

Sobol' Matrices For Multi-Output Models With Quantified Uncertainty*

Robert A. Milton[†], Solomon F. Brown[‡], and Aaron S. Yeardley[§]

Abstract. Variance based global sensitivity usually measures the relevance of inputs to a single output using Sobol' indices. This paper extends the definition in a natural way to multiple outputs, directly measuring the relevance of inputs to the linkages between outputs in a correlation-like matrix of indices. The usual Sobol' indices constitute the diagonal of this matrix. Existence, uniqueness and uncertainty quantification are established by developing the indices from a putative regression model. Analytic expressions for generalized Sobol' indices and their standard deviations are given, and tested computationally against test functions whose ANOVA can be performed.

Key words. Global Sensitivity Analysis, Sobol' Index, Surrogate Model, Multi-Output, Gaussian Process, Uncertainty Quantification

MSC codes. 60G07, 60G15, 62J10, 62H99

1. Introduction. This paper is concerned with analysing the results of experiments or computer simulations in a design matrix of $M \geq 1$ input and $L \geq 1$ output columns, over N rows (datapoints). Global Sensitivity Analysis (GSA) [28] examines the relevance of the various inputs to the various outputs. When pursued via ANOVA decomposition of a single output $L = 1$, this leads naturally to the well known Sobol' indices, which have by now been applied across most fields of science and engineering [31, 16]. This paper extends the definition in a natural way to multiple outputs $L \geq 1$. Incidentally, the preference throughout this work is for uppercase constants and lowercase variables.

The Sobol' decomposition apportions the variance of a single output to sets of one or more inputs [35]. We shall use ordinals of inputs, tuples which are conveniently also naive sets.

$$(1.1) \quad \mathbf{m} := (0, \dots, m-1) \subseteq (0, \dots, M-1) =: \mathbf{M}$$

Obviously this restricts the subsets of inputs being studied, but this is a loss of convenience not generality, as any desired subset may be studied by ordering the inputs appropriately in advance. There is no generally accepted extension of Sobol' indices to multiple outputs $L > 1$. The discussion thus far, and almost all prior GSA, has dealt with a single (i.e scalar) output. With multiple (i.e vector) outputs, the Sobol' decomposition apportions the covariance matrix of outputs rather than the variance of a single output. With L outputs, the closed Sobol' index $S_{\mathbf{m}}$ is generally a symmetric $\mathbf{L} \times \mathbf{L}$ matrix. The diagonal elements express the relevance of inputs to the output variables themselves. The off-diagonal elements express relevance to the

*Submitted to the editors DATE.

Funding: This work was funded by the Fog Research Institute under contract no. FRI-454.

[†]Department of Chemical and Biological Engineering, University of Sheffield, Sheffield, S1 3JD, United Kingdom (r.a.milton@sheffield.ac.uk, <https://www.browngroupsheffield.com/>).

[‡]Department of Chemical and Biological Engineering, University of Sheffield, Sheffield, S1 3JD, United Kingdom (s.f.brown@sheffield.ac.uk, <https://www.browngroupsheffield.com/>).

[§]Department of Chemical and Biological Engineering, University of Sheffield, Sheffield, S1 3JD, United Kingdom (asyeardley1@sheffield.ac.uk, <https://www.browngroupsheffield.com/>).

linkages between outputs. This may be of considerable interest when outputs are, for example, yield and purity of a product, or perhaps a single output measured at various times. The Sobol indices reveal (amongst other things) which inputs it is worthwhile varying in an effort to alter the linkages between outputs. Prior work on Sobol' indices with multiple outputs [14, 43, 15] has settled ultimately on just the diagonal elements of the covariance matrix, so this linkage remains unexamined. Although output covariance has been incorporated indirectly in prior studies by performing principal component analysis (PCA) on outputs prior to GSA on the (diagonal) variances of the resulting output basis [7]. This has been used in particular to study synthetic "multi-outputs" which are actually the dynamic response of a single output over time [22, 46].

Accurate calculation of Sobol' indices even for a single output is computationally expensive and requires 10,000+ datapoints [23]. A (sometimes) more efficient approach is calculation via a surrogate model, such as Polynomial Chaos Expansion (PCE) [17, 45, 44], low-rank tensor approximation [10, 21], and support vector regression [13]. As well as being efficient, surrogate models also smooth out noise in the output, which is often highly desirable in practice. This paper employs one of the most popular surrogates, the Gaussian Processes (GP) [30, 27] as it is highly tractable. We shall follow the multi-output form (MOGP) described in [4], in order to examine the linkages between outputs. This paper deals exclusively with the anisotropic Radial Basis Function kernel, known as RBF/ARD, which is widely accepted as the kernel of choice for smooth outputs [26]. This uses the classic Gaussian bell curve to express the proximity of two input points, described in detail in [26].

Semi-analytic expressions for Sobol' indices are available for scalar PCEs [38], and the diagonal elements of multi-output PCEs [15]. Semi-analytic expressions for Sobol' indices of GPs have been provided in integral form by [25] and alternatively by [9]. These approaches are implemented, examined and compared in [24, 37]. Both [25, 24] estimate the errors on Sobol' indices in semi-analytic, integral form. Fully analytic, closed form expressions have been derived without error estimates for uniformly distributed inputs [42] with an RBF kernel. There are currently no closed form expressions for MOGPs, or the errors on Sobol' indices, or any GPs for which inputs are not uniformly distributed.

In this paper we provide explicit, closed-form analytic formulae for the $\mathbf{L} \times \mathbf{L}$ matrices of closed Sobol' indices and their errors, for a class of MOGP with an RBF/ARD kernel applicable to smoothly varying outputs. We transform uniformly distributed inputs \mathbf{u} to normally distributed inputs \mathbf{z} prior to fitting a GP and performing analytic calculation of closed Sobol' indices. This leads to relatively concise expressions in terms of exponentials, and enables ready calculation of the errors (variances) of these expressions. It also allows for an arbitrary rotation Θ of inputs, as normal variables are additive, whereas summing uniform inputs does not produce uniform inputs. If the goal is reducing inputs, rotating their basis first boosts the possibilities immensely [12]. It presents the possibility of choosing Θ to maximise the closed Sobol' index of the first few inputs.

The quantities to be calculated and their formal context are introduced in [2], assuming only that the output is an integrable function of the input. Our approach effectively regards a regression model which quantifies uncertainty with each prediction as just another name for a stochastic process. A great deal of progress is made in section 6 using general stochastic (not necessarily Gaussian) processes. This approach is analytically cleaner, as it is not obfuscated

77 by the GP details. Furthermore, it turns out that the desirable properties of the Gaussian (lack
78 of skew, simple kurtosis) are not actually helpful, as these terms cancel of their own accord.
79 This development leaves just two terms to be calculated, which require the stochastic process
80 to be specified. MOGPs with an RBF/ARD kernel are tersely developed and described in ??,
81 then used to calculate the two unknown terms in ?????. Methods to reduce computational
82 complexity are discussed in ??.Conclusions are drawn in ??.
83 The void ordinal $\mathbf{0} := ()$ voids any tensor it subscripts.

2. Multi-output models with quantified uncertainty.

The purpose of this Section is to set the scene for this study, serve as a glossary of notation, and prove the formal foundation of the construction to follow. Our notation is not entirely standard, but is preferred for lightness and fluency once grasped. Regarding notation and other topics, the intention is to facilitate computation by combining efficient, tensorized operations such as `einsum` [1, 2, 3] into numerically stable calculations. As such, the devices employed herein will be familiar to practitioners of machine learning.

Throughout this paper, square bracketed quantities such as $[u]_{\mathbf{M} \times \mathbf{o}}$ are tensors subscripted by (the cartesian product of) their ranks, for bookkeeping. Ranks are expressed as ordinals of axes, tuples which are conveniently also naive sets and will be subtracted as such

$$\begin{aligned} \mathbf{0} &:= () \subseteq \mathbf{m} := (0, \dots, m-1) \subseteq \mathbf{M} := (0, \dots, M-1) \\ \mathbf{M} - \mathbf{m} &:= \{m' \in \mathbf{M} \mid m' \notin \mathbf{m}\} = (m, \dots, M-1) \end{aligned}$$

GSA decomposes input space by sets of axes, and this notation facilitates that. It does assume that input axes are expediently ordered already, which might impair convenience but not generality. Conventional indexing is permitted under this scheme, by allowing singleton axes to be written in *italic* without parentheses.

From a Bayesian perspective a determined tensor is a function of determined input(s), a random variable (RV) a function of undetermined input, and a stochastic process (SP) a function of both determined and undetermined inputs. In any case, all sources of (sans-serif) uncertainty may be gathered in a single input $[u]_{\mathbf{M}}$, final to the definition of input space:

$$(2.1) \quad \begin{array}{ll} \mathbf{M} := (0, \dots, M-1) & \text{determined inputs on the unit interval} \quad [u]_{\mathbf{M}} \in [0, 1]^M \\ M =: \mathbf{M} + \mathbf{1} - \mathbf{M} & \text{an undetermined input on the unit interval} \quad [u]_{\mathbf{M}} \sim \mathcal{U}(0, 1) \end{array}$$

Exponentiation is always categorical – repeated cartesian \times or tensor \otimes – unless stated otherwise. It is crucial to this work that all $\mathbf{M} + \mathbf{1}$ inputs vary completely independently of each other – they are in no way codependent or correlated. Uncertainty is distributed uniformly $[u]_{\mathbf{M}} \sim \mathcal{U}(0, 1)$ to exploit the “universality of the uniform” [6, pp.224], better known as the inverse probability integral transform [5]. This explicitly provides a bijective mapping to any RV $[\tilde{u}]_{\mathbf{M}}$ by inverting its cumulative distribution function (CDF) $\tilde{F}: [\tilde{u}]_{\mathbf{M}} \mapsto [u]_{\mathbf{M}}$ in the generalized sense

$$\tilde{F}^{-1}([u]_{\mathbf{M}}) := \inf \left\{ [\tilde{u}]_{\mathbf{M}} \mid \tilde{F}([\tilde{u}]_{\mathbf{M}}) \geq [u]_{\mathbf{M}} \right\}$$

This mapping $\tilde{F}: [u]_{\mathbf{M}} \mapsto [\tilde{u}]_{\mathbf{M}}$ is the quantile function of $[\tilde{u}]_{\mathbf{M}}$ on $[0, 1]$, and embedding it in an output model y can express any continuous RV at a given determined input $[u]_{\mathbf{M}}$. In other words, the output model could be any continuous SP. The amenity of definition (2.1) is that each input $[u]_{\mathbf{m}}$ is automatically a probability measure, for any $m \in \mathbf{M}$.

