Derivative based global sensitivity analysis and its entropic link

A Preprint

Jiannan Yang
School of Physics, Engineering and Technology
University of York
Heslington, York, YO10 5DD, UK
jiannan.yang@york.ac.uk

September 29, 2023

Abstract

Distribution-based global sensitivity analysis (GSA), such as variance-based and entropy-based approaches, can provide quantitative sensitivity information. However, they can be expensive to evaluate and are thus limited to low dimensional problems. Derivative-based GSA, on the other hand, require much fewer model evaluations. It is known that derivative-based GSA is closely linked to variance-based total sensitivity index, while its relationship with the entropy-based measure is unclear. To fill this gap, we introduce a log-derivative based functional to demonstrate that the entropy-based and derivative-based sensitivity measures are strongly connected. In particular, we give proofs that, similar to the case with variance-based GSA, there is an inequality relationship between entropy-based and derivative-based important measures. Both analytical and numerical verifications are provided. Examples show that the derivative-based methods give similar variable rankings as entropy-based index and can thus be potentially used as a proxy for both variance-based and entropy-based distribution-type GSA.

Keywords entropic sensitivity index; sensitivity inequality; conditional entropy; exponential entropy; Ishigami function

1 Introduction

The most widely adopted global sensitivity analysis (GSA) methods are derivative-based and distribution-based.

The most common distribution-based approach examines variability using the output variance. Variance-based methods, also called Sobol's indices, decompose the function output into a linear combinations of input and interaction of increasing dimensionality, and estimate contribution of each input factor to the variance of the output [1]. As only the 2nd order moments are considered, it was pointed out in [2] that the variance based sensitivity measure is not well suited for heavy tailed or multimodal distributions. Entropy is a measure of uncertainty similar to variance: higher entropy tends to indicate higher variance (for Gaussian, entropy is proportional to log variance). Nevertheless, it was shown in [2] that entropy-based methods and variance-based methods can sometimes produce significantly different results.

Both variance-based and entropy-baesd global sensitivity analysis (GSA) can provide quantitative contributions of each input variable to the output quantity of interest. However, the estimation of variance and entropy based sensitivity indices can become expensive in terms of number of model evaluations. This limits the the appliation of both variance-based and entropy-based methods to low dimensional problems.

In contrast, derivative-based methods are much more efficient as only the average of the functional gradients across the input space is needed. It is thus often used for screening of a large number of input variables. For example, the Morris' method [3] constructs a global sensitivity measure by computing a weighted mean of the finite difference approximation to the partial derivatives, and it requires only a few model evaluations.

It is thus interesting to find out how to use derivative-based methods to mitigate the dimensionality issues of variance and entropy based approaches.

Previous studies have found a link between the derivative-based and variance-based indices. In [4], a sensitivity measure $\mu*$ is proposed based on the absolute values of the partial derivatives. It is empirically demonstrated that for some practical problems $\mu*$ is similar to the variance based total index. In [5], Sobol and Kucherenko have proposed the so-called derivative-based global sensitivity measures (DGSM). This importance criterion is similar to the modified Morris measure, except that the squared partial derivatives are used instead of their absolute values. In addition, an inequality link between variance based global sensitivity indices and the DGSM is established in the case of uniform or Gaussian input variables.

This inequality between DGSM and variance-based GSA has been extended to input variables belonging to the large class of Boltzmann probability measures in [6]. A new sensitivity index, which is defined as a constant times the crude derivative-based sensitivity, is shown to be a maximal bound of the variance based total sensitivity index. Furthermore, in [7], the variance-based sensitivity indices are interpreted as difference-based measures, where the total sensitivity index is equivalent to taking a difference in the output when perturbing one of the parameters with the other parameters fixed. The similarity to partial derivatives helps to explain why the mean of absolute elementary effects from the Morris' method can be a good proxy for the total sensitivity index.

However, the relationship between derivative-based and entropy-based sensitivity indices is still unknown. In this paper, we will provide a novel proof that, for a monotonic function, entropy-based total sensitivity index is equivalent to the global sensitivity index based on log derivatives. More generally, we demonstrate that the exponential of the entropy-based sensitivity measure is bounded by the derivative-based indices, including the modified Morris' index and DGSM proposed in the literature. The missing link between the derivative-based and entropy-based GSA is thus established and that is the main contribution of this paper.

In this paper, we focus on entropy-based GSA and explore its link with derivative-based GSA. However, it should be noted that there are many other moment-indepdent sensitivity measures [8], which are often based on a distance metric to measure the discrepancy between the conditional and unconditional output probability density functions (PDFs). For example, sensitivity indices based on the modification of the input PDFs have been proposed in [9] for reliability sensitivity analysis, where the input perturbation is derived from minimizing the probability divergence under constraints. [10] proposed a moment independent δ-indicator that looks at the entire input/output distribution. The definition of δ-indicator examines the expected total shift between the conditional and unconditional output PDFs, where the shift is conditional on one or more of the random input variables. Recently, the Fisher Information Matrix has been proposed to examine the perturbation of the entire joint probability density function (jPDF) of the outputs, and is closely linked to the relative entropy between the jPDF of the outputs and its perturbation due to an infinitesimal variation of the input distributions [11, 12].

In what follows, we will first review both derivative-based and entropy-based global sensitivity measures in Section 2. In Section 3, the relationship between these two indices are established, where mathematical proofs are given for general multivariate functions. For the purpose of verification, demonstrating examples are given in Section 4 for both monotonic and general functions. Concluding remarks are given in Section 5.

