A Unified Framework for Uncertainty and Sensitivity Analysis of Computational Models with Many Input Parameters

Li Gu and C. F. Jeff Wu

H. Milton Stewart School of Industrial and Systems Engineering
Georgia Institute of Technology
Atlanta, Georgia
Email: lgu@gatech.edu and jeffwu@isye.gatech.edu

***Abstract*—Computational models have found wide applications in simulating physical systems. Uncertainties in input parameters of the system can greatly influence the outputs, which are studied by Uncertainty Analysis (UA) and Sensitivity Analysis (SA). As the system becomes more complex, the number of input parameters can be large and existing methods for UA and SA are computationally intensive or prohibitive. We propose a unified framework by using a hierarchical variable selection approach to connect UA and SA with one design. By incorporating the effect hierarchy principle and the effect heredity principle, the method works well especially when the number of input parameters is large. Since the procedure requires only one design, it is economical in run size and computationally efficient.**

Keywords–Uncertainty analysis; Sensitivity analysis; Screening; Effect hierarchy principle; Effect heredity principle; Polynomial chaos expansions.

I. INTRODUCTION

The understanding and analysis of complex physical systems relies increasingly on computer simulations. Simulations are based on computational models with input parameters. In many practical situations, there is uncertainty in the choice of the input parameter values. Thus, the understanding of uncertainties in the input parameters and their impact on the output becomes essential. Two major tools, Uncertainty Analysis (UA) and Sensitivity Analysis (SA), are often employed for this purpose.

UA studies how uncertainties in the input parameters can be mapped to uncertainties in the outputs. A typical method for UA is Monte Carlo simulation, which works as follows. The inputs are determined by generating random samples, following a distribution that characterize the uncertainties of the input parameters. Then analyze the empirical cumulative distribution function of the outputs. See [1] for a review. SA studies how the total output uncertainty can be attributed to uncertainties in each input parameter. It can be done locally or globally, but we only focus on global SA in this paper. Global SA methods can be gathered into two categories: direct computation and metamodeling. Methods in the first category directly use Monte Carlo simulation to compute sensitivity indices, like FAST [2] and Sobol' indices [3]; Methods in the second category build a metamodel to replace the computational model for subsequent statistical analysis. Various metamodels were considered for SA, such as linear regression model [4], Gaussian process model [5], polynomial chaos expansions [6], and smoothing spline [7].

SA becomes complicated when the number of input parameters is not small. Based on the effect sparsity principle [8], only a few (but unknown) parameters are significant among the many candidates. Therefore variable selection, which selects a short list of important parameters, is applied before SA to reduce the computational burden, and to improve the accuracy. Variable selection methods usually need extra function evaluations with the use of special designs, such as systematic fractional replicate design [9], sequential bifurcation [10], or elementary effect method [11].

One can choose the aforementioned methods to sequentially conduct UA, variable selection, and SA. However, this will require too many function evaluations because three separate methods and designs are involved. Especially for variable selection, existing methods are computationally intensive or prohibitive when the number of input parameters is very large and the computational model nonlinear.

In this paper, we propose a new framework, where only one design is used for the three steps. For UA, we run Monte Carlo simulation with Latin hypercube samples [12]. The *Polynomial Chaos* (PC) expansions is used as the metamodel. By keeping the same samples for UA, a variable selection approach designed to handle many input parameters is employed for selecting significant linear, nonlinear and interaction effects. Then, the sensitivity indices (Sobol' indices) of the selected parameters are computed *analytically* from the PC coefficients without any extra function evaluations [6]. Because we need only one design in the three steps, the proposed approach can drastically reduce the computational time in the simulation. Also, it works efficiently for high dimensions (*i.e.*, the number of input parameters p is large). We focus on the selection of input parameters, which is a bridge connecting UA and SA. Details of UA and SA can be found in [1] [6] [13], and are hence omitted.

The rest of the paper is organized as follows. Section II reviews the PC expansions, which is the metamodel we use. Two variable selection methods (the sure independence screening and the lasso) are reviewed in Section III. Section IV introduces a hierarchical variable selection approach which can handle many input parameters. A numerical example is given in Section V. Concluding remarks are given in Section VI.

 ISBN: 978-1-61208-371-1

II. POLYNOMIAL CHAOS EXPANSIONS

Computational models are black-box functions, which may have complex relationships, such as nonlinear effects and interactions. In most cases, the linear model does not work well as a metamodel to approximate computational models. We consider the PC expansions as the metamodel. A brief review of the PC expansions is given in this section.

Denote the computational model by

$$y = f(\boldsymbol{x}), \tag{1}$$

where f is the computer code, $\boldsymbol{x} \in \mathbb{R}^p$ is the p-variate input, and y is the output. Suppose the basis $\Psi(\boldsymbol{x})$ is constructed by the tensor product of orthogonal polynomials $\phi(x_j)$, $j = 1, \ldots, p$, as follows:

$$\Psi_{\boldsymbol{\alpha}}(\boldsymbol{x}) = \phi_{\alpha_1}(x_1) \cdots \phi_{\alpha_p}(x_p), \tag{2}$$

where $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_p)$, and α_j corresponds to the order of ϕ_{α_j}. The order of $\Psi_{\boldsymbol{\alpha}}$ is defined by $|\boldsymbol{\alpha}| = \sum_{j=1}^{p} \alpha_j$. (The choice of orthogonal polynomials varies depending on the distribution of x, which can be found in [14].)