Throughout this work, expectations are subscripted by the inputs marginalized

$$\mathbb{E}_{\#}[\bullet] := \int_{[0]_{\#}}^{[1]_{\#}} [\bullet] \, d[u]_{\#} \quad \forall \# \subseteq \mathbf{M} + \mathbf{1} \text{ or } \# \in \mathbf{M} + \mathbf{1}$$

118 Covariances invoke the tensor product (summing the ranks of the arguments), and carry the
 119 subscript of their underlying expectation

$$120 \quad \mathbb{V}_{\#}[[\bullet], [\ast]] := \mathbb{E}_{\#}[[\bullet] \otimes [\ast]] - \mathbb{E}_{\#}[\bullet] \otimes \mathbb{E}_{\#}[\ast] \quad \forall \# \subseteq \mathbf{M} + \mathbf{1} \text{ or } \# \in \mathbf{M} + \mathbf{1}$$

121 The covariance of anything with itself is expressed with a single argument $\mathbb{V}_{\#}[\bullet] := \mathbb{V}_{\#}[\bullet, \bullet]$,
 122 as is customary.

123 A multi-output model with quantified uncertainty (MQU) is defined as any Lebesgue
 124 integrable function of input space (2.1)

$$125 \quad (2.2) \quad y: [0, 1]^{M+1} \rightarrow \mathbb{R}^L \quad \text{such that} \\ \mathbb{V}_{\mathbf{M}} \mathbb{E}_M [y([u]_{\mathbf{M}+\mathbf{1}})]_l > \mathbb{E}_{\mathbf{M}} \mathbb{E}_M [y([u]_{\mathbf{M}+\mathbf{1}})]_l = 0 \quad \forall l \in \mathbf{L}$$

126 To avoid division by zero in GSA, every output component $l \in \mathbf{L}$ must depend on at least
 127 one determined input: constant functions and pure noise are not allowed. Without loss
 128 of generality we have also offset the model to have zero mean $[0]_{\mathbf{L}}$. This is algebraically
 129 unnecessary, but it can be crucial to numerical stability of computation.

130 Our notation will reflect machine learning practice [1, 2, 3] which facilitates parallel com-
 131 putation by tensorizing function application over batch dimensions such as \mathbf{o} according to

$$132 \quad [y([u]_{\mathbf{M}+\mathbf{1} \times \mathbf{o}})]_{\mathbf{L} \times \mathbf{o}} := y([u]_{\mathbf{M}+\mathbf{1} \times \mathbf{o}}) \quad \forall \mathbf{o} \in \mathbf{o}$$

133 To elucidate the meaning and purpose of an MQU, and set the stage for GSA, consider a
 134 design matrix of $[u]_{\mathbf{M} \times \mathbf{o}}$ inputs alongside $[Y]_{\mathbf{L} \times \mathbf{o}}$ outputs:

$$135 \quad \begin{array}{cccccc} [u]_{0 \times 0} & \cdots & [u]_{M-1 \times 0} & [Y]_{0 \times 0} & \cdots & [Y]_{L-1 \times 0} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ [u]_{0 \times o-1} & \cdots & [u]_{M-1 \times o-1} & [Y]_{0 \times o-1} & \cdots & [Y]_{L-1 \times o-1} \end{array}$$

136 The sole function of an MQU is actually to provide a single number for any such design matrix
 137 of $o \in \mathbb{Z}^+$ samples (rows): namely, the probability that the output samples result from the
 138 input samples (to within some error $\sigma > 0$)

$$139 \quad \mathbb{P} \left[(y([u]_{\mathbf{M}+\mathbf{1} \times \mathbf{o}}) - [Y]_{\mathbf{L} \times \mathbf{o}})^2 < \sigma^2 \mid [u]_{\mathbf{M} \times \mathbf{o}} = [u]_{\mathbf{M} \times \mathbf{o}} \right]$$

140 This is what it means to generate L quantifiably uncertain outputs from M determined inputs.
 141 Following the Kolmogorov extension theorem [29, pp.124], the meaning of an MQU is nothing
 142 other than an SP. This includes fully determined MQUs – $y([u]_{\mathbf{M}+\mathbf{1}})$ which do not depend
 143 on the final, undetermined input $[u]_M$ – as SPs with zero variance. Kolmogorov's exten-
 144 sion theorem formally identifies an SP – a collection $y([u]_{\mathbf{M}+\mathbf{1}}) = \{y([u]_{\mathbf{M}+\mathbf{1}}) \mid [u]_{\mathbf{M}} = [u]_{\mathbf{M}}\}$
 145 of RVs indexed by determined inputs $[u]_{\mathbf{M}}$ – with all its finite dimensional distributions
 146 $\{y([u]_{\mathbf{M}+\mathbf{1} \times \mathbf{o}}) \mid [u]_{\mathbf{M} \times \mathbf{o}} = [u]_{\mathbf{M} \times \mathbf{o}}\}$. The latter viewpoint is the formal version of the random
 147 field [20] or random function interpretation of SPs [33, pp.42] frequently alluded to in machine
 148 learning [27]. It can be a helpful perspective on the objects appearing in this work.

In conception MQUs include any explicit or black-box function (simulation), surrogate (emulator, response surface, meta-model) or regression which supplies a probability distribution (even a zero variance one) on its outputs (predictions). In practice, such an MQU is usually inferred from a design of experiments (DOE) which samples the determined inputs $[u]_{\mathbf{M}}$ from $U(0, 1)^M$. This is not restrictive, a preferred set of determined inputs $[x]_{\mathbf{M}}$ may have any sampling distribution whatsoever, so long as it is continuous and its M dimensions mutually independent. Simply take the CDF of $[x]_{\mathbf{M}}$ to obtain $[u]_{\mathbf{M}}$ before beginning, and apply the quantile function of $[x]_{\mathbf{M}}$ (described above and in [5]) after finishing. In general and in summary, simply treat any $\tilde{y}([x]_{\mathbf{M}+1})$ as the pullback of an MQU by the CDF of $[x]_{\mathbf{M}+1}$.

To close this Section we shall formally secure the construction which follows. An MQU y is Lebesgue integrable by definition (2.2), so it must be measurable. All measures throughout this work are finite (probability measures, in fact), so integrability of y implies integrability of y^n for all $n \in \mathbb{Z}^+$ [39]. Therefore, Fubini's theorem [41, pp.77] allows all expectations to be taken in any order (which will be crucial later). We can thus safely construct any SP which is polynomial in y and its marginals; and freely extract finite dimensional distributions from it by the Kolmogorov extension theorem [29, pp.124]. This guarantees existence and uniqueness of every device in this paper.

3. GSA Sobol' indices. This Section recapitulates the definition of classical Sobol' indices in the context of an MQU, to prepare the way for easy generalization to Sobol' matrices in the next Section. Sobol' indices are traditionally constructed [35] from a High Dimensional Model Representation (HDMR or Hoeffding-Sobol' decomposition [8]). This is enlightening, but not strictly necessary. Instead, we shall swiftly construct the indices from an MQU. Later, this will allow us to derive standard errors on the Sobol' indices.

In essence, GSA is performed by marginalizing (obscuring) determined input axes selectively. A reduced model (mQU) is an SP defined as

$$(3.1) \quad [y_{\mathbf{m}}]_{\mathbf{L}} := \mathbb{E}_{\mathbf{M}-\mathbf{m}}[y([u]_{\mathbf{M}+1})]_{\mathbf{L}} \quad \forall \mathbf{m} \subseteq \mathbf{M}$$

When $\mathbf{M} - \mathbf{m} = \mathbf{0}$, nothing is marginalized and $y_{\mathbf{M}}$ is called the full model (MQU). The undetermined input $[u]_{\mathbf{M}}$ is never marginalized, and each mQU depends on \mathbf{m} determined inputs $[u]_{\mathbf{m}}$ alongside $[u]_{\mathbf{M}}$. Any distributional assumption about inputs simply pulls back (factors through) expectations here with no impact whatsoever: $U(0, 1)$ is no more than a proxy for the CDF of any continuous distribution.

An mQU $[y_{\mathbf{m}}]_{\mathbf{L}}$ is not automatically converted to a new MQU over $[u]_{\mathbf{m}+1}$ by re-indexing $[u]_{\mathbf{M}}$ to $[u]_{\mathbf{m}}$. This would not be the same as simply hiding columns $\mathbf{M} - \mathbf{m}$ in the design matrix of the previous Section. In general the new MQU would not adequately reflect the original full model, because any output variation due to marginalized inputs $[u]_{\mathbf{M}-\mathbf{m}}$ has simply been lost: it cannot be re-allocated to the new undetermined input $[u]_{\mathbf{m}}$. This gets to the heart of GSA, assessing the degree to which a reduced model mQU mimics its full model. Which is a matter of assessing how much output variation has been lost through marginalization.