2 Global sensitivity analysis

In this section, both derivative-based and entropy-based sensitivity measures will be briefly reviewed, for the purpose of establishing a link between them in the next section.

2.1 Derivative-based indices

Consider the function $y = g(\mathbf{x})$ where the function $g(\cdot)$ is differentiable. Functionals based on $\partial g/\partial x_i$ have been proposed to examine the global sensitivity with respect to the parameter x_i.

For example, the modified Morris sensitivity index μ^* [13] is an approximation of the functional:

$$\mu_i = \mathbb{E}_X \left[\left| \frac{\partial g}{\partial x_i} \right| \right] \tag{1}$$

where in [5], a similar functional called DGSM has been studied:

$$\nu_i = \mathbb{E}_X \left[\left| \frac{\partial g}{\partial x_i} \right|^2 \right] \tag{2}$$

Table 1: Variance-based indices vs. entropy-based indices

	Variance-based indices	Entropy-based indices
Main Effect	$\frac{V(Y) - \mathbb{E}[V(Y\|x_i)]}{V(Y)}$	$\frac{H(Y) - \mathbb{E}[H(Y\|x_i)]}{H(Y)}$
Total Effect	$\frac{\mathbb{E}[V(Y\|x_{\sim i})]}{V(Y)}$	$\frac{\mathbb{E}[H(Y\|x_{\sim i})]}{H(Y)}$

In addition, we will introduce the following functional based on log derivatives:

$$l_i = \mathbb{E}_X \left[\ln \left| \frac{\partial g}{\partial x_i} \right| \right] \tag{3}$$

to explore a link between entropy and derivatives based global sensitivity indices. Note that different from μ_i and ν_i, the log-derivative based index l_i can be negative.

Note that these there sensitivity indices are closely related as:

$$e^{l_i} \leq \mu_i \leq \sqrt{\nu_i} \tag{4}$$

where it is evident that $\mu_i \leq \sqrt{\nu_i}$ based on Cauchy-Schwarz inequality. In addition, we have $e^{l_i} \leq \mu_i$ using Jensen's inequality as the logrithmic function is a concave function.

2.2 Entropy-based indices

Entropy is a measure of the average uncertainty, or the degree of non-uniformality, represented by the distribution. Global sensitivity indices can be formulated using entropy, analogously to the variance based sensitivity indices [14]:

$$\eta_i = \frac{H(Y) - \mathbb{E}[H(Y|x_i)]}{H(Y)} = \frac{I(X_i, Y)}{H(Y)} \tag{5}$$

where $H(Y)$ is the entropy of Y, and $H(Y|X_i) = \mathbb{E}[H(Y|x_i)]$ is the expected conditional entropy of Y given X_i. $I(X_i, Y)$ is the mutual information which measures how much knowing X_i reduces uncertainty of Y or vice versa.

The index η_i in Eq 5 measures the excepted reduction in the entropy of of Y by fixing X_i. This can be regarded as the main effect contribution of X_i to the entropy of Y, in analogy to the main effect index from variance-based indices as seen in Table 1.

Similarly, an entropy-based total sensitivity index can be defined as [15]:

$$\eta_{Ti} = \frac{\mathbb{E}[H(Y|\mathbf{x}_{\sim i})]}{H(Y)} \tag{6}$$

where the average is calculated over all possible values of $\mathbf{X}_{\sim i}$. This measures the remaining entropy of Y if the true values of $\mathbf{X}_{\sim i}$ can be determined, in analogy to the total effect index from variance-based indices as seen in Table 1.

3 Link between derivative-based and entropy-based GSA

In this section, the relationship between the derivative-based and entropy-based sensitivity indices is explored. We will consider the function $y = g(\mathbf{x})$ where the function $g(\cdot)$ is differentiable. Let X be a continuous random variable with probability density function (PDF) given by $f_X(x)$. $Y = g(X)$ can be regarded as a transformed variable and the transformed PDF $f_Y(y)$ can then be found via the Jacobian matrix. We will start from a one-dimensional input variable and extend it to two-dimensional and multi-dimensional cases using the method of dummy variables.

3.1 one dimensional variable

If $y = g(x)$ has a unique inverse $x = g^{-1}(y)$, the conditional PDF of the output Y given the input X can be written as:

$$f_{Y|X}(x, y) = \delta(y - g(x)) = \frac{\delta(x - g^{-1}(y))}{|g'(g^{-1}(y))|} \tag{7}$$

where $\delta(\cdot)$ is Dirac's delta distribution. The marginal PDF of Y is thus given as:

$$f_Y(y) = \frac{f_X(g^{-1}(y))}{|g'(g^{-1}(y))|} = \frac{f_X(x)}{|g'(x)|} \tag{8}$$

The entropy of the continuous random variable Y is thus:

$$\begin{aligned} H(Y) &= -\int f_Y(y) \ln f_Y(y) dy \\ &= -\int f_X(x) \ln \frac{f_X(x)}{|g'(x)|} dx \\ &= H(X) + \int f_X(x) \ln |g'(x)| dx \end{aligned} \tag{9}$$

where $dy = g'(x)\, dx$ for the change of variable.

