ishigami function example

A common benchmark function for the sensitivity analysis, the Ishigami function, is analysed here [6]. The Ishigami function has three independent, identically distributed inputs, and is given by

$$Y = \sin(X_1) + 7\sin^2(X_2) + \frac{X_3^4}{10}\sin(X_1). \tag{8}$$

The distributions of the three input RVs and the associated first partial derivatives of the Ishigami function are summarized in table 6. Note the oscillatory terms in the function and that

Table 5. Comparison of sensitivity measures for subsection 4.1 example.

Quantity (∗)	VG[a] (G_*)	FOSI[b] (S_{T_*})	TSI[b] (S_*)	NLSC[c]	NLUC[c]	SC[d]	UC[d]	Linearized VG[e] (G_*^{L})
$m_{\mathrm{R,c}}$	0.4	0.439	0.439	1	0.050	1	0.050	0.862
$\delta m_{\mathrm{R,c}}$	0.07	0.0702	0.0702	1	0.020	1	0.020	0.138
ρ_{a}	0.5	0.002 51	0.491	0	0	0	0	0
ρ_{W}	0.5	0	0.487	0	0	0	0	0
ρ_{R}	0.001	0	0.001 19	0	0	0	0	0

[a] VGs computed to one significant digit using 10^6 Monte Carlo simulations of (6).
[b] FOSIs and TSIs computed exactly with a CAS and rounded to three significant digits.
[c] NLSCs derived from the NLUCs. NLUCs computed exactly by taking the square root of the variance of (6) with all inputs but one fixed at their expected values.
[d] UCs derived from the SCs and computed exactly for the linearized measurement function (7).
[e] VGs for the linearized measurement function (7), computed exactly as squares of the σ-normalized derivatives and rounded to three significant digits.

Table 6. Inputs for subsection 4.2 example.

Input RV (∗)	$\frac{\partial Y}{\partial *}$
$X_1 \sim \mathsf{R}(a_{X_1} = -\pi, b_{X_1} = \pi)$	$\left(1 + \frac{X_3^4}{10}\right) \cos(X_1)$
$X_2 \sim \mathsf{R}(a_{X_2} = -\pi, b_{X_2} = \pi)$	$14 \sin(X_2) \cos(X_2)$
$X_3 \sim \mathsf{R}(a_{X_3} = -\pi, b_{X_3} = \pi)$	$\frac{2}{5} X_3^3 \sin(X_1)$

the range of each input RV is apparently misstated in [6] as $(-\frac{\pi}{10}, \frac{\pi}{10})$.

A CAS [42] was used to compute exactly all results in the uncertainty budget given by table 7. The exact computation of the estimate and the standard uncertainty of Y was verified by 10^6 Monte Carlo simulations of (8). Using the exact values of μ_Y and σ_Y^2 in (5), the exact computation of the VG for each input was also verified by 10^6 Monte Carlo simulations. To further verify the computation of the VGs, the variance of each of the inputs in table 6 was reduced by 10% (one input at a time with expectations unchanged), and the ratio of the relative reduction in the measurand's variance to the relative reduction in each input's variance was exactly computed. The resulting ratios agreed with the VGs reported in table 7 to at least one significant digit. The FOSIs and TSIs reported here agree well with values reported in and inferred from [6].

Two results are particularly interesting in this benchmark problem. First is the presence of a *negative* VG for the first input X_1 and a VG *greater than 1* for the third input X_3. Second is that the VGs *change* the ranking of the inputs indicated by the TSIs. Thus, the TSIs (in addition to the FOSIs) may misdirect variance reduction effort. Such misdirection was verified by the relative reduction of the input variances by 10% (one input at a time with expectations unchanged), for which the VGs were sufficient indicators of the realized relative variance reduction/increase in the output Y. We note that a further decomposition of the output's variance into an apportionment among the inputs (using SIs) does not improve upon the FOSI/TSI result.