The marginal variance of an mQU is a tensor-valued RV (function of $[u]_{\mathbf{M}}$) defined as

$$(3.2) \quad \mathbb{V}_{\mathbf{m}}[y_{\mathbf{m}}]_{\mathbf{L} \times \mathbf{L}} := \mathbb{V}_{\mathbf{m}}[[y_{\mathbf{m}}]_{\mathbf{L}}]$$

It is – for all realisations $[u]_{\mathbf{M}} = [u]_{\mathbf{M}}$ – a symmetric tensor by definition and positive semi-definite by Jensen's inequality. We may therefore define a vector RV which is the square root

191 of its diagonal, namely the standard deviation $\mathbb{D}_{\mathbf{m}}[\mathbf{y}_{\mathbf{m}}]_{\mathbf{L}}$

$$192 \quad (3.3) \quad \mathbb{D}_{\mathbf{m}}[\mathbf{y}_{\mathbf{m}}]_l := \sqrt{\mathbb{V}_{\mathbf{m}}[\mathbf{y}_{\mathbf{m}}]_{l \times l}} > 0 \quad \forall l \in \mathbf{L}$$

193 This is positive definite due to the positive variance clause in the definition (2.2) of an MQU.

194 The closed Sobol' index of input axes \mathbf{m} with respect to scalar output $[y([\mathbf{u}]_{\mathbf{M}+1})]_l$ is
195 defined as an RV

$$196 \quad (3.4) \quad [\mathbf{S}_{\mathbf{m}}]_l := \frac{\mathbb{V}_{\mathbf{m}}[\mathbf{y}_{\mathbf{m}}]_{l \times l}}{\mathbb{V}_{\mathbf{M}}[\mathbf{y}_{\mathbf{M}}]_{l \times l}}$$

197 The definition originated [34] for fully determined functions (y independent of $[\mathbf{u}]_{\mathbf{M}}$). It
198 was later extended to random inputs [36] and Gaussian processes [25, 24]. The latter work
199 introduced the Sobol' index as an RV constructed from an important category of SP.

200 The complement of a closed index is called a total index [18]

$$201 \quad (3.5) \quad [\mathbf{S}_{\mathbf{M}-\mathbf{m}}^T]_l := 1 - [\mathbf{S}_{\mathbf{m}}]_l$$

202 Sobol' indices are readily interpreted, given that codependent inputs (reviewed in [40]) are
203 beyond the scope of this work. By Definition (3.4) a closed index $S_{\mathbf{m}} \in [0, 1]$ measures the
204 proportion of output variance captured by the reduced model $\mathbf{y}_{\mathbf{m}}$. This, in turn, indicates the
205 influence or relevance of input axes \mathbf{m} . The full model $\mathbf{y}_{\mathbf{M}}$ explains everything explicable, so
206 its Sobol' index is $[\mathbf{S}_{\mathbf{M}}]_l = [\mathbf{S}_{\mathbf{M}}^T]_l = 1$. The void model $\mathbf{y}_{\mathbf{0}}$ is just the mean output – an RV
207 depending only on undetermined noise $[\mathbf{u}]_{\mathbf{M}}$ – which explains nothing, so its Sobol' index is
208 $[\mathbf{S}_{\mathbf{0}}]_l = [\mathbf{S}_{\mathbf{0}}^T]_l = 0$.

209 The closed index of a single input axis $m \in \mathbf{M}$ is called a first-order index

$$210 \quad (3.6) \quad [\mathbf{S}_m]_l := [\mathbf{S}_{\mathbf{m}+1-\mathbf{m}}]_l$$

211 Because inputs may cooperate to affect the output, a closed index often exceeds the sum of
212 its first-order contributions, obeying (for any realization $[\mathbf{u}]_{\mathbf{M}} = [u]_{\mathbf{M}}$)

$$213 \quad (3.7) \quad \sum_{\dot{\mathbf{m}} \in \mathbf{m}} [\mathbf{S}_{\dot{\mathbf{m}}}]_l \leq [\mathbf{S}_{\mathbf{m}}]_l \leq [\mathbf{S}_{\mathbf{m}}^T]_l$$

214 The final inequality observes that a closed index only includes cooperation of input axes \mathbf{m}
215 with each other, whereas a total index also includes cooperation between \mathbf{m} and $\mathbf{M} - \mathbf{m}$ (but
216 excludes cooperation of input axes $\mathbf{M} - \mathbf{m}$ with each other. This is the only difference between
217 closed and total indices, but it is an important one.

218 Regarding the covariance between any reduced model and its full model, Fubini's theorem
219 safeguards the intuitive identity

$$\begin{aligned} 220 \quad (3.8) \quad \mathbb{V}_{\mathbf{M}}[\mathbf{y}_{\mathbf{m}}, \mathbf{y}_{\mathbf{M}}]_{\mathbf{L} \times \mathbf{L}} &:= \mathbb{E}_{\mathbf{M}}[[\mathbf{y}_{\mathbf{m}}]_{\mathbf{L}} \otimes [\mathbf{y}_{\mathbf{M}}]_{\mathbf{L}}] - \mathbb{E}_{\mathbf{M}}[\mathbf{y}_{\mathbf{m}}]_{\mathbf{L}} \otimes \mathbb{E}_{\mathbf{M}}[\mathbf{y}_{\mathbf{M}}]_{\mathbf{L}} \\ &= \mathbb{E}_{\mathbf{m}}[[\mathbf{y}_{\mathbf{m}}]_{\mathbf{L}} \otimes [\mathbf{y}_{\mathbf{M}}]_{\mathbf{L}}] - [\mathbf{y}_{\mathbf{0}}]_{\mathbf{L}} \otimes [\mathbf{y}_{\mathbf{0}}]_{\mathbf{L}} \\ &=: \mathbb{V}_{\mathbf{m}}[\mathbf{y}_{\mathbf{m}}]_{\mathbf{L} \times \mathbf{L}} \end{aligned}$$

221 Using this, the correlation between (the scalar output predicted by) an mQU and its parent
 222 MQU is an RV

$$223 \quad (3.9) \quad [R_{\mathbf{m}}]_l := \frac{\mathbb{V}_{\mathbf{M}}[y_{\mathbf{m}}, y_{\mathbf{M}}]_{l \times l}}{\mathbb{D}_{\mathbf{m}}[y_{\mathbf{m}}]_l \mathbb{D}_{\mathbf{M}}[y_{\mathbf{M}}]_l} = \frac{\mathbb{D}_{\mathbf{m}}[y_{\mathbf{m}}]_l}{\mathbb{D}_{\mathbf{M}}[y_{\mathbf{M}}]_l}$$

224 The square of the correlation is a discerning measure of the quality of regression [11] called
 225 the coefficient of determination. Which is now clearly identical to the closed Sobol' index

$$226 \quad (3.10) \quad [R_{\mathbf{m}}]_l^2 = \frac{\mathbb{V}_{\mathbf{m}}[y_{\mathbf{m}}]_l}{\mathbb{D}_{\mathbf{M}}[y_{\mathbf{M}}]_l \mathbb{D}_{\mathbf{M}}[y_{\mathbf{M}}]_l} =: [S_{\mathbf{m}}]_l$$

227 Sobol' indices are used to identify reduced models $[y_{\mathbf{m}}]_l$ which adequately mimic the full model
 228 $[y_{\mathbf{M}}]_l$. A closed index close to 1 confirms that the two models make nearly identical predictions.
 229 Simplicity and economy (not least of calculation) motivate the adoption of a reduced model,
 230 a closed Sobol' index close to 1 is what justifies it. This is precisely equivalent to screening
 231 out (obscuring) $\mathbf{M} - \mathbf{m}$ input axes on the grounds that their influence on $[y_{\mathbf{M}}]_l$ – measured
 232 by their total index $[S_{\mathbf{M}-\mathbf{m}}^T]_l$ – is close to 0.

233 **4. GSA Matrices.** This Section introduces a tensor RV called the Sobol' matrix of an
 234 mQU. This is defined by tensorizing the definitions of the previous Section, replacing l for a
 235 scalar output with $\mathbf{L} \times \mathbf{L}$ for a vector output. All division is performed element-wise, inverting
 236 Hadamard multiplication \circ

$$237 \quad (4.1) \quad [*]_{\#} = \frac{[\bullet]_{\#}}{[\star]_{\#}} \iff [*]_{\#} \circ [\star]_{\#} = [\bullet]_{\#}$$

238 The tensorization of definition (3.4) is problematic, as it may invoke division by zero.
 239 Instead, we shall tensorize the alternative definition (3.10) of a closed Sobol' index as a
 240 coefficient of determination. We shall then compare the two definitions in the light of a toy
 241 example.