$$\mathbb{E}[\ln |g'(x)|] = H(Y) - H(X) \tag{10}$$

3.2 two dimensional variables

In the two-dimensional case, we will make use of the concept of dummy variables to utilise the Jacobian matrix. Let $y_1 = g_1(x_1, x_2) = g(\mathbf{x})$ and $y_2 = g_2(x_1, x_2) = x_2$ as a dummy variable. We assume that the functions $g_1(\cdot)$ and $g_2(\cdot)$ have unique inverses (this condition will be generalised in Section 3.4). Then the transformed joint PDF of $\mathbf{Y} = \{Y_1, Y_2\}$ is:

$$f(y_1, y_2) = \frac{f(x_1, x_2)}{|\det \mathbb{J}|} \tag{11}$$

where $\mathbb{J}$ is the Jacobian matrix:

$$\det \mathbb{J} = \begin{vmatrix} \frac{\partial g_1}{\partial x_1} & \frac{\partial g_1}{\partial x_2} \\ \frac{\partial g_2}{\partial x_1} & \frac{\partial g_2}{\partial x_2} \end{vmatrix} = \frac{\partial g_1}{\partial x_1} \tag{12}$$

The entropy of $\mathbf{Y}$ can then be found as:

$$\begin{aligned} H(Y_1, Y_2) &= -\int f_Y(y_1, y_2) \ln f_Y(y_1, y_2) dy_1 dy_2 \\ &= -\int f_X(x_1, x_2) \ln \frac{f_X(x_1, x_2)}{|g_1'(x_1)|} dx_1 dx_2 \\ &= H(X_1, X_2) + \int f_X(x_1, x_2) \ln |g_1'(x_1)| dx_1 dx_2 \end{aligned} \tag{13}$$

Assuming the input variables are independent, i.e., $H(X_1, X_2) = H(X_1) + H(X_2)$, and note that $y1 = y$ and $y2 = x2$ is a dummy variable:

$$H(Y|X_2) = H(X_1) + \mathbb{E}\left[\ln \left|\frac{\partial g(\mathbf{x})}{\partial x_1}\right|\right] \tag{14}$$

where the expectation is with respect to the jPDF of the input variables. We have used the property of the conditional entropy, where $H(Y, X_2) = H(Y|X_2) + H(X_2)$. Note that $H(Y|X_2) = \mathbb{E}\left[H(Y|x_2)\right]$ is the expected conditional entropy.

3.3 multi-dimensional variables

Let $y_1 = g_1(\mathbf{x})$ where $\mathbf{x} = (x_1, x_2, \ldots, x_n)$, and introduce dummy variables $y_i = g_i(\mathbf{x}) = x_i$ with $i = 2, \ldots, n$. Assuming the functions $g_i(\cdot)$ have unique inverses, the transformed jPDF is:

$$f_Y(\mathbf{y}) = \frac{f(\mathbf{x})}{|\det \mathbb{J}|} \tag{15}$$

where $\mathbb{J}$ is the Jacobian matrix with $\mathbb{J}_{ij} = \partial g_i / \partial x_j$.

Note that for the Jacobian matrix from the 2nd row onwards, i.e., $i \geq 2$, $\partial g_i/\partial x_j = 1$ when $i = j$ and $\partial g_i/\partial x_j = 0$ when $i \neq j$. Therefore, the Jacobian matrix in this case is a triangular matrix. As a result, the Jacobian determinant is the product of the diagonal entries:

$$\det \mathbb{J} = \left| \frac{\partial g_1}{\partial x_1} \times \underbrace{1 \times \cdots \times 1}_{2 \text{ to } n} \right| = \left| \frac{\partial g_1}{\partial x_1} \right| \tag{16}$$

The entropy of $\mathbf{Y}$ is thus:

$$\begin{aligned} H(\mathbf{Y}) &= -\int f_Y(\mathbf{y}) \ln f_Y(\mathbf{y}) d\mathbf{y} \\ &= -\int f_X(\mathbf{x}) \ln \frac{f_X(\mathbf{x})}{\left|\frac{\partial g_1(\mathbf{x})}{\partial x_1}\right|} d\mathbf{x} \\ &= H(\mathbf{X}) + \int f_X(\mathbf{x}) \ln \left| \frac{\partial g_1(\mathbf{x})}{\partial x_1} \right| d\mathbf{x} \end{aligned} \tag{17}$$

where $\mathbf{Y} = \{Y_1, X_2, X_3, \ldots, X_n\}$.

Note that $H(\mathbf{Y}) = H(Y_1|\mathbf{X}_{2\sim n}) + H(\mathbf{X}_{2\sim n})$, and assuming the input variables are independent, we then have:

$$H(Y|\mathbf{X}_{2\sim n}) = H(X_1) + \mathbb{E}\left[\ln \left| \frac{\partial g(\mathbf{x})}{\partial x_1} \right| \right] \tag{18}$$

where the expectation is with respect to the jPDF of the input variables. Note that $H(Y|\mathbf{X}_{2\sim n}) = \mathbb{E}\left[H(Y|\mathbf{x}_{2\sim n})\right]$ is the expected conditional entropy.

The reasoning above uses the first variable x_1 as an example. However, the results hold for any variables via simple row/column exchanges, which only affects the sign of the determinant but not its modulus.

3.4 summary

In summary, for a differentiable function $y = g(\mathbf{x})$, where $\mathbf{x} = (x_1, x_2, \ldots, x_n)$, there exists a relationship between the entropy and the expectation of the partial derivatives:

$$\kappa_i = \mathbb{E}\left[H(Y|\mathbf{x}_{\sim i})\right] - H(X_i) = \mathbb{E}\left[\ln \left| \frac{\partial g(\mathbf{x})}{\partial x_i} \right| \right] \tag{19}$$

where $\sim i$ indicates the index ranges from 1 to n excluding i. Note that $H(Y|\mathbf{X}_{\sim i}) = \mathbb{E}\left[H(Y|\mathbf{x}_{\sim i})\right]$, where the expectation is with respect to all possible values of $\mathbf{X}_{\sim i}$.