5. Discussion

In this concluding section, the use of VGs for uncertainty budgets in the GUM framework is summarized, and a comparison is made with other common sensitivity measures. Finally, some possible extensions and future work related to the application of VGs are briefly given.

5.1. VGs for uncertainty budgets in the GUM framework

As illustrated in the above examples, VGs are a useful quantitative tool for metrologists to prioritize which inputs' variances should be reduced in order to optimally reduce the variance of the measurand. VGs aid the construction of uncertainty budgets in settings with nonlinearizable measurement functions with independent inputs. VG sensitivity measures are *global* in the sense that they account for nonlinear effects over the entire support of the independent

Table 7. Uncertainty budget for subsection 4.2 example.

Quantity (∗)	Est.[a] (μ)	Std. Unc.[a] (σ)	VG[b] (G_*)	TSI[c] (S_{T_*})	FOSI[c] (S_*)
X_1	0	$\pi/\sqrt{3} \approx 1.814$	−0.2788	0.5576	0.3139
X_2	0	$\pi/\sqrt{3} \approx 1.814$	0.2212	0.4424	0.4424
X_3	0	$\pi/\sqrt{3} \approx 1.814$	1.8045	0.2437	0
Y	7/2	$\sqrt{\frac{33\,975+36\pi^4+\pi^8}{1800} - \frac{49}{4}} \approx 3.721$	$\sum G_* = 1.7469$	—	$\sum S_* = 0.7563$

[a] Input estimates and standard uncertainties considered to be exact values. Output estimates and standard uncertainties computed exactly with a CAS and rounded to three decimal places.
[b] VGs computed exactly with a CAS and rounded to four decimal places.
[c] TSIs and FOSIs computed exactly with a CAS and rounded to four decimal places.

inputs. This includes nonlinear effects within individual inputs as well as interaction effects between inputs.

Small variance reductions in the inputs can be considered individually, or in groups via simple addition of the corresponding VGs. For linear(ized) measurement functions, VGs reduce to a commonly understood variance decomposition, while providing an alternative interpretation to variance decomposition. Altogether, VGs appear to be useful for a reasonably large class of measurement functions with independent inputs. However, additional research is needed to more completely elucidate the class of measurement functions and input RVs for which VGs are guaranteed to be well-defined by (4) and computable using (5). Easily verifiable conditions under which the existence of the (finite) expectation in (5) guarantees the existence of the VG would be particularly useful.

Complementing the GUM-S1 methodology, a Monte Carlo method can be employed to compute VGs for differentiable measurement functions. This requires non-prohibitive computational expense of the measurement function and its first partial derivatives. Numerical differentiation adds uncertainty to the computation of VGs. Finally, the variance-based global sensitivity analysis using VGs first requires an uncertainty analysis that computes μ_Y and σ_Y^2 used in (4) or (5), which may require additional consideration of the propagation of sampling uncertainties into the computation of VGs.

5.2. Comparison with other sensitivity measures

VGs give quantitative information about how small relative changes in the variances of the inputs proportionally translate into relative changes in the output variance. This includes all higher order effects of the nonlinearity within individual inputs and from interaction effects between inputs. Unlike the variance-based SIs advanced by Saltelli *et al* [5], the definition of VGs does *not* imply that the collection of VGs sums to 1. Unlike FOSIs and TSIs, VGs need not sum to 1 for additive functions with independent inputs. In particular, even for the simplest type of (nonlinear) additive function, i.e. $Y = g(X_1)$, it is possible that $G_{X_1} > 1$ or $G_{X_1} < 1$ (including $G_{X_1} < 0$). However, if the function is linear, then $\sum_{n=1}^{N} G_{X_n} = 1$. Thus, the VGs introduced here should *not* be viewed generally as a decomposition of the output variance into an apportionment among the inputs.

As indicated above by the Ishigami benchmark function, the quantitative meaning of VGs makes them advantageous over SIs for variance reduction. The nesting of expectation operators in the computation of SIs also makes them computationally intensive, whereas the inner expectations in VGs, namely μ_Y and μ_{X_n}, are computed only once as a necessary part of the uncertainty analysis. However, differentiability of the measurement function and computation of its partial derivatives are typically required for VGs, while not required for SIs[10].