242 The closed Sobol' matrix of input axes \mathbf{m} with respect to vector output $[y([u]_{\mathbf{M}+\mathbf{1}})]_{\mathbf{L}}$ is
 243 defined as the tensor RV

$$244 \quad (4.2) \quad [S_{\mathbf{m}}]_{\mathbf{L} \times \mathbf{L}} := \frac{\mathbb{V}_{\mathbf{m}}[y_{\mathbf{m}}]_{\mathbf{L} \times \mathbf{L}}}{\mathbb{D}_{\mathbf{M}}[y_{\mathbf{M}}]_{\mathbf{L}} \otimes \mathbb{D}_{\mathbf{M}}[y_{\mathbf{M}}]_{\mathbf{L}}}$$

245 By this definition, the closed Sobol' matrix of the full model is just the ordinary correlation
 246 matrix of the components of output vector y

$$247 \quad (4.3) \quad [R]_{\mathbf{L} \times \mathbf{L}} := \frac{\mathbb{V}_{\mathbf{M}}[y_{\mathbf{M}}]_{\mathbf{L} \times \mathbf{L}}}{\mathbb{D}_{\mathbf{M}}[y_{\mathbf{M}}]_{\mathbf{L}} \otimes \mathbb{D}_{\mathbf{M}}[y_{\mathbf{M}}]_{\mathbf{L}}} =: [S_{\mathbf{M}}]_{\mathbf{L} \times \mathbf{L}}$$

248 The complement of a closed Sobol' matrix is called a total Sobol' matrix

$$249 \quad (4.4) \quad [S_{\mathbf{M}-\mathbf{m}}^T]_{\mathbf{L} \times \mathbf{L}} := [S_{\mathbf{M}}]_{\mathbf{L} \times \mathbf{L}} - [S_{\mathbf{m}}]_{\mathbf{L} \times \mathbf{L}}$$

Let us emphasise that these definitions are equivalent to the Sobol' index definitions (3.4) and (3.5) on the Sobol' matrix diagonal

$$[S_m]_{l \times l} = [S_m]_l \quad \text{and} \quad [S_m^T]_{l \times l} = [S_m^T]_l \quad \forall l \in \mathbf{L}$$

Just as a Sobol' index is a proportion of output variance, each element of a Sobol' matrix is a component of output correlation. Unlike the former, the latter can be negative. It is still sometimes useful to examine the first-order matrix for a single input axis m

$$(4.5) \quad [S_m]_{\mathbf{L} \times \mathbf{L}} := [S_{\mathbf{m}+1-\mathbf{m}}]_{\mathbf{L} \times \mathbf{L}}$$

However, there is no ordering analagous to (3.7) as contributions of opposing sign may offset each other. All one can rely on – elementwise for any realization $[u]_M = [u]_M - i$

$$(4.6a) \quad [S_0]_{\mathbf{L} \times \mathbf{L}} = [S_0^T]_{\mathbf{L} \times \mathbf{L}} = [0]_{\mathbf{L} \times \mathbf{L}}$$

$$(4.6b) \quad [S_M]_{\mathbf{L} \times \mathbf{L}} = [S_M^T]_{\mathbf{L} \times \mathbf{L}} = [R]_{\mathbf{L} \times \mathbf{L}}$$

$$(4.6c) \quad [0]_{\mathbf{L} \times \mathbf{L}} \leq |[S_m]_{\mathbf{L} \times \mathbf{L}}|, |[S_m^T]_{\mathbf{L} \times \mathbf{L}}| \leq |[R]_{\mathbf{L} \times \mathbf{L}}|$$

So the void model $[y_0]_{\mathbf{L}}$ still explains nothing, the full model $[y_0]_{\mathbf{L}}$ still explains everything explicable, and every other mQU lies somewhere in between.

Let us pursue an informative toy example to illustrate this. Take some even M and define

$$w: [0, 1]^{M/2} \rightarrow \mathbb{R} \quad \text{such that} \quad \mathbb{E}_{\mathbf{M}/2}[w] = 0 \quad \text{and} \quad \mathbb{V}_{\mathbf{M}/2}[w] = 1$$

to construct the $(L = 2)$ MQU

$$y([u]_{\mathbf{M}+1}) = \begin{bmatrix} w([u]_{\mathbf{M}/2}) + w([u]_{\mathbf{M}-\mathbf{M}/2}) \\ w([u]_{\mathbf{M}/2}) - w([u]_{\mathbf{M}-\mathbf{M}/2}) \end{bmatrix}$$

The Sobol' matrix of the full model is easily calculated

$$S_M = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Two important reduced models have the Sobol' matrices

$$S_{\mathbf{M}/2} = S_{\mathbf{M}/2}^T = \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix} \quad ; \quad S_{\mathbf{M}-\mathbf{M}/2} = S_{\mathbf{M}-\mathbf{M}/2}^T = \begin{bmatrix} 1/2 & -1/2 \\ -1/2 & 1/2 \end{bmatrix}$$

The Sobol' matrices reveal that input axes $\mathbf{M}/2$ influence both outputs in the same sense, while input axes $\mathbf{M} - \mathbf{M}/2$ influence the two outputs in opposite sense. In real-world examples, this information is often valuable. Perhaps the two components of y are the yield and the purity of a pharamaceutical product, or the efficacy and the cost of CO₂ capture at a power plant. We can also ascertain that axes in $\mathbf{M}/2$ do not cooperate with axes in $\mathbf{M} - \mathbf{M}/2$, as witnessed by total Sobol' matrices equal to closed Sobol' matrices.

The direct tensorization of Sobol' index definition (3.4) would define the matrix

$$\left[\hat{\mathbf{S}}_{\mathbf{m}}\right]_{\mathbf{L} \times \mathbf{L}} := \frac{[\mathbf{S}_{\mathbf{m}}]_{\mathbf{L} \times \mathbf{L}}}{[\mathbf{S}_{\mathbf{M}}]_{\mathbf{L} \times \mathbf{L}}} \quad \text{which is undefined whenever} \quad [\mathbf{S}_{\mathbf{M}}]_{l \times l'} = 0 \text{ for some } l, l' \in \mathbf{L}$$

In our toy example, the most interesting information – the off-diagonal indices for two important reduced models – is undefined and completely lost by this definition.

We close this section by considering how a Sobol' matrix may be summarised in a single number. This is often desired to present “the” sensitivity of the vector output to a reduced set of inputs \mathbf{m} . The simple answer is to define a seminorm [32, pp.314] (like a norm, only positive semidefinite instead of positive definite) on the Sobol' matrix. The seminorm should be chosen according to one's interest in the outputs. For example, one could use the determinant $|[\mathbf{S}_{\mathbf{m}}]_{\mathbf{l} \times \mathbf{l}}|$ of a chosen submatrix $\mathbf{l} \subseteq \mathbf{L}$, or the modulus in case \mathbf{l} is singleton. The Sobol' matrix provides a platform for investigating a variety of such measures. However, all matrices and measures remain RVs at this point, and the next two sections are devoted to extracting statistics – i.e. determined quantities – from these.

5. GSA Statistics. At this stage, Sobol' matrices have been robustly defined as tensor RVs – i.e. functions of the undetermined input $[\mathbf{u}]_{\mathbf{M}}$. This section extracts two determined matrix statistics from each RV: its expected value and its standard deviation or error.

Expected values are written as the italic version of the underlying RV, starting with

$$(5.1a) \quad \text{marginal variance} \quad [\mathbf{V}_{\mathbf{m}}]_{\mathbf{L} \times \mathbf{L}} := \mathbb{E}_{\mathbf{M}} \mathbb{V}_{\mathbf{m}}[\mathbf{y}_{\mathbf{m}}]_{\mathbf{L} \times \mathbf{L}}$$

$$(5.1b) \quad \text{marginal deviation} \quad [\mathbf{D}_{\mathbf{m}}]_{\mathbf{L}} := \mathbb{E}_{\mathbf{M}} \mathbb{D}_{\mathbf{m}}[\mathbf{y}_{\mathbf{m}}]_{\mathbf{L}} \implies [\mathbf{D}_{\mathbf{m}}]_l = \sqrt{[\mathbf{V}_{\mathbf{m}}]_{l \times l}} \quad \forall l \in \mathbf{L}$$

The GSA statistics used to assess input relevance are the expected Sobol' matrices

$$(5.2a) \quad \text{closed Sobol' matrix} \quad [\mathbf{S}_{\mathbf{m}}]_{\mathbf{L}^2} := \mathbb{E}_{\mathbf{M}} [\mathbf{S}_{\mathbf{m}}]_{\mathbf{L}^2} = \frac{[\mathbf{V}_{\mathbf{m}}]_{\mathbf{L}^2}}{[\mathbf{D}_{\mathbf{m}}]_{\mathbf{L}} \otimes [\mathbf{D}_{\mathbf{m}}]_{\mathbf{L}}}$$

$$(5.2b) \quad \text{total Sobol' matrix} \quad [\mathbf{S}_{\mathbf{M}-\mathbf{m}}^T]_{\mathbf{L}^2} := \mathbb{E}_{\mathbf{M}} [\mathbf{S}_{\mathbf{M}-\mathbf{m}}^T]_{\mathbf{L}^2} = [\mathbf{S}_{\mathbf{M}}]_{\mathbf{L}^2} - [\mathbf{S}_{\mathbf{m}}]_{\mathbf{L}^2}$$

In order to express standard errors – which are due to undetermined input $[\mathbf{u}]_{\mathbf{M}}$ – the notation introduced in (3.3) is extended to define the matrix standard deviation $\mathbb{D}_{\mathbf{M}}[\bullet]_{\mathbf{L} \times \mathbf{L}}$ of a matrix RV $[\bullet]_{\mathbf{L} \times \mathbf{L}}$

$$(5.3) \quad \mathbb{D}_{\mathbf{M}}[\bullet]_{l \times l'} := \sqrt{\mathbb{V}_{\mathbf{M}}[[\bullet]_{l \times l'}]} \quad \forall l, l' \in \mathbf{L}$$

The GSA statistics used to assess the standard error of Sobol' matrices are

$$(5.4a) \quad \text{closed Sobol' matrix error} \quad [\mathbf{T}_{\mathbf{m}}]_{\mathbf{L}^2} := \mathbb{D}_{\mathbf{M}}[\mathbf{S}_{\mathbf{m}}]_{\mathbf{L}^2}$$

$$(5.4b) \quad \text{total Sobol' matrix error} \quad [\mathbf{T}_{\mathbf{M}-\mathbf{m}}^T]_{\mathbf{L}^2} := [\mathbf{T}_{\mathbf{M}}]_{\mathbf{L}^2} + [\mathbf{T}_{\mathbf{m}}]_{\mathbf{L}^2} \geq \mathbb{D}_{\mathbf{M}}[\mathbf{S}_{\mathbf{M}-\mathbf{m}}^T]_{\mathbf{L}^2}$$

The total Sobol' matrix error is a conservative statistic which achieves equality on the diagonal, and is robust and sufficiently precise for most practical purposes off the diagonal. Attempts at greater precision are apt to produce wildly unreliable estimates when implemented on a

computer. This is because the total Sobol' matrix (4.4) is the difference between two terms which are often highly correlated with each other. Tiny differences between correlated terms are swamped by numerical error in implementation, so the resulting computation is apt to yield garbage in practice.