The above relationship holds only if the transformation $g(\cdot)$ has a unique inverse. For a general differentiable function, we have the following inequality instead [16]:

$$\kappa_i \leq l_i \tag{20}$$

where l_i is used to denote $\mathbb{E}\left[\ln \left| \frac{\partial g(\mathbf{x})}{\partial x_i} \right| \right]$ as in Eq 3.

As the exponential function is a monotonic increasing function, and taking account of the relationship in Eq 4, we can then relate entropy to the derivative based global sensitivity indices as:

$$e^{\kappa_i} \leq e^{l_i} \leq \mu_i \leq \sqrt{\nu_i} \tag{21}$$

where l_i, μ_i and ν_i are derivative-based sensitivity indices as given in Section 2.1. For example, μ_i is used to denote $\mathbb{E}\left[|\partial g(\mathbf{x})/\partial x_i|\right]$, which can be regarded as the limiting value of the modified Morris' global sensitivity index [5] as given in Eq 1, while ν_i is the DGSM sensitivity index proposed in [5]. Equation 21 thus establises an inequality relationship between exponential entropy and the global sensitivity measures based on partial derivatives.

Note that it is the exponential entropy that has a direct relationship with the derivative-based global sensitivity measures. Recall that the entropy of a random variable with a uniform distribution is $\ln(b - a)$, where a, b are the bounds of the distribution. Taking the natural exponential of the entropy in this case results in $b - a$, which is the range of the uniform distribution. Therefore, the exponential entropy e^H can be regarded as a measure of the extent, or effective support, of a distribution [4]. As $H(Y|\mathbf{X}_{\sim i})$ measures the remaining entropy in average if the true values of $\mathbf{X}_{\sim i}$ can be determined, the exponential entropy based index e^{κ_i} thus indicates the ratio between the range of output distribution, conditioning on $\mathbf{X}_{\sim i}$ are known, and the range of input X_i. And this ratio is upper bounded by derivative-based sensitivity indices, including the modified Morris' sensitivity index μ_i.

Studies in [2] already noted that an exponetial transformation of the standard entropy-based sensitivity measures may improve its discrimination power. In addition, as pointed out in [15], entropy for continuous random variables (aka differential entropy) can become negative. This is a drawback of entropy-based sensitivity method. In contrast, e^{κ_i} is based on exponential entropy and thus always positive.

For the purpose of screening and identifying non-influential inputs, the inputs are typically assumed to have the same level of uncertainties, i.e., $H(X_i) = H(X_j)$. As the output entropy $H(Y)$ does not affect the relative values, the new index e^{κ_i} is essentially the same as the entropy-based total index η_i given in 1.

4 Examples

The mathematical relationship proved in Section 3, between derivative-based and entropy-based GSA, is numerically verified in this section. The equality between κ_i and l_i in Eq 19 is first examined for monotonic functions, and the Ishigami function and G-function are then used to demonstrate the inequality in Eq 21 for general functions.

All input variables are assumed to have the same uniform distribution for each function, while Gaussian distributions are used for example 5. Having the same distributions for all inputs, on the one hand, is for simplicity of the demonstrating examples; on the other hand, is a commonly adopted approach for input variable screenings, which is the main function of the entropy-based total sensitivity index.

4.1 Monotonic functions

For verification purposes, all the examples in this section are chosen to have tractable expressions for both the integral of derivatives and the conditional entropies. For examples $1-3$, the conditional entropies are also numerically estimated using the method given in Appendix A. This is to demonstrate that numerical estimation of entropy-based sensitivity indices can be very expensive. And that is the main motivation to establish a link with the derivative-based method, so that it can be used as a potential proxy for entropy-based GSA.

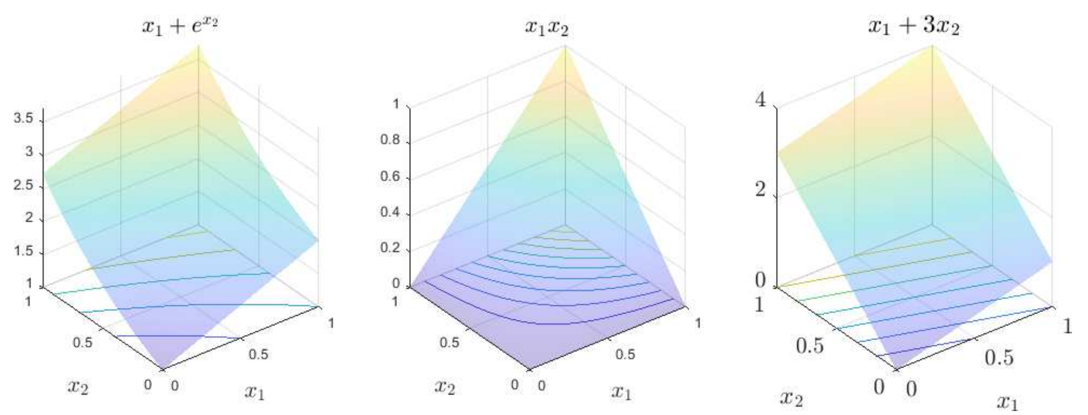

Figure 1: Surface plots, with contours shown underneath, for the monotonic functions in examples 1-3.