Then, the PC expansions is

$$g(\boldsymbol{x}) = \sum_{k=0}^{\infty} \beta_k \Psi_{\boldsymbol{\alpha}_k}(\boldsymbol{x}), \tag{3}$$

where β_k are unknown coefficients. The order of $g(\boldsymbol{x})$ is defined by $\max_k |\boldsymbol{\alpha}_k|$. Note that $g(\boldsymbol{x})$ with a proper choice of β_k converges in quadratic mean to $f(\boldsymbol{x})$ [15]. The *truncated* PC expansions involving only a finite number of bases is used for practical computations, and is denoted by

$$g_P(\boldsymbol{x}) = \sum_{k=0}^{P} \beta_k \Psi_{\boldsymbol{\alpha}_k}(\boldsymbol{x}), \tag{4}$$

where P is usually decided by restricting the order of the PC expansions.

The coefficients β_k in (4) need to be determined. In stochastic mechanics, stochastic spectral methods using a Galerkin minimization technique that minimizes the residual in the balance equation have been widely used [16]. Such methods are intrusive, because an *ad hoc* modification on the computer code of each problem is required. Some non-intrusive methods were recently proposed for estimating coefficients, where the computation can be done with a set of deterministic model evaluations. Sudret [6] used the regression method. A projection method was discussed in [17]. Throughout this paper, we follow [6] to estimate coefficients. Suppose we have n inputs, $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$. Then $\boldsymbol{\beta} = (\beta_0, \ldots, \beta_P)^T$ in (4) are estimated by minimizing the sum of squares of residuals as follows:

$$\hat{\boldsymbol{\beta}} = \arg\min_{\boldsymbol{\beta}} \sum_{i=1}^{n} \left\{ f(\boldsymbol{x}_i) - \sum_{k=0}^{P} \beta_k \Psi_{\boldsymbol{\alpha}_k}(\boldsymbol{x}_i) \right\}^2. \tag{5}$$

The solution is

$$\hat{\boldsymbol{\beta}} = (\boldsymbol{\Psi}^T \boldsymbol{\Psi})^{-1} \boldsymbol{\Psi}^T \boldsymbol{y}, \tag{6}$$

where $\boldsymbol{\Psi}$ denotes the $n \times P$ matrix with (i,k)th entry $\Psi_{\boldsymbol{\alpha}_k}(\boldsymbol{x}_i)$, and $\boldsymbol{y} = (f(\boldsymbol{x}_1), \ldots, f(\boldsymbol{x}_n))^T$.

Variable selection is necessary when the number of bases P is large. Note that selecting input parameters is equivalent to selecting bases in the PC expansions. Blatman and Sudret [13] proposed a method based on stepwise regression to fit the PC expansions sequentially. A similar idea using least angle regression was discussed in [18]. However, these methods work only if the number of input parameters is small. For high dimensions, if we consider the PC expansions with order L, the number of candidate bases $P = (p+L)!/p!L!$ increases rapidly, and hence the computation becomes burdensome.

III. BRIEF REVIEW OF THE SURE INDEPENDENCE SCREENING AND THE LASSO

In this section, we review two variable selection methods for linear models with high dimensions and moderate dimensions respectively. These methods are embedded in the proposed approach after suitable modifications of the selection of bases. We should point out that the choice is not unique. That is, other variable selection methods [19] [20] [21] that work for linear models can perform the same function in the proposed approach.

Fan and Lv [22] proposed the *Sure Independence Screening* (SIS) on linear models for high dimensional variable selection at a relatively low computational cost. We slightly modify the SIS here to make it applicable on selecting bases, instead of linear predictors. It works as follows. Denote $\Psi_{\boldsymbol{\alpha}_k}(X) = (\Psi_{\boldsymbol{\alpha}_k}(\boldsymbol{x}_1), \ldots, \Psi_{\boldsymbol{\alpha}_k}(\boldsymbol{x}_n))^T$. Compute the correlation between $\Psi_{\boldsymbol{\alpha}_k}(X)$ and $\boldsymbol{y}$, which is given by

$$\omega_k = \mathrm{corr}(\Psi_{\boldsymbol{\alpha}_k}(X), \boldsymbol{y}), \quad \text{for } k = 1, \ldots, P. \tag{7}$$

Then for any given $\gamma \in (0,1)$, the selected model is

$$\mathcal{A}_\gamma = \{1 \le k \le P : |\omega_k| \text{ is among the first } [\gamma n] \text{ largest of all}\}, \tag{8}$$

where $[\gamma n]$ is the integer part of γn. $\gamma n = n - 1$ or $n/\log(n)$ were recommended.