To calculate the closed Sobol' matrix error $[T_{\mathbf{m}}]_{\mathbf{L} \times \mathbf{L}}$ we use the Taylor series method [19, pp.353], which is valid provided $\mathbb{V}_{\mathbf{M}}[y_{\mathbf{M}}]_{l \times l}$ is well approximated by its mean

$$[V_{\mathbf{M}}]_{l \times l} \gg |\mathbb{V}_{\mathbf{M}}[y_{\mathbf{M}}]_{l \times l} - [V_{\mathbf{M}}]_{l \times l}|$$

This is essentially the positive variance clause in the definition (2.2) of an MQU, designed to prohibit constant or pure noise outputs which do not depend on any of the determined input axes \mathbf{M} .

The covariance between two marginal covariance matrix RVs is the determined 4-tensor

$$(5.5) \quad [W_{\mathbf{mm}'}]_{\mathbf{L}^4} := \mathbb{V}_M[\mathbb{V}_{\mathbf{m}}[y_{\mathbf{m}}]_{\mathbf{L} \times \mathbf{L}}, \mathbb{V}_{\mathbf{m}'}[y_{\mathbf{m}'}]_{\mathbf{L} \times \mathbf{L}}] \quad \forall \mathbf{m}, \mathbf{m}' \in \mathbf{M}$$

used to define the determined matrix $[Q_{\mathbf{m}}]_{\mathbf{L} \times \mathbf{L}}$ elementwise

$$(5.6) \quad [Q_{\mathbf{m}}]_{l \times l'} := [W_{\mathbf{mm}}]_{(l \times l')^2} - [V_{\mathbf{m}}]_{l \times l'} \sum_{l^{\ddagger} \in \{l, l'\}} \frac{[W_{\mathbf{mm}}]_{l^{\ddagger} \times l^{\ddagger} \times l \times l'}}{[D_{\mathbf{M}}]_{l^{\ddagger}}^2} + \frac{[V_{\mathbf{m}}]_{l \times l'}^2}{4} \sum_{l^{\ddagger} \in \{l, l'\}} \frac{[W_{\mathbf{MM}}]_{l^{\ddagger} \times l^{\ddagger} \times l \times l}}{[D_{\mathbf{M}}]_{l^{\ddagger}}^2 [D_{\mathbf{M}}]_l^2} + \frac{[W_{\mathbf{MM}}]_{l^{\ddagger} \times l^{\ddagger} \times l' \times l'}}{[D_{\mathbf{M}}]_{l^{\ddagger}}^2 \otimes [D_{\mathbf{M}}]_{l'}^2}$$

The Taylor series estimate of the closed Sobol matrix error (5.4a) is finally calculated elementwise as

$$(5.7) \quad [T_{\mathbf{m}}]_{l \times l'} := \mathbb{D}_M[S_{\mathbf{m}}]_{l \times l'} = \frac{\sqrt{[Q_{\mathbf{m}}]_{l \times l'}}}{[D_{\mathbf{M}}]_l [D_{\mathbf{M}}]_{l'}}$$

It is satisfying to note that these equations enforce

$$[T_{\mathbf{M}}]_{l \times l} = 0 \quad \text{on the diagonal} \quad [S_{\mathbf{M}}]_{l \times l} = 1$$

6. GSA via mQU moments. The central problem in calculating errors on Sobol' matrices is that they involve ineluctable covariances over ungoverned noise between differently marginalized SPs. But marginalization and covariance determination are both a matter of taking expectations. Using Fubini's theorem [41, pp.77] the ineluctable can be avoided by reversing the order of expectations – taking moments over ungoverned noise, then marginalizing.

The covariance function or kernel [27] of an MQU measures the similarity between the output by two determined inputs $[u]_{\mathbf{M}}, [u']_{\mathbf{M}'}$ as

$$(6.1) \quad [k([u]_{\mathbf{M}}, [u']_{\mathbf{M}'})]_{\mathbf{L} \times \mathbf{L}'} := \mathbb{V}_M[y([u]_{\mathbf{M}}), y([u']_{\mathbf{M}'})]$$

Throughout this work, primes on uppercase constants are for bookkeeping only, they do not affect the value $\mathbf{M}^{n'} := \mathbf{M}$; $\mathbf{L}^{n'} := \mathbf{L}$. Primes on lowercase variables do affect the value, so it may or may not be the case that $\mathbf{m}^{n'} = \mathbf{m}$; $\mathbf{l}^{n'} = \mathbf{l}$. All one needs to secure a covariance function is the ability to calculate the finite-dimensional distribution

$$\mathbb{P} \left[\left(y([\mathbf{u}]_{\mathbf{M}+1 \times 2}) - \begin{bmatrix} [Y]_{\mathbf{L}} \\ [Y']_{\mathbf{L}'} \end{bmatrix} \right)^2 < \sigma^2 \mid [\mathbf{u}]_{\mathbf{M} \times 2} = \begin{bmatrix} [\mathbf{u}]_{\mathbf{M}} \\ [\mathbf{u}']_{\mathbf{M}'} \end{bmatrix} \right]$$

for the design matrix

$$\begin{bmatrix} [u]_0 & \cdots & [u]_{M-1} & [Y]_0 & \cdots & [Y]_L \\ [u']_0 & \cdots & [u']_{M'-1} & [Y']_0 & \cdots & [Y']_{L'} \end{bmatrix}$$

This is the primary function of an MQU, as discussed in [section 2](#).

Using this finite-dimensional distribution one can always calculate the non-centralized mQU moments

$$(6.2) \quad [\mu_{\mathbf{m}}]_{\mathbf{L}} = \mathbb{E}_{\mathbf{M}-\mathbf{m}} \mathbb{E}_M [Y]_{\mathbf{L}} = \mathbb{E}_{\mathbf{M}-\mathbf{m}} \mathbb{E}_M [y([\mathbf{u}]_{\mathbf{M}})]_{\mathbf{L}}$$

349

$$\begin{aligned} (6.3) \quad [\mu_{\mathbf{m}\mathbf{m}'}]_{\mathbf{L} \times \mathbf{L}'} &= \mathbb{E}_{\mathbf{M}-\mathbf{m}} \mathbb{E}_{\mathbf{M}'-\mathbf{m}'} \mathbb{V}_M [[Y]_{\mathbf{L}}, [Y']_{\mathbf{L}'}] \\ &= \mathbb{E}_{\mathbf{M}-\mathbf{m}} \mathbb{E}_{\mathbf{M}'-\mathbf{m}'} \mathbb{V}_M [k([\mathbf{u}]_{\mathbf{M}}, [\mathbf{u}']_{\mathbf{M}'})]_{\mathbf{L} \times \mathbf{L}'} + [\mu_{\mathbf{m}}]_{\mathbf{L}} \otimes [\mu_{\mathbf{m}'}]_{\mathbf{L}'} \end{aligned}$$

This is how to calculate mQU moments, it is not how to define them.

In order to calculate the Sobol' matrix error, let us define the mQU moments

$$(6.4) \quad [\mu_{\mathbf{m} \dots \mathbf{m}^{(n-1)'}}]_{\mathbf{L}^n} := \mathbb{E}_M [[y_{\mathbf{m}}]_{\mathbf{L}} \otimes \cdots \otimes [y_{\mathbf{m}^{(n-1)'}}]_{\mathbf{L}}]$$

Simply reversing the order in which expectations are taken – using Fubini's theorem – establishes equations (6.2) and (6.3). The law of iterated expectations further entails

$$(6.5) \quad [\mu_{\mathbf{0} \dots \mathbf{0}\mathbf{m}^{j'} \dots \mathbf{m}^{(n-1)'}}]_{\mathbf{L}^n} = \mathbb{E}_{\mathbf{m}} [\mu_{\mathbf{m} \dots \mathbf{m}\mathbf{m}^{j'} \dots \mathbf{m}^{(n-1)'}}]_{\mathbf{L}^n} = \mathbb{E}_{\mathbf{M}} [\mu_{\mathbf{M} \dots \mathbf{M}\mathbf{m}^{j'} \dots \mathbf{m}^{(n-1)'}}]_{\mathbf{L}^n} \quad \forall j < n$$