Example 1. Consider the function $y = x_1 + e^{x_2}$, where the partial derivatives are $\partial y/\partial x_1 = 1$ and $\partial y/\partial x_2 = e^{x_2}$. For $x_i \sim \mathbb{U}(0, 1)$, the expected value $l_i = \mathbb{E}\left[\ln\left|\frac{\partial g(\mathbf{x})}{\partial x_i}\right|\right]$ can be integrated analytically as 0 and 1/2 for x_1 and x_2 respectively. The sensitivities indices based on conditional entropy can also be calculated analytically in this case. $\kappa_1 = H(Y|X_2) - H(X_1) = 0$, where $H(Y|x_2) = H(X_1) = 0$ as the differential entropy remains constant under addition of a constant (e^{x_2} is a constant for $H(Y|x_2)$). As the transformed variable $V = e^x$ has a PDF $V \sim 1/v, 1 < v < e$, $\kappa_2 = H(Y|X_1) - H(X_2) = \mathbb{E}_{X_1}[H(Y|x_1)] - 0 = \mathbb{E}_{X_1}\left[-\int_1^e \frac{1}{y}\ln\frac{1}{y}\right] = 1/2$. Therefore, $\kappa_1 = l_1$ and $\kappa_2 = l_2$, as given by Eq 19.

Example 2. Consider the function $y = x_1 \times x_2$, where the partial derivatives are $\partial y/\partial x_1 = x_2$ and $\partial y/\partial x_2 = x_1$. For $x_i \sim \mathbb{U}(0, 1)$, the expected value $\mathbb{E}\left[\ln\left|\frac{\partial g(\mathbf{x})}{\partial x_i}\right|\right]$ are -1 for both x_1 and x_2. Recall that the differential entropy

Table 2: Sensitivity results for the monotonic functions in examples 1-3 (surface plots in Figure 1). Conditional entropy results are obtained from Monte Carlo sampling with number of samples ranging from 10^3 to 10^8. The results from 10^8 samples are compared to the exact results and the relative error are less than 1% for all functions. Also given are the corresponding analytical results for derivative-based indices.

Number of Samples $x_i \sim \mathbb{U}(0,1)$	$y = x_1 + e^{x_2}$		$y = x_1 \times x_2$		$y = x_1 + 3x_2$	
	κ_1	κ_2	κ_1	κ_2	κ_1	κ_2
1.00E+03	0.0934	0.5251	-0.9057	-0.8468	0.2279	1.1269
1.00E+04	0.0602	0.5225	-0.9111	-0.9211	0.1043	1.1093
1.00E+05	0.0315	0.5131	-0.9528	-0.9593	0.0526	1.1047
1.00E+06	0.0148	0.5069	-0.9760	-0.9755	0.0242	1.1019
1.00E+07	0.0068	0.5034	-0.9879	-0.9878	0.0113	1.1000
1.00E+08	0.0032	0.5015	-0.9939	-0.9939	0.0052	1.0993
Exact results	0.0000	0.5000	-1.0000	-1.0000	0.0000	1.0986
error	-	-0.30%	0.61%	0.61%	-	-0.06%
	l_1	l_2	l_1	l_2	l_1	l_2
$\mathbb{E}\left[\ln\left\|\frac{\partial y}{\partial x_i}\right\|\right]$	0.0000	0.5000	-1.0000	-1.0000	0.0000	1.0986

increases additively upon multiplication with a constant, $\kappa_1 = H(Y|X_2) - H(X_1) = \mathbb{E}_{X_2}[H(x_1) + \ln|x_2|] = \int_0^1 \ln x_2 dx_2 = -1$. Therefore, $\kappa_i = l_i, i = 1,2$ as given by Eq 19.

Example 3. Consider the function $y = x_1 + 3x_2$, where the partial derivatives are 1 and 3 for x_1 and x_2 respectively. For $x_i \sim \mathbb{U}(0,1)$, the expected value $\mathbb{E}\left[\ln\left|\frac{\partial g(\mathbf{x})}{\partial x_i}\right|\right]$ can be integrated analytically as 0 and $\ln 3 \simeq 1.0986$. It is straightforward to show that, as in previous examples, $\kappa_1 = 0$ and $\kappa_2 = \ln 3$, which are the same as l_1 and l_2.

Example 4. Consider the product function $y = x_1 x_2^r$, where $r \geq 1$ (for $r = 1$, we recover the function in example 2). For $x_1, x_2 \sim \mathbb{U}(0,1)$, it is straightforward to show that $s_1 = -r$ and $s_2 = \ln r - r$. For entropy based indices, we have $\kappa_1 = H(Y|X_2) - H(X_1) = \mathbb{E}_{X_2}[H(x_1) + \ln|x_2|] = \int_0^1 \ln x_2^r dx_2 = -r$, which is the same as l_1. $\kappa_2 = H(Y|X_1) - H(X_2) = \mathbb{E}_{X_1}[H(Y|x_1) + \ln|x_1|] - 0 = \mathbb{E}_{X_1}\left[-\int_0^1 p(y|x_1)\ln p(y|x_1)\right] - 1 = \ln r - r$, where $p(y|x_1) = \frac{1}{r}y^{\frac{1}{r}-1}$ as the transformed variable $V = x^r$ has a PDF $V \sim \frac{1}{r}v^{\frac{1}{r}-1}, 0 < v < 1$. Therefore, $\kappa_1 = l_1$ and $\kappa_2 = l_2$ as shown by Eq 19.