As shown here and elsewhere [6], UCs/SCs and NLUCs/NLSCs may be inadequate sensitivity measures for nonlinear measurement functions. However, FOSIs and TSIs (or other SIs) can optionally be included in uncertainty budgets in addition to VGs, because they give additional sensitivity information that is sometimes useful. For example, comparing respective FOSIs and TSIs gives a good indication of interaction effects and the additivity of the model [5].

A TSI considers the effect on the output's variance of a probabilistically weighted fixation of an input to any possible value in its state-of-knowledge probability distribution. Thus, a TSI near zero indicates that the input is globally non-influential [5]. A near-zero VG gives a more local indication about an input's influence, *per se*, which is relative to an infinitesimally small relative reduction of the input's variance with its expected value fixed (recall definition (3))[11]. Despite the differences between VGs and SIs, a common idea concerning *relative* variance reduction underlies both these sensitivity measures. This is illustrated by comparing (1) and (4).

Recall that a complete decomposition/apportionment of the output variance using the entire set of SIs may be computationally prohibitive. For computationally expensive measurement functions with many inputs, pre-screening inputs with elementary effect tests would reduce the computational burden of both SIs and VGs. The establishment of benchmark problems with verified solutions is advantageous in these and related matters. Such benchmarks should correspond to the problem setting at hand, which may correspond to a specific application area such as metrology. Other computational techniques such the Fourier amplitude sensitivity test (FAST) and random balanced design (RBD) [5] can be useful for benchmarking. These were not considered here because of their implementation complexity relative to the Monte Carlo method of GUM-S1.

5.3. Extensions for future work

In closing, we consider some extensions of the application of VGs. First, a direct extension of VGs to problems with jointly distributed inputs would be useful. This might be accomplished by infinitesimally reducing a variance-related parameter of a copula [17] or other representation of the joint distribution in a manner analogous to definition (3). Second, as mentioned in subsection 3.7, VGs enable an iterative scheme for an optimized reduction in the variance of the measurand. Such a scheme could account for relative ease/cost of reducing variances in the various inputs. Finally, effective measures/tests for the appropriateness of measurement function linearization are useful in applications of the GUM framework. We have shown that if the measurement function is linear, then $\sum_{n=1}^{N} G_{X_n} = 1$. Equivalently, if $\sum_{n=1}^{N} G_{X_n} \neq 1$, then the measurement function must be nonlinear. However, the use of VGs to diagnose 'degrees' of nonlinearity and the implications of $\sum_{n=1}^{N} G_{X_n} = 1$ require further investigation.

VGs are applicable to measurement functions with independent inputs whose underlying state-of-knowledge

[10] Consider the function $Y = g(X_1, X_2) = \sqrt{X_1^2 + X_2^2}$, where X_1 and X_2 are normal i.i.d., which gives Y a Rayleigh distribution. The derivative of this function does not exist at the origin, and to compute VGs, one may resort to the limit definition (4), with the associated computational difficulties.

[11] Occasionally, a near-zero VG (or TSI) indicates that a mistake was made somewhere, so that the analysis and computations should be re-checked.

probability distributions have finite variance. The inputs' uncertainty evaluation may be either type A or type B. In many type A evaluations, however, one attempts to reduce an input's variance by taking additional indications from a measurement process, or by reducing the variability in the measurement process itself. VGs may potentially be extended to these specific situations.

Because of the stochastic nature of measurement processes, one could consider the *expected* relative reduction in the output variance given one or more additional indications of an input. (Indeed, a rare event producing a single additional indication could increase an input's variance instead of decreasing it.) The *type* of input/output state-of-knowledge distributions may change with additional indication(s), such as different shifted and scaled t-distributions with changing degrees of freedom [2, 3]. The discrete nature of this setting is more analogous to finite differences than gradients/derivatives. Similar expectations could be computed for a fixed number of indications to be taken from a measurement process whose variance were reduced infinitesimally, giving yet another gradient-related variance-based sensitivity measure.