This reduction will be used repeatedly in the remainder of this section. Definitions (2.2), (3.2), and (5.1) conveniently zero the fully marginalized moments

$$(6.6) \quad \mathbb{V}_{\mathbf{0}}[y_{\mathbf{0}}]_{\mathbf{L}^2} = [V_{\mathbf{0}}]_{\mathbf{L}^2} = [V_{\mathbf{0}}]_{\mathbf{L}^2} = [\mu_{\mathbf{0}}]_{\mathbf{L}}^2 = [\mu_{\mathbf{00}}]_{\mathbf{L}^2} = [0]_{\mathbf{L}^2}$$

Centralizing the mQUs

$$(6.7) \quad [\mathbf{e}_{\mathbf{m}}]_{\mathbf{L}} := [y_{\mathbf{m}}]_{\mathbf{L}} - [\mu_{\mathbf{m}}]_{\mathbf{L}} \quad \forall \mathbf{m} \subseteq \mathbf{M}$$

the expected conditional variance in (5.1) amounts to

$$\begin{aligned} (6.8) \quad [V_{\mathbf{m}}]_{\mathbf{L}^2} &= \mathbb{E}_{\mathbf{m}} \mathbb{E}_M [[\mathbf{e}_{\mathbf{m}} + \mu_{\mathbf{m}}]_{\mathbf{L}}^2] - \mathbb{E}_M [[\mathbf{e}_{\mathbf{0}} + \mu_{\mathbf{0}}]_{\mathbf{L}}^2] \\ &= \mathbb{E}_{\mathbf{m}} [[\mu_{\mathbf{m}}]_{\mathbf{L}}^2] - [\mu_{\mathbf{0}}]_{\mathbf{L}}^2 + \mathbb{E}_{\mathbf{m}} [\mu_{\mathbf{m}\mathbf{m}}]_{\mathbf{L}^2} - [\mu_{\mathbf{00}}]_{\mathbf{L}^2} \\ &= \mathbb{E}_{\mathbf{m}} [[\mu_{\mathbf{m}}]_{\mathbf{L}}^2] \end{aligned}$$

364 Substituting this formula in (5.1a) and (5.2) determines the closed and total Sobol' matrices
 365 in terms of mQU moments.

366 The covariance between conditional variances in (5.5) is

$$\begin{aligned}
 [W_{\mathbf{m}\mathbf{m}'}]_{\mathbf{L}^2 \times \mathbf{L}'^2} &:= \mathbb{V}_M[\mathbb{V}_{\mathbf{m}}[y_{\mathbf{m}}]_{\mathbf{L}^2}, \mathbb{V}_{\mathbf{m}'}[y_{\mathbf{m}'}]_{\mathbf{L}'^2}] \\
 &= \mathbb{V}_M \left[\mathbb{E}_{\mathbf{m}} \left[[y_{\mathbf{m}}]_{\mathbf{L}}^2 - [y_0]_{\mathbf{L}}^2 \right], \mathbb{E}_{\mathbf{m}'} \left[[y_{\mathbf{m}'}]_{\mathbf{L}'}^2 - [y_0]_{\mathbf{L}'}^2 \right] \right] \\
 &= \mathbb{E}_M \left[\mathbb{E}_{\mathbf{m}} \left[[y_{\mathbf{m}}]_{\mathbf{L}}^2 - [y_0]_{\mathbf{L}}^2 \right] \otimes \mathbb{E}_{\mathbf{m}'} \left[[y_{\mathbf{m}'}]_{\mathbf{L}'}^2 - [y_0]_{\mathbf{L}'}^2 \right] \right] - [V_{\mathbf{m}}]_{\mathbf{L}^2} \otimes [V_{\mathbf{m}'}]_{\mathbf{L}'^2} \\
 &= [A_{\mathbf{m}\mathbf{m}'} - A_{\mathbf{0}\mathbf{m}'} - A_{\mathbf{m}\mathbf{0}} + A_{\mathbf{00}}]_{\mathbf{L}^2 \times \mathbf{L}'^2}
 \end{aligned}$$

368 where

$$\begin{aligned}
 [A_{\mathbf{m}\mathbf{m}'}]_{\mathbf{L}^2 \times \mathbf{L}'^2} &:= \mathbb{E}_M \mathbb{E}_{\mathbf{m}'} \mathbb{E}_{\mathbf{m}} \left[[y_{\mathbf{m}}]_{\mathbf{L}}^2 \otimes [y_{\mathbf{m}'}]_{\mathbf{L}'}^2 \right] - [V_{\mathbf{m}}]_{\mathbf{L}^2} \otimes [V_{\mathbf{m}'}]_{\mathbf{L}'^2} \\
 &= \mathbb{E}_{\mathbf{m}} \mathbb{E}_{\mathbf{m}'} \mathbb{E}_M \left[[\mathbf{e}_{\mathbf{m}} + \mu_{\mathbf{m}}]_{\mathbf{L}}^2 \otimes [\mathbf{e}_{\mathbf{m}'} + \mu_{\mathbf{m}'}]_{\mathbf{L}'}^2 - [\mu_{\mathbf{m}}]_{\mathbf{L}}^2 \otimes [\mu_{\mathbf{m}'}]_{\mathbf{L}'}^2 \right]
 \end{aligned}$$

370 exploiting the fact that $V_0 = [0]_{\mathbf{L}^2}$. Equation (6.5) cancels all terms beginning with $[\mathbf{e}_{\mathbf{m}}]_{\mathbf{L}}^2$,
 371 first across $A_{\mathbf{m}\mathbf{m}'} - A_{\mathbf{0}\mathbf{m}'}$ then across $A_{\mathbf{m}\mathbf{0}} - A_{\mathbf{00}}$. All remaining terms ending in $[\mu_{\mathbf{m}'}]_{\mathbf{L}'}^2$ are
 372 eliminated by centralization $\mathbb{E}_M[\mathbf{e}_{\mathbf{m}}] = 0$. Similar arguments eliminate $[\mathbf{e}_{\mathbf{m}'}]_{\mathbf{L}'}^2$ and $[\mu_{\mathbf{m}}]_{\mathbf{L}}^2$.
 373 Effectively then

$$[A_{\mathbf{m}\mathbf{m}'}]_{\mathbf{L}^4} = \sum_{\pi(\mathbf{L}^2)} \sum_{\pi(\mathbf{L}'^2)} \mathbb{E}_{\mathbf{m}} \mathbb{E}_{\mathbf{m}'} [\mu_{\mathbf{m}} \otimes \mu_{\mathbf{m}\mathbf{m}'} \otimes \mu_{\mathbf{m}'}]_{\mathbf{L}^2 \times \mathbf{L}'^2}$$

375 so (6.6) entails

$$(6.9) \quad [W_{\mathbf{m}\mathbf{m}'}]_{\mathbf{L}^4} = \sum_{\pi(\mathbf{L}^2)} \sum_{\pi(\mathbf{L}'^2)} \mathbb{E}_{\mathbf{m}} \mathbb{E}_{\mathbf{m}'} [\mu_{\mathbf{m}} \otimes \mu_{\mathbf{m}\mathbf{m}'} \otimes \mu_{\mathbf{m}'}]_{\mathbf{L}^2 \times \mathbf{L}'^2}$$

377 where each summation is over permutations of tensor axes

$$\pi(\mathbf{L}^2) := \{(\mathbf{L} \times \mathbf{L}''), (\mathbf{L}'' \times \mathbf{L})\} \quad ; \quad \pi(\mathbf{L}'^2) := \{(\mathbf{L}' \times \mathbf{L}'''), (\mathbf{L}''' \times \mathbf{L}')\}$$

379 Substituting (6.9) in (5.4a), (5.4), and (5.6) determines the closed and total Sobol' matrixe
 380 errors in terms of mQU moments.

381 **7. Old.** The correlation between (the vector predictions of) two mQUs is the tensor RV

$$(7.1) \quad [R_{\mathbf{m}\mathbf{m}'}]_{\mathbf{L} \times \mathbf{L}} := \frac{\mathbb{V}_{\mathbf{M}}[y_{\mathbf{m}}, y_{\mathbf{m}'}]_{\mathbf{L} \times \mathbf{L}}}{\mathbb{D}_{\mathbf{m}}[y_{\mathbf{m}}]_{\mathbf{L}} \otimes \mathbb{D}_{\mathbf{m}'}[y_{\mathbf{m}'}]_{\mathbf{L}}} \quad \forall \mathbf{m}, \mathbf{m}' \subseteq \mathbf{M}$$

383 The square of the correlation is a discerning judge of regression models [11] called the
 384 coefficient of determination, which is clearly identical to the closed Sobol' index

$$(7.2) \quad [R_{\mathbf{m}\mathbf{M}}^2]_{l \times l} := [R_{\mathbf{m}\mathbf{M}}]_{l \times l}^2 = [S_{\mathbf{m}}]_{l \times l}$$

Perhaps the most significant use of closed Sobol' indices is to identify a representative reduced model of $m \leq M$ inputs within the full model \mathbf{M} . A closed index close to 1 confirms that the two models make nearly identical predictions. Simplicity and economy (not least of calculation) motivate the adoption of a reduced model, a closed Sobol' index close to 1 is what permits it. This is precisely equivalent to ignoring $\mathbf{M} - \mathbf{m}$ inputs on the grounds that their influence on y – measured by their total index $[S_{\mathbf{M}-\mathbf{m}}^T]_l$ – is close to 0.