Example 5. Consider the linear function $y = \sum_{i=1}^n a_i x_i$. This function has been used in [14] to demonstrate the equivalence between entropy based and variance based sensitivity indices for Guassian random inputs. In the case with independent inputs, the sensitivity index based on the conditional entropy can be obtained as $\kappa_i = H(Y|\mathbf{X}_{\sim i}) - H(X_i) = 1/2\ln 2\pi e a_i^2\sigma_i^2 - 1/2\ln 2\pi e\sigma_i^2 = \ln|a_i|$, where σ_i^2 is the variance for the random input x_i. The results are the same as the derivative based index $l_i = \ln|\partial y/\partial x_i| = \ln|a_i|$, which demonstrates the equality given in Eq 19.

4.2 General functions

Both Ishigami function and G-function are commonly used test functions for global sensitivity analysis, due to the presence of strong interactions. These two functions, each with three input variables, are used in this section to demostrate the inequality relationship derived in Eq 21.

The conditional entropies, for the estimation of the sensitivity index κ_i, are estimated numerically using Monte Carlo sampling as described in Appendix A. Different number of samples are used, ranging from $1e6$ to $1e8$. For each estimation, the computation is repeated for 20 times and both the mean value and the standard deviation (std) are reported in Table 3 and 4.

For the estimation of derivative-based indices, l_i and μ_i, Matlab's inbuild numerical integrator "integral" is used with default tolerance setting. Note that the index ν_i is not compared here, as its relationship with μ_i has been studied elsewhere [5].

To rank the input variables, sensitivity indices in this section are normalised. For example, $e^{\kappa_i} / \sum e^{\kappa_i}$ is used for the entropy-based sensitivity results. As a result, the rankings in Table 3 and 4 are shown as percentages.

Example 6. The Ishigami function, $y = \sin(x_1) + a\sin^2(x_2) + bx_3^4 \sin(x_1)$, is often used as an example for uncertainty and sensitivity analysis. It exhibits strong nonlinearity and nonmonotonicity, as can be seen in Figure 2. In this case, $a = 7$ and $b = 0.1$ are used, and the input random variables have uniform distributions, i.e., $x_i \sim \mathbb{U}(-\pi, \pi)$ for $i = 1, 2, 3$.

The sensitivity results are listed in Table 3. The index κ_i is estimated with different number of samples, and each estimation with 20 repetitiions for variability assurance. It is clear from the small std that the estimation is well converged. The index e^{κ_i} is then estimated using the mean value from the results with $1e8$ samples, and compared with derivative-based indices e^{l_i} and μ_i.

First, from Table 3, it is clear that the inequality among e^{κ_i}, e^{l_i} and μ_i from Eq 21 is satisfied. Second, the ranking is similar from these three indices, with x_2 being the most important, thus suggesting using the derivative-based indices as a proxy for entropy-based importance measure. Third, we note in passing that the quantitative ranking based on e^{κ_i} is very close to the results reported in [2]. This is different from the variance-based total sensitivity results, as compared in [2], which ranks x_1 being the most important.

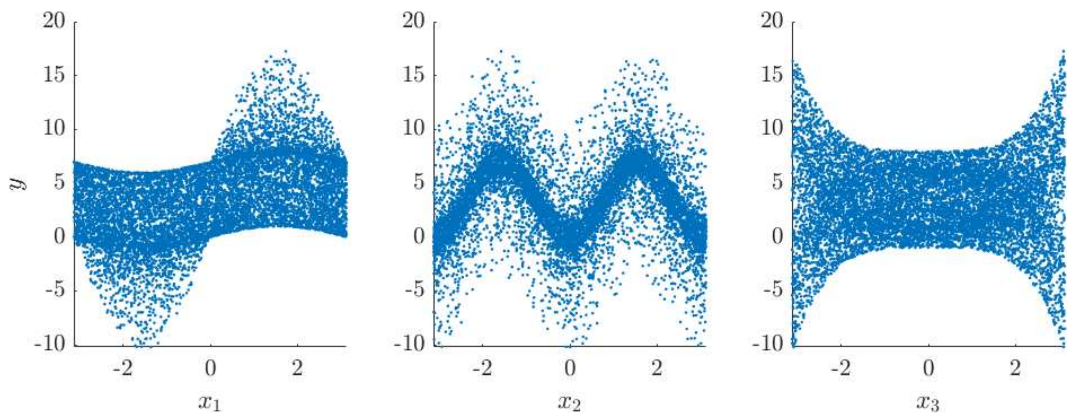

Figure 2: Scatter plot for the Ishigami function, $y = \sin(x_1) + a\sin^2(x_2) + bx_3^4 \sin(x_1)$.

Example 7. Consider the so-called G-function, $y = \prod_{i=1}^{3}(|4x_i - 2| + a_i)/(1 + a_i)$, which is often used for numerical experiments in sensitivity analysis. It is a highly nonlinear function, as can be seen in Figure 3 for a two-variable example. In this case, $a_i = (i - 2)/2$, for $i = 1, 2, 3$. The input random variables have uniform distributions, i.e., $x_i \sim \mathbb{U}(0, 1)$ for $i = 1, 2, 3$. A lower value of a_i indicates a higher importance of the input variable x_i, i.e., x_1 is the most important, while x_3 is the least important in this case.

The sensitivity results for the G-function, in the same format as Table 3, are reported in Table 4. It is clear that, not only the inequality relationship in Eq 21 is satisfied, the three indices give the same quantitative ranking of the three variables. Note that e^{l_i} and μ_i produce the same resutls, as the effect of logrithm and exponential operations cancels out for this product function.

5 Conclusions

- A link between derivative-based and entropy-based global sensitivity measures has been establised.
- For monotonic functions, mathematical proofs are given for the equality between the expected log-derivative sensitivities and entropy based global measures.
- For general functions, the exponential of the entropy-based sensitivity measures is found to be upper bounded by derivative-based global sensitivity indices.