Acknowledgments

The primary author would like to thank his postdoctoral advisor at NIST, Sudarsan Rachuri, for introducing him to the GUM. The primary author would also like to thank Al Jones at NIST and the two anonymous referees for their valuable comments.

Appendix. Taking derivatives inside expectation operators

The derivation of formula (5) for computing VGs requires taking a derivative with respect to a parameter inside two different expectation operators. This appendix provides the theoretical justification for the stated sufficient conditions for the validity of (5) and illustrates how to construct dominating RVs that justify this operation for the Ishigami function from subsection 4.2.

The following theorem is adapted directly from results proved in [39, section 9.2] and [40, corollary 5.9], which are based upon the Lebesgue dominated convergence theorem (LDCT). The hypotheses of the following theorem provide the sufficient conditions stated in subsection 3.2 for the validity of (5).

Theorem 1. *$\{Y(z)\}_{z\in[a,b]}$ is a parametrized collection of RVs defined on some probability triple $(\Omega, \mathcal{F}, \mathbf{P})$. Suppose for some $z_0 \in [a,b]$ that $\mathsf{E}(Y(z_0)) < \infty$ and for all $z \in [a,b]$ that the derivative $\frac{\partial Y}{\partial z}(z)$ exists. Furthermore, suppose there exists a non-negative RV W defined on $(\Omega, \mathcal{F}, \mathbf{P})$ with finite expectation such that for all $z \in [a,b]$*

$$\left|\frac{\partial}{\partial z}Y(z)\right| \leqslant W.$$

It follows that $\frac{\partial}{\partial z}\mathsf{E}(Y(z)) = \mathsf{E}\big(\frac{\partial}{\partial z}Y(z)\big)$ for all $z \in [a,b]$.

We note that establishing uniform integrability [39, section 9.1] of difference quotients of RVs with respect to a parameter is a possible alternative to invoking the LDCT.

Recall from (8) that the Ishigami function is given by

$$Y = \sin(X_1) + 7\sin^2(X_2) + \frac{X_3^4}{10}\sin(X_1).$$

The distributions of the three input RVs and the associated first partial derivatives of the Ishigami function are summarized in table 6. As described in subsection 3.2, a sufficient condition for the validity of (5) for computing G_{X_n} for some $n = 1, 2, 3$ is the existence of dominating RVs, i.e. W_1 and W_2, each with finite expectation, such that

$$\left|\frac{\partial}{\partial z}\widetilde{Y}_n(z)\right| \leqslant W_1 \quad \text{and} \quad \left|\frac{\partial}{\partial z}\big(\widetilde{Y}_n(z)\big)^2\right| \leqslant W_2,$$

for all z in an open interval containing zero. By assumption, $\mathsf{E}(\widetilde{Y}_n(0)) = \mathsf{E}(Y) < \infty$ and $\mathsf{E}((\widetilde{Y}_n(0))^2) = \mathsf{V}(Y) + (\mathsf{E}(Y))^2 < \infty$, and we can take $z \in [-\varepsilon, \varepsilon]$ for some $0 < \varepsilon < 1$. Note that applying the chain rule for derivatives gives

$$\frac{\partial}{\partial z}\widetilde{Y}_n(z) = \left.\frac{\partial Y}{\partial X_n}\right|_{\widetilde{\mathbf{X}}_n(z)} \frac{\partial}{\partial z}\widetilde{X}_n(z)$$

and

$$\frac{\partial}{\partial z}\big(\widetilde{Y}_n(z)\big)^2 = 2\widetilde{Y}_n(z)\left.\frac{\partial Y}{\partial X_n}\right|_{\widetilde{\mathbf{X}}_n(z)} \frac{\partial}{\partial z}\widetilde{X}_n(z).$$