Our aim is to compare predictions from a reduced regression model \mathbf{y}_m with those from the full regression model \mathbf{y}_M . Correlation between these predictions is squared – using element-wise (Hadamard) multiplication \circ and division $/$ – to form an RV called the coefficient of determination

$$(7.3) \quad [R_{\mathbf{mM}}^2]_{\mathbf{L}^2} := \frac{\mathbb{V}_{\mathbf{M}}[\mathbf{y}_m, \mathbf{y}_M] \circ \mathbb{V}_{\mathbf{M}}[\mathbf{y}_m, \mathbf{y}_M]}{\mathbb{V}_{\mathbf{m}}[\mathbf{y}_m] \circ \mathbb{V}_{\mathbf{M}}[\mathbf{y}_M]} = \frac{\mathbb{V}_{\mathbf{m}}[\mathbf{y}_m]}{\mathbb{V}_{\mathbf{M}}[\mathbf{y}_M]}$$

However, this is undefined whenever $\mathbb{V}_{\mathbf{M}}[\mathbf{y}_M]_{l \times l'} = 0$, obscuring potentially useful information about $\mathbb{V}_{\mathbf{m}}[\mathbf{y}_m]_{l \times l'}$. Introducing 1-tensors representing the square root diagonal of a covariance matrix

$$(7.4) \quad \left[\sqrt{\mathbb{V}[\cdot, \cdot]_{\mathbf{L}^2}} \right]_l := \mathbb{V}[\cdot, \cdot]_{l^2}^{1/2}$$

the correlation coefficient between output dimensions is

$$(7.5) \quad [R_{\mathbf{m}}]_{\mathbf{L} \times \mathbf{L}'} := \frac{\mathbb{V}_{\mathbf{m}}[\mathbf{y}_m]_{\mathbf{L} \times \mathbf{L}'}}{\sqrt{\mathbb{V}_{\mathbf{m}}[\mathbf{y}_m]_{\mathbf{L}^2}} \otimes \sqrt{\mathbb{V}_{\mathbf{m}}[\mathbf{y}_m]_{\mathbf{L}'^2}}} \quad \forall \mathbf{m} \subseteq \mathbf{M}$$

Let us define the multi-output closed Sobol' index as the product of the full correlation between output dimensions and the coefficient of determination

$$(7.6) \quad [S_{\mathbf{m}}]_{\mathbf{L} \times \mathbf{L}'} := [R_{\mathbf{m}}]_{\mathbf{L} \times \mathbf{L}'} \circ [R_{\mathbf{mM}}^2]_{\mathbf{L} \times \mathbf{L}'}$$

and the multi-output total Sobol' index as its complement

$$(7.7) \quad [S_{\mathbf{M}-\mathbf{m}}^T]_{\mathbf{L} \times \mathbf{L}'} := [S_{\mathbf{M}}]_{\mathbf{L} \times \mathbf{L}'} - [S_{\mathbf{m}}]_{\mathbf{L} \times \mathbf{L}'}$$

These definitions coincide precisely with the traditional Sobol' index along the diagonal $\sqrt{[S_{\mathbf{m}}]_{\mathbf{L}^2}} \circ \sqrt{[S_{\mathbf{m}}]_{\mathbf{L}'^2}}$, which has been very much the focus of prior literature [14, 43, 15]. The off-diagonal elements are bound by the diagonal as

$$(7.8) \quad -[S_{\mathbf{m}}]_{l^2}^{1/2} [S_{\mathbf{m}}]_{l'^2}^{1/2} \leq [S_{\mathbf{m}}]_{l \times l'} = [R_{\mathbf{m}}]_{l \times l'} [S_{\mathbf{m}}]_{l^2}^{1/2} [S_{\mathbf{m}}]_{l'^2}^{1/2} \leq [S_{\mathbf{m}}]_{l^2}^{1/2} [S_{\mathbf{m}}]_{l'^2}^{1/2}$$

To calculate moments over ungoverned noise we use the Taylor series method [19, pp.353], which is valid provided $\mathbb{V}_{\mathbf{M}}[\mathbf{y}_M]_{l^2}$ is well approximated by its mean

$$(7.9) \quad [V_{\mathbf{M}}]_{l^2} := \mathbb{E}_{\mathbf{M}} \mathbb{V}_{\mathbf{M}}[\mathbf{y}_M]_{l^2} \gg |\mathbb{V}_{\mathbf{M}}[\mathbf{y}_M]_{l^2} - [V_{\mathbf{M}}]_{l^2}|$$

415 This provides the mean Sobol' index

$$416 \quad (7.10) \quad [S_{\mathbf{m}}]_{\mathbf{L} \times \mathbf{L}'} := \mathbb{E}_M [S_{\mathbf{m}}]_{\mathbf{L} \times \mathbf{L}'} = \frac{[V_{\mathbf{m}}]_{\mathbf{L} \times \mathbf{L}'}}{\sqrt{[V_{\mathbf{M}}]_{\mathbf{L}^2}} \otimes \sqrt{[V_{\mathbf{M}}]_{\mathbf{L}'^2}}}$$

$$417 \quad (7.11) \quad \text{where } [V_{\mathbf{m}}]_{\mathbf{L} \times \mathbf{L}'} := \mathbb{E}_M \mathbb{V}_{\mathbf{m}}[y_{\mathbf{m}}]_{\mathbf{L} \times \mathbf{L}'} \quad \forall \mathbf{m} \subseteq \mathbf{M}$$

418 with standard deviation due to ungoverned noise of

$$419 \quad (7.12) \quad [T_{\mathbf{m}}]_{\mathbf{L} \times \mathbf{L}'}^2 := \mathbb{V}_M [S_{\mathbf{m}}]_{(\mathbf{L} \times \mathbf{L}')^2} = \frac{[Q_{\mathbf{m}}]_{(\mathbf{L} \times \mathbf{L}')^2}}{[V_{\mathbf{M}}]_{\mathbf{L}^2}^{2/2} \otimes [V_{\mathbf{M}}]_{\mathbf{L}'^2}^{2/2}}$$

420 where improper fractions exponentiate a square root diagonal of $V_{\mathbf{M}}$, and

$$421 \quad (7.13) \quad [Q_{\mathbf{m}}]_{(\mathbf{L} \times \mathbf{L}')^2} := [W_{\mathbf{mm}}]_{(\mathbf{L} \times \mathbf{L}')^2} - [V_{\mathbf{m}}]_{\mathbf{L} \times \mathbf{L}'} \circ \sum_{\mathbf{L}^\circ \in \{\mathbf{L}, \mathbf{L}'\}} \frac{[W_{\mathbf{Mm}}]_{\mathbf{L}^\circ \times \mathbf{L} \times \mathbf{L}'}}{[V_{\mathbf{M}}]_{\mathbf{L}^\circ}^{2/2}} \\ 422 \quad + \frac{[V_{\mathbf{m}}]_{\mathbf{L} \times \mathbf{L}'}}{4} \circ \sum_{\mathbf{L}^\circ \in \{\mathbf{L}, \mathbf{L}'\}} \frac{[W_{\mathbf{MM}}]_{\mathbf{L}^\circ \times \mathbf{L}^2}}{[V_{\mathbf{M}}]_{\mathbf{L}^\circ}^{2/2} \otimes [V_{\mathbf{M}}]_{\mathbf{L}^2}^{2/2}} + \frac{[W_{\mathbf{MM}}]_{\mathbf{L}^\circ \times \mathbf{L}'^2}}{[V_{\mathbf{M}}]_{\mathbf{L}^\circ}^{2/2} \otimes [V_{\mathbf{M}}]_{\mathbf{L}'^2}^{2/2}} \\ 423$$

$$424 \quad (7.14) \quad [W_{\mathbf{mm}'}]_{\mathbf{L}^4} := \mathbb{V}_M [\mathbb{V}_{\mathbf{m}}[y_{\mathbf{m}}], \mathbb{V}_{\mathbf{m}'}[y_{\mathbf{m}'}]]_{\mathbf{L}^4}$$

425 REFERENCES

- 426 [1] *NumPy*, 2023-08-19, <https://numpy.org>.
 427 [2] *PyTorch*, 2023-08-19, <https://pytorch.org>.
 428 [3] *TensorFlow*, 2023-08-19, <https://tensorflow.org>.
 429 [4] M. A. ALVAREZ, L. ROSASCO, AND N. D. LAWRENCE, *Kernels for vector-valued functions: a review*,
 430 2011, <https://arxiv.org/abs/http://arxiv.org/abs/1106.6251v2>.
 431 [5] J. E. ANGUS, *The probability integral transform and related results*, SIAM Review, 36 (1994), pp. 652–654,
 432 <https://doi.org/10.1137/1036146>.
 433 [6] J. K. BLITZSTEIN AND J. HWANG, *Introduction to Probability, Second Edition*, Taylor & Francis Group,
 434 2019.
 435 [7] K. CAMPBELL, M. D. MCKAY, AND B. J. WILLIAMS, *Sensitivity analysis when model outputs are func-*
 436 *tions*, Reliability Engineering & System Safety, 91 (2006), pp. 1468–1472, [https://doi.org/10.1016/j.](https://doi.org/10.1016/j.ress.2005.11.049)
 437 [ress.2005.11.049](https://doi.org/10.1016/j.ress.2005.11.049).
 438 [8] G. CHASTAING, F. GAMBOA, AND C. PRIEUR, *Generalized hoeffding-sobol decomposition for dependent*
 439 *variables -application to sensitivity analysis*, (2011), <https://doi.org/10.48550/ARXIV.1112.1788>,
 440 <https://arxiv.org/abs/1112.1788>.
 441 [9] W. CHEN, R. JIN, AND A. SUDJANTO, *Analytical variance-based global sensitivity analysis in simulation-*
 442 *based design under uncertainty*, Journal of Mechanical Design, 127 (2005), p. 875, [https://doi.org/](https://doi.org/10.1115/1.1904642)
 443 [10.1115/1.1904642](https://doi.org/10.1115/1.1904642).
 444 [10] M. CHEVREUIL, R. LEBRUN, A. NOUY, AND P. RAI, *A least-squares method for sparse low rank ap-*
 445 *proximation of multivariate functions*, SIAM/ASA Journal on Uncertainty Quantification, 3 (2015),
 446 pp. 897–921, <https://doi.org/10.1137/13091899X>, [https://arxiv.org/abs/http://arxiv.org/abs/1305.](https://arxiv.org/abs/http://arxiv.org/abs/1305.0030v2)
 447 [0030v2](https://arxiv.org/abs/http://arxiv.org/abs/1305.0030v2).
 448 [11] D. CHICCO, M. J. WARRENS, AND G. JURMAN, *The coefficient of determination r-squared is more*
 449 *informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation*, PeerJ
 450 Computer Science, 7 (2021-07), p. e623, <https://doi.org/10.7717/peerj-cs.623>.