Table 3: Sensitivity results for the Ishigami function. The conditional entropy based sensitivity results are obtained for different number of samples. This is repeated for 20 times and the mean and standard deviation (std) are given. Also shown are the three indices e^{κ_i}, e^{l_i} and μ_i, for which the inequality in Eq 21 is clearly satisfied.

	Ishigami function $y = \sin(x_1) + a\sin^2(x_2) + bx_3^4\sin(x_1)$					
Number of Samples	κ_1 mean	κ_1 std	κ_2 mean	κ_2 std	κ_3 mean	κ_3 std
1.00E+06	-0.4474	7.40E-04	-0.0764	1.07E-03	-0.8669	1.61E-03
3.16E+06	-0.4994	2.87E-04	-0.1079	3.61E-04	-0.9726	5.24E-04
1.00E+07	-0.5400	2.40E-04	-0.1356	1.57E-04	-1.0685	4.06E-04
3.16E+07	-0.5756	1.97E-04	-0.1582	1.17E-04	-1.1543	3.24E-04
1.00E+08	-0.6043	9.77E-05	-0.1770	8.56E-05	-1.2313	1.69E-04
	x_1 sensitivity	x_1 ranking	x_2 sensitivity	x_2 ranking	x_3 sensitivity	x_3 ranking
e^{κ_i}	0.5464	32.6%	0.8378	50.0%	0.2919	17.4%
e^{l_i}	1.0667	21.9%	3.5000	71.8%	0.3087	6.3%
μ_i	1.8769	22.6%	4.4563	53.6%	1.9739	23.8%

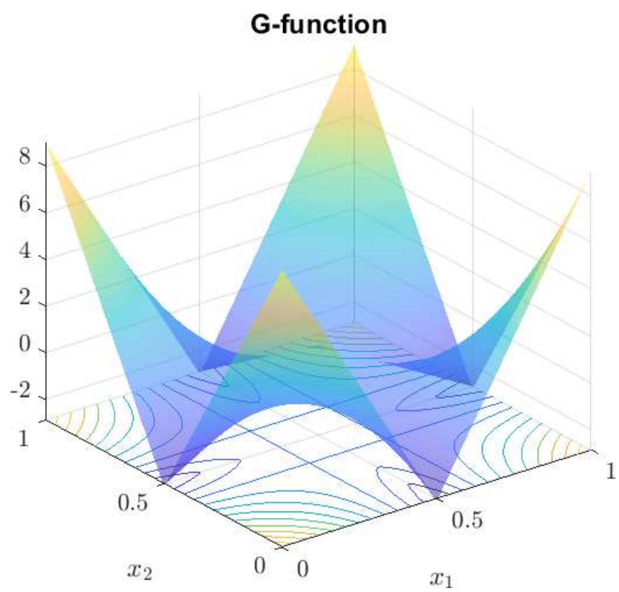

Figure 3: Surface plots, with contours shown underneath, for an example of G-function with 2 variables.

- Numerical examples show that the derivative-based methods give similar variable rankings as entropy-based index and can thus be potentially used as a proxy for entropy-based GSA.

Acknowledgment

For the purpose of open access, the author has applied a Creative Commons Attribution (CC BY) licence to any Author Accepted Manuscript version arising.

Data availability statement

The datasets generated during and/or analysed during the current study are available in the GitHub repository: `XXX`

Table 4: Sensitivity results for the G-function, where the inequality in Eq 21 is clearly satisfied. Same key as Table 3

	G-function $y = \prod_{i=1}^{3} \frac{\lvert 4x_i - 2\rvert + a_i}{1 + a_i}$					
Number of Samples	κ_1 mean	κ_1 std	κ_2 mean	κ_2 std	κ_3 mean	κ_3 std
1.00E+06	0.3473	1.40E-03	-0.1383	1.48E-03	-0.3993	1.55E-03
3.16E+06	0.3431	9.71E-04	-0.1584	8.19E-04	-0.4273	7.72E-04
1.00E+07	0.3401	6.17E-04	-0.1735	5.41E-04	-0.4480	4.77E-04
3.16E+07	0.3384	4.66E-04	-0.1841	3.75E-04	-0.4630	3.20E-04
1.00E+08	0.3377	3.58E-04	-0.1917	2.67E-04	-0.4739	2.34E-04
	x_1 sensitivity	x_1 ranking	x_2 sensitivity	x_2 ranking	x_3 sensitivity	x_3 ranking
e^{κ_i}	1.4018	49.2%	0.8255	29.0%	0.6226	21.8%
e^{l_i}	4.0000	46.2%	2.6667	30.8%	2.0000	23.1%
μ_i	4.0000	46.2%	2.6667	30.8%	2.0000	23.1%

References

[1] SALTELLI, A.: *Global sensitivity analysis: the primer*. John Wiley, 2008 `http://books.google.at/books?id=wAssmt2vumgC`. – ISBN 9780470059975

[2] AUDER, B ; IOOSS, B: Global sensitivity analysis based on entropy. In: *Proceedings of the ESREL 2008 Conference* (2008), S. 2107–2115

[3] MORRIS, Max D.: Factorial Sampling Plans for Preliminary Computational Experiments. In: *Technometrics* 33 (1991), Nr. 2, 161-174. `http://dx.doi.org/10.1080/00401706.1991.10484804`. – DOI 10.1080/00401706.1991.10484804

[4] CAMPBELL, L L.: Exponential entropy as a measure of extent of a distribution. In: *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* 5 (1966), Nr. 3, S. 217–225