For G_{X_1}, note that

$$\widetilde{Y}_1(z) = \sin\big(\widetilde{X}_1(z)\big) + 7\sin^2(X_2) + \frac{X_3^4}{10}\sin(\widetilde{X}_1(z)),$$

and, because $\widetilde{X}_1(z) = \sqrt{1-z}(X_1 - \mu_{X_1}) + \mu_{X_1}$ and $\mu_{X_1} = 0$,

$$\frac{\partial}{\partial z}\widetilde{Y}_1(z) = \left(\cos(\widetilde{X}_1(z)) + \frac{X_3^4}{10}\cos(\widetilde{X}_1(z))\right)\frac{X_1}{-2\sqrt{1-z}}.$$

Using the triangle inequality, the dominating function W_1 is derived as follows:

$$\begin{aligned}\left|\frac{\partial}{\partial z}\widetilde{Y}_1(z)\right| &= \left|\cos(\widetilde{X}_1(z)) + \frac{X_3^4}{10}\cos(\widetilde{X}_1(z))\right|\left|\frac{X_1}{-2\sqrt{1-z}}\right| \\ &\leqslant |\cos(\widetilde{X}_1(z))| + \frac{X_3^4}{10}|\cos(\widetilde{X}_1(z))|\frac{|X_1|}{2\sqrt{1-\varepsilon}} \\ &\leqslant \left(1 + \frac{X_3^4}{10}\right)\frac{|X_1|}{2\sqrt{1-\varepsilon}} = W_1,\end{aligned}$$

which has finite expectation for the rectangular distributions given for X_1, X_2 and X_3. Likewise, the dominating function W_2 is derived as follows:

$$\begin{aligned}&\left|\frac{\partial}{\partial z}(\widetilde{Y}_1(z))^2\right| \\ &= 2\left|\sin(\widetilde{X}_1(z)) + 7\sin^2(X_2) + \frac{X_3^4}{10}\sin(\widetilde{X}_1(z))\right| \\ &\quad\times\left|\cos(\widetilde{X}_1(z)) + \frac{X_3^4}{10}\cos(\widetilde{X}_1(z))\right|\left|\frac{X_1}{-2\sqrt{1-z}}\right|\end{aligned}$$

$$\leqslant 2\left(|\sin(\widetilde{X}_1(z))| + 7\sin^2(X_2) + \frac{X_3^4}{10}|\sin(\widetilde{X}_1(z))|\right)$$
$$\times\left(|\cos(\widetilde{X}_1(z))| + \frac{X_3^4}{10}|\cos(\widetilde{X}_1(z))|\right)\frac{|X_1|}{2\sqrt{1-\varepsilon}}$$
$$\leqslant \left(8+\frac{X_3^4}{10}\right)\left(1+\frac{X_3^4}{10}\right)\frac{|X_1|}{\sqrt{1-\varepsilon}} = W_2,$$

which has finite expectation for the rectangular distributions given for X_1, X_2 and X_3.

For G_{X_2}, a similar argument shows that

$$\left|\frac{\partial}{\partial z}\widetilde{Y}_2(z)\right| \leqslant \frac{7\,|X_2|}{\sqrt{1-\varepsilon}} = W_1$$

and

$$\left|\frac{\partial}{\partial z}(\widetilde{Y}_2(z))^2\right| \leqslant \left(8+\frac{X_3^4}{10}\right)\frac{14|X_2|}{\sqrt{1-\varepsilon}} = W_2,$$

both of which have finite expectation.

Finally, for G_{X_3}, a similar argument shows that

$$\left|\frac{\partial}{\partial z}\widetilde{Y}_3(z)\right| \leqslant \frac{(1+\varepsilon)X_3^4}{5} = W_1$$

and

$$\left|\frac{\partial}{\partial z}(\widetilde{Y}_3(z))^2\right| \leqslant 2\left(8+\frac{(1+\varepsilon)^2X_3^4}{10}\right)\left(\frac{(1+\varepsilon)X_3^4}{5}\right) = W_2,$$

both of which have finite expectation.