- [12] P. G. CONSTANTINE, *Active Subspaces: Emerging Ideas for Dimension Reduction in Parameter Studies*, Society for Industrial and Applied Mathematics, mar 2015, <https://doi.org/10.1137/1.9781611973860>.
- [13] C. CORTES AND V. VAPNIK, *Support-vector networks*, Machine Learning, 20 (1995), pp. 273–297, <https://doi.org/10.1007/bf00994018>.
- [14] F. GAMBOA, A. JANON, T. KLEIN, AND A. LAGNOUX, *Sensitivity indices for multivariate outputs*, Comptes Rendus Mathematique, 351 (2013), pp. 307–310, <https://doi.org/10.1016/j.crma.2013.04.016>.
- [15] O. GARCIA-CABREJO AND A. VALOCCHI, *Global sensitivity analysis for multivariate output using polynomial chaos expansion*, Reliability Engineering & System Safety, 126 (2014), pp. 25–36, <https://doi.org/10.1016/j.ress.2014.01.005>.
- [16] R. GHANEM, D. HIGDON, H. OWHADI, ET AL., *Handbook of uncertainty quantification*, vol. 6, Springer, 2017.
- [17] R. G. GHANEM AND P. D. SPANOS, *Spectral techniques for stochastic finite elements*, Archives of Computational Methods in Engineering, 4 (1997), pp. 63–100, <https://doi.org/10.1007/BF02818931>.
- [18] T. HOMMA AND A. SALTELLI, *Importance measures in global sensitivity analysis of nonlinear models*, Reliability Engineering & System Safety, 52 (1996), pp. 1–17, [https://doi.org/10.1016/0951-8320\(96\)00002-6](https://doi.org/10.1016/0951-8320(96)00002-6).
- [19] M. G. KENDALL, *Kendall's advanced theory of statistics*, John Wiley & Sons, 1994.
- [20] D. KHOSHNEVISAN, *Multiparameter Processes*, Springer Monographs in Mathematics Ser., Springer New York, New York, NY, 2002. Description based on publisher supplied metadata and other sources.
- [21] K. KONAKLI AND B. SUDRET, *Global sensitivity analysis using low-rank tensor approximations*, Reliability Engineering & System Safety, 156 (2016), pp. 64–83, <https://doi.org/10.1016/j.ress.2016.07.012>.
- [22] M. LAMBONI, H. MONOD, AND D. MAKOWSKI, *Multivariate sensitivity analysis to measure global contribution of input factors in dynamic models*, Reliability Engineering & System Safety, 96 (2011), pp. 450–459, <https://doi.org/10.1016/j.ress.2010.12.002>.
- [23] B. LAMOUREUX, N. MECBAL, AND J. R. MASSÉ, *A combined sensitivity analysis and kriging surrogate modeling for early validation of health indicators*, Reliability Engineering and System Safety, 130 (2014), pp. 12–26, <https://doi.org/10.1016/j.ress.2014.03.007>.
- [24] A. MARREL, B. IOOSS, B. LAURENT, AND O. ROUSTANT, *Calculations of sobol indices for the gaussian process metamodel*, Reliability Engineering & System Safety, 94 (2009), pp. 742–751, <https://doi.org/10.1016/j.ress.2008.07.008>.
- [25] J. E. OAKLEY AND A. O'HAGAN, *Probabilistic sensitivity analysis of complex models: a bayesian approach*, Journal of the Royal Statistical Society: Series B (Statistical Methodology), 66 (2004), pp. 751–769, <https://doi.org/10.1111/j.1467-9868.2004.05304.x>.
- [26] C. E. RASMUSSEN, *Some useful gaussian and matrix equations*, 2016, <http://mlg.eng.cam.ac.uk/teaching/4f13/1617/gaussian%20and%20matrix%20equations.pdf>.
- [27] C. E. RASMUSSEN AND C. K. I. WILLIAMS, *Gaussian Processes for Machine Learning (Adaptive Computation and Machine Learning series)*, The MIT Press, 2005.
- [28] S. RAZAVI, A. JAKEMAN, A. SALTELLI, C. PRIEUR, B. IOOSS, E. BORGONOVO, E. PLISCHKE, S. L. PIANO, T. IWANAGA, W. BECKER, S. TARANTOLA, J. H. GUILLAUME, J. JAKEMAN, H. GUPTA, N. MELILLO, G. RABITTI, V. CHABRIDON, Q. DUAN, X. SUN, S. SMITH, R. SHEIKHOESLAMI, N. HOSSEINI, M. ASADZADEH, A. PUY, S. KUCHERENKO, AND H. R. MAIER, *The future of sensitivity analysis: An essential discipline for systems modeling and policy support*, Environmental Modelling & Software, 137 (2021-03), p. 104954, <https://doi.org/10.1016/j.envsoft.2020.104954>.
- [29] L. C. G. ROGERS AND D. WILLIAMS, *Diffusions, Markov Processes, and Martingales*, Cambridge University Press, 2000.
- [30] J. SACKS, W. J. WELCH, T. J. MITCHELL, AND H. P. WYNN, *Design and analysis of computer experiments*, Statistical Science, 4 (1989), pp. 409–423.
- [31] A. SALTELLI, K. ALEKSANKINA, W. BECKER, P. FENNEL, F. FERRETTI, N. HOLST, S. LI, AND Q. WU, *Why so many published sensitivity analyses are false: A systematic review of sensitivity analysis practices*, Environmental Modelling & Software, 114 (2019-04), pp. 29–39, <https://doi.org/10.1016/j.envsoft.2019.01.012>.
- [32] E. SCHECHTER, *Handbook of analysis and its foundations*, Academic Press, San Diego, 1997.
- [33] V. SKOROKHOD, *Basic Principles and Applications of Probability Theory*, Springer Berlin Heidelberg,

- 2005.
- [34] I. SOBOL, *Sensitivity estimates for nonlinear mathematical models*, Mathematical Modelling and Computational Experiments, 4 (1993), pp. 407–414.
- [35] I. M. SOBOL, *Global sensitivity indices for nonlinear mathematical models and their monte carlo estimates*, Mathematics and Computers in Simulation, 55 (2001), pp. 271–280.
- [36] I. M. SOBOL AND S. S. KUCHERENKO, *Global sensitivity indices for nonlinear mathematical models. review*, Wilmott, 2005 (2005), pp. 56–61, <https://doi.org/10.1002/wilm.42820050114>.
- [37] A. SRIVASTAVA, A. K. SUBRAMANIYAN, AND L. WANG, *Analytical global sensitivity analysis with gaussian processes*, Artificial Intelligence for Engineering Design, Analysis and Manufacturing, 31 (2017), pp. 235–250, <https://doi.org/10.1017/s0890060417000142>.
- [38] B. SUDRET, *Global sensitivity analysis using polynomial chaos expansions*, Reliability Engineering & System Safety, 93 (2008), pp. 964–979, <https://doi.org/10.1016/j.res.2007.04.002>.
- [39] A. VILLANI, *Another note on the inclusion $L_p(\mu) \subset L_q(\mu)$* , The American Mathematical Monthly, 92 (1985), p. 485, <https://doi.org/10.2307/2322503>.
- [40] P. WIEDERKEHR, *Global sensitivity analysis with dependent inputs*, master's thesis, 2018.
- [41] D. WILLIAMS, *Probability with Martingales*, Cambridge University Press, 1991.
- [42] Z. WU, D. WANG, P. O. N, F. HU, AND W. ZHANG, *Global sensitivity analysis using a gaussian radial basis function metamodel*, Reliability Engineering & System Safety, 154 (2016), pp. 171–179, <https://doi.org/10.1016/j.res.2016.06.006>.
- [43] S. XIAO, Z. LU, AND F. QIN, *Estimation of the generalized sobol's sensitivity index for multivariate output model using unscented transformation*, Journal of Structural Engineering, 143 (2017), [https://doi.org/10.1061/\(asce\)st.1943-541x.0001721](https://doi.org/10.1061/(asce)st.1943-541x.0001721).
- [44] D. XIU, *Numerical Methods for Stochastic Computations: A Spectral Method Approach*, Princeton University Press, 2010.
- [45] D. XIU AND G. E. KARNIAKAKIS, *The wiener–askey polynomial chaos for stochastic differential equations*, SIAM Journal on Scientific Computing, 24 (2002), pp. 619–644, <https://doi.org/10.1137/s1064827501387826>.
- [46] K. ZHANG, Z. LU, K. CHENG, L. WANG, AND Y. GUO, *Global sensitivity analysis for multivariate output model and dynamic models*, Reliability Engineering & System Safety, 204 (2020), p. 107195, <https://doi.org/10.1016/j.res.2020.107195>.