[5] SOBOL', I.M. ; KUCHERENKO, S.: Derivative based global sensitivity measures and their link with global sensitivity indices. In: *Mathematics and Computers in Simulation* 79 (2009), Nr. 10, S. 3009–3017

[6] LAMBONI, M. ; IOOSS, B. ; POPELIN, A.-L. ; GAMBOA, F.: Derivative-based global sensitivity measures: General links with Sobol' indices and numerical tests. In: *Mathematics and Computers in Simulation* 87 (2013), 45-54. `http://dx.doi.org/https://doi.org/10.1016/j.matcom.2013.02.002`. – DOI https://doi.org/10.1016/j.matcom.2013.02.002. – ISSN 0378–4754

[7] WAINWRIGHT, Haruko M. ; FINSTERLE, Stefan ; JUNG, Yoojin ; ZHOU, Quanlin ; BIRKHOLZER, Jens T.: Making sense of global sensitivity analyses. In: *Computers & Geosciences* 65 (2014), 84-94. `http://dx.doi.org/https://doi.org/10.1016/j.cageo.2013.06.006`. – DOI https://doi.org/10.1016/j.cageo.2013.06.006. – ISSN 0098–3004. – TOUGH Symposium 2012

[8] BORGONOVO, Emanuele ; PLISCHKE, Elmar: Sensitivity analysis: A review of recent advances. In: *European Journal of Operational Research* 248 (2016), Nr. 3, 869-887. `http://dx.doi.org/https://doi.org/10.1016/j.ejor.2015.06.032`. – DOI https://doi.org/10.1016/j.ejor.2015.06.032. – ISSN 0377–2217

[9] LEMAÎTRE, P. ; SERGIENKO, E. ; ARNAUD, A. ; BOUSQUET, N. ; GAMBOA, F. ; IOOSS, B.: Density modification-based reliability sensitivity analysis. In: *Journal of Statistical Computation and Simulation* 85 (2015), Nr. 6, 1200-1223. `http://dx.doi.org/10.1080/00949655.2013.873039`. – DOI 10.1080/00949655.2013.873039

[10] BORGONOVO, E.: A new uncertainty importance measure. In: *Reliability Engineering & System Safety* 92 (2007), Nr. 6, 771-784. `http://dx.doi.org/https://doi.org/10.1016/j.ress.2006.04.015`. – DOI https://doi.org/10.1016/j.ress.2006.04.015. – ISSN 0951–8320

[11] YANG, Jiannan: A general framework for probabilistic sensitivity analysis with respect to distribution parameters. In: *Probabilistic Engineering Mechanics* 72 (2023), S. 103433

[12] YANG, Jiannan: Decision-Oriented Two-Parameter Fisher Information Sensitivity Using Symplectic Decomposition. In: *Technometrics* 0 (2023), Nr. 0, 1-12. `http://dx.doi.org/10.1080/00401706.2023.2216251`. – DOI 10.1080/00401706.2023.2216251

[13] CAMPOLONGO, Francesca ; CARIBONI, Jessica ; SALTELLI, Andrea: An effective screening design for sensitivity analysis of large models. In: *Environmental modelling & software* 22 (2007), Nr. 10, S. 1509–1518

[14] KRZYKACZ-HAUSMANN, Bernard: Epistemic sensitivity analysis based on the concept of entropy. In: *Proceedings of SAMO2001* (2001), S. 31–35

[15] KALA, Zdeněk: Global sensitivity analysis based on entropy: From differential entropy to alternative measures. In: *Entropy* 23 (2021), Nr. 6, S. 778

[16] PAPOULIS, Athanasios: *Probability, random variables and stochastic processes*. 1984

[17] MODDEMEIJER, Rudy: On estimation of entropy and mutual information of continuous distributions. In: *Signal processing* 16 (1989), Nr. 3, S. 233–248

Appendix A Numerical estimation of entropy

Adopting the approach from [17], the xy-plane is gridded by equal size cells ($\Delta x \times \Delta y$) with coordinates (i, j). The probability of observing a sample in cell (i, j) is:

$$p_{ij} = \iint_{\text{cell}(i,j)} f(x,y)dxdy \approx f(x_i, y_j)\Delta x \Delta y \tag{A.1}$$

where (x_i, y_j) is the centre of the cell.

Assuming the jPDF is approximately constant within a cell, the joint entropy can be represented as:

$$\begin{aligned} H(X,Y) &= -\int f(x,y)\ln f(x,y)dxdy \\ &\approx -\sum f(x_i,y_j)\ln f(x_i,y_j)\Delta x\Delta y \\ &\approx -\sum p_{ij}\left(\ln p_{ij} - \ln(\Delta x \Delta y)\right) \\ &\approx -\sum \left(\frac{k_{ij}}{N}\ln\frac{k_{ij}}{N}\right) + \ln(\Delta x\Delta y) \end{aligned} \tag{A.2}$$

where k_{ij} represents the number of samples observed in the cell (i, j), and N is the total number of samples.

Similarly, the conditional entropy can be approximated as:

$$\begin{aligned} H(Y|X) &= -\int f(x,y)\frac{\ln f(x,y)}{f(x)}dxdy \\ &\approx -\sum p_{ij}\left(\ln p_{ij} - \ln p_i - \ln \Delta y\right) \\ &\approx -\sum \left(\frac{k_{ij}}{N}\ln\frac{k_{ij}}{k_i}\right) + \ln \Delta y \end{aligned} \tag{A.3}$$

where $k_i = \sum_j k_{ij}$ and similar expressions can be derived when $\mathbf{X}$ is a vector variable.