References

[1] Joint Committee for Guides in Metrology (BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP and OIML) 2012 *International Vocabulary of Metrology—Basic and General Concepts and Associated Terms (VIM) (JCGM 200:2012, 2008 Version with Minor Corrections)* 3rd edn (Paris: International Bureau of Weights and Measures (BIPM)) (http://www.bipm.org/en/publications/guides/vim.html)

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

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

[4] Committee on Mathematical Foundations of Verification, Validation and Uncertainty Quantification 2012 *Assessing the Reliability of Complex Models: Mathematical and Statistical Foundations of Verification, Validation, and Uncertainty Quantification* (Washington, DC: National Research Council of the National Academy of Sciences)

[5] Saltelli A, Ratto M, Andres T, Campolongo F, Cariboni J, Gatelli D, Saisana M and Tarantola S 2008 *Global Sensitivity Analysis: The Primer* (New York: Wiley)

[6] Allard A and Fisher N 2009 Sensitivity analysis in metrology: study and comparison on different indices for measurement uncertainty *Advanced Mathematical and Computational Tools in Metrology and Testing: VIII* ed F Pavese, M Bär, A B Forbes, J M Linares, C Perruchet and N F Zhang (Singapore: World Scientific) pp 1–6

[7] Saltelli A, Tarantola S, Campolongo F and Ratto M 2004 *Sensitivity Analysis in Practice: A Guide to Assessing Scientific Models* (New York: Wiley)

[8] Joint Committee for Guides in Metrology (BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP and OIML) 2011 *Evaluation of Measurement Data—Supplement 2 to the 'Guide to the Expression of Uncertainty in Measurement'—Extension to Any Number of Output Quantities (JCGM 102:2011)* (Paris: International Bureau of Weights and Measures (BIPM)) (http://www.bipm.org/en/publications/guides/gum.html)

[9] Joint Committee for Guides in Metrology (BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP and OIML) 2009 *Evaluation of Measurement Data—An Introduction to the 'Guide to the Expression of Uncertainty in Measurement' and Related Documents (JCGM 104:2009)* (Paris: International Bureau of Weights and Measures (BIPM)) (http://www.bipm.org/en/publications/guides/gum.html)

[10] Kacker R, Sommer K-D and Kessel R 2007 Evolution of modern approaches to express uncertainty in measurement *Metrologia* **44** 513–29

[11] Bich W, Callegaro L and Pennecchi F 2006 Nonlinear models and best estimates in the GUM *Metrologia* **43** S196–9

[12] Callegaro L and Pennecchi F 2007 Why always seek the expected value? A discussion relating to the L_p norm *Metrologia* **44** L68–70

[13] Possolo A and Toman B 2007 Assessment of measurement uncertainty via observation equations *Metrologia* **44** 464–75

[14] Sommer K D and Siebert B R L 2006 Systematic approach to the modelling of measurements for uncertainty evaluation *Metrologia* **43** S200–10

[15] Cox M G and Seibert B R L 2006 The use of a Monte Carlo method for evaluating uncertainty and expanded uncertainty *Metrologia* **43** S178–88

[16] Possolo A, Toman B and Estler T 2009 Contribution to a conversation about the supplement 1 to the GUM *Metrologia* **46** L1–7

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

[18] Wübbeler G, Harris P M, Cox M G and Elster C 2010 A two-stage procedure for determining the number of trials in the application of a Monte Carlo method for uncertainty evaluation *Metrologia* **47** 317–24

[19] Robert C P and Casella G 2004 *Monte Carlo Statistical Methods* 2nd edn (New York: Springer)

[20] McKay M D, Beckman R J and Conover W J 1979 A comparison of three methods for selecting values of input variables in the analysis of output from a computer code *Technometrics* **21** 239–45

[21] Bich W 2008 How to revise the GUM? *Accred. Qual. Assur.* **13** 271–5

[22] Bich W *et al* 2012 Revision of the '*Guide to the Expression of Uncertainty in Measurement*' *Metrologia* **49** 702–5

[23] Elster C and Toman B 2009 Bayesian uncertainty analysis under prior ignorance of the measurand versus analysis using the supplement 1 to the *Guide*: a comparison *Metrologia* **46** 261–6

[24] Guthrie W F, Liu H-K, Rukhin A L, Toman B, Wang J C M and Zhang N F 2009 Three statistical paradigms for the assessment and interpretation of measurement uncertainty *Data Modeling for Metrology and Testing in Measurement Science* ed F Pavese and A B Forbes (Basel: Birkhäuser) pp 71–115

[25] Elster C, Wöger W and Cox M G 2007 Draft GUM supplement 1 and Bayesian analysis *Metrologia* **44** L31–2
[26] Lira I and Wöger W 2006 Comparison between the conventional and Bayesian approaches to evaluate measurement data *Metrologia* **43** S249–59
[27] Kacker R, Toman B and Huang D 2006 Comparison of ISO-GUM, draft GUM Supplement 1 and Bayesian statistics using simple linear calibration *Metrologia* **43** S167–77
[28] Kacker R and Jones A 2003 On use of Bayesian statistics to make the *Guide to the Expression of Uncertainty in Measurement* consistent *Metrologia* **40** 235–48
[29] Weise K and Wöger W 1993 A Bayesian theory of measurement uncertainty *Meas. Sci. Technol.* **4** 1–11
[30] Kessel R and Kacker R 2009 A test of linearity using covering arrays for evaluating uncertainty in measurement *Advanced Mathematical and Computational Tools in Metrology and Testing: VIII* ed F Pavese, M Bär, A B Forbes, J M Linares, C Perruchet and N F Zhang (Singapore: World Scientific) pp 195–203
[31] Wübbeler G, Krystek M and Elster C 2008 Evaluation of measurement uncertainty and its numerical calculation by a Monte Carlo method *Meas. Sci. Technol.* **19** 084009
[32] European Co-operation for Accreditation Expression of the uncertainty of measurement in calibration 1999 *Document* EA-4/02
[33] Kessel R, Kacker R and Berglund M 2006 Coefficient of contribution to the combined standard uncertainty *Metrologia* **43** S189–95
[34] Cox M G and Harris P M 2006 Software specifications for uncertainty evaluation *Technical Report* DEM-ES-010 (CMSC 40/04 revised) National Physical Laboratory, Teddington, UK (revised/corrected November 2008)
[35] Cox M G and Harris P M 2010 Software support for metrology best practice guide no. 6, uncertainty evaluation *Technical Report* MS 6, National Physical Laboratory, Teddington, UK
[36] Borgonovo E 2006 Measuring uncertainty importance: investigation and comparison of alternative approaches *Risk Anal.* **26** 1349–61
[37] Rosenblatt M 1952 Remarks on a multivariate transformation *Ann. Math. Stat.* **23** 470–2
[38] International Organization for Standardization (ISO) 2006 *Statistics—Vocabulary and symbols. Part 1: General Statistical Terms and Terms Used in Probability (ISO 3534-1:2006, Corrected Version 2007)* 2nd edn (Geneva: ISO)
[39] Rosenthal J S 2006 *A First Look at Rigorous Probability Theory* 2nd edn (Hackensack, NJ: World Scientific)
[40] Bartle R G 1966 *The Elements of Integration and Lebesgue Measure* (New York: Wiley) (Wiley Classics Library Edition Published in 1995)
[41] Stein W A *et al* 2012 *Sage Mathematics Software (Version 5.3)* The Sage Development Team (http://www.sagemath.org)
[42] Wolfram Research, Inc. 2012 Mathematica (Version 8.0.4.0) Wolfram Research, Inc., Champaign, IL, USA (http://www.wolfram.com/mathematica/)
[43] Boudjemaa R, Cox M G, Forbes A B and Harris P M 2004 Automatic differentiation and its application in metrology *Advanced Mathematical and Computational Tools in Metrology and Testing: VI* ed P Ciarlini, F Cox, M G Pavese and G B Rossi (Singapore: World Scientific) pp 170–9