When the dimension is moderate, the lasso [23] is commonly used for variable selection, which can be modified for selecting bases as follows:

$$\hat{\boldsymbol{\beta}} = \arg\min_{\boldsymbol{\beta}} \sum_{i=1}^{n} \left\{ f(\boldsymbol{x}_i) - \sum_{k=0}^{P} \beta_k \Psi_{\boldsymbol{\alpha}_k}(\boldsymbol{x}_i) \right\}^2 + \lambda \sum_{k=0}^{P} |\beta_k|, \tag{9}$$

where λ is a tuning parameter that controls the number of selected bases, and usually is determined by cross-validation. The lasso does both variable selection and shrinkage due to the use of the L_1-penalty (last term in (9)). Note that the lasso is more accurate but computationally much heavier than SIS, and hence is more suitable for a moderate number of candidate bases.

IV. HIERARCHICAL SCREENING METHOD CONNECTING UA AND SA

Using the PC expansions as the metamodel, we propose a variable selection approach for computational models with many input parameters. Since the method connects UA and SA using one design, we call it *Uncertainty-Sensitivity-Analysis* (or abbreviated as USA). The USA has a hierarchical structure with multiple layers. The following two principles give justification for the structure of the approach.

The effect hierarchy principle and the effect heredity principle [8] are commonly considered in variable selection, in

order to reduce computational burden due to a large number of candidate bases. The effect hierarchy principle states that lower order effects are more likely to be important than higher order effects, which gives a reasonable reference for the ranking of importances of bases. By acknowledging the effect hierarchy principle, it becomes possible to arrange the sequence of bases under consideration. The bases that are more likely to be important will be assigned in lower layers, which gives them higher priorities to be tested. The effect heredity principle states that an interaction can be active only if one or all of its parent effects are also active. It can significantly reduce the number of bases to be considered in higher layers, by selecting candidate bases given the selected bases in lower layers. See [24] for a discussion on the principles.

The main idea of the USA is the following. Based on the effect hierarchy principle, the hierarchical structure of candidate bases is built to rank the priorities of the bases with different orders. We do variable selection from lower layers to higher layers. In each layer, depending on the number of candidate bases, the SIS and the lasso are used alternately. Then, the effect heredity principle transfers the selected bases to the next layer, thus connecting any two adjacent layers. The diagram of the USA is sketched in Figure 1.

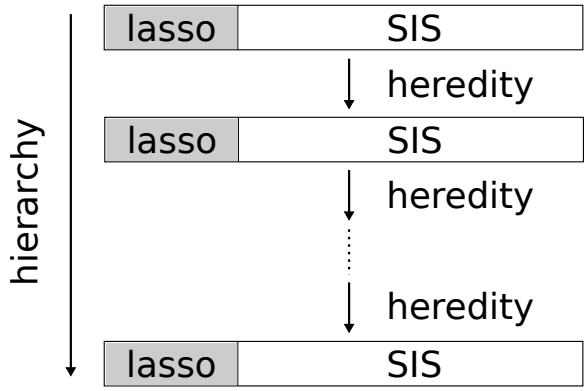

Figure 1. Diagram of the proposed approach.

Let $\mathcal{C} = \{\boldsymbol{\alpha}_k : k = 1, \ldots, P\}$ be the set of the candidate bases. (Note that $k = 0$ is the intercept and is always selected in the model.) Using effect hierarchy, we only consider bases with orders no more than l in the lth layer. Thus, let $\mathcal{M}^{(l)} = \{\boldsymbol{\alpha}_k : k = i_{k_1}, \ldots, i_{k_l};$ and $|\boldsymbol{\alpha}_k| \leq l\}$ be the set of selected bases in the lth layer. Based on $\mathcal{M}^{(l-1)}$, we can generate a set of candidate bases $\mathcal{C}^{(l)}$ for the lth layer. The details will be discussed later. Then, the variable selection within the layer works in the following way. Consider the response $\boldsymbol{y}$. We apply the SIS on $\mathcal{C}^{(l)}$ to select a subset of bases $\mathcal{A}_1^{(l)}$ with a moderate size. Apply the lasso on $\mathcal{A}_1^{(l)}$ to obtain the set of selected bases $\mathcal{S}_1^{(l)}$. In the mth step, we have $\mathcal{S}_{m-1}^{(l)}$ selected in the previous steps. The response is changed to the residual $\boldsymbol{r}_{m-1}$, which is a vector with ith entry $f(\boldsymbol{x}_i) - \sum_{\boldsymbol{\alpha}_k \in \mathcal{S}_{m-1}^{(l)}} \hat{\beta}_k \Psi_{\boldsymbol{\alpha}_k}(\boldsymbol{x}_i)$. $\hat{\beta}$ is estimated by (5) and (6) with current selected bases. We then apply the SIS on $\mathcal{C}^{(l)} \setminus \mathcal{S}_{m-1}^{(l)}$ to get $\mathcal{A}_m^{(l)}$. The set of selected bases $\mathcal{S}_m^{(l)}$ is obtained by using the lasso on $\mathcal{S}_{m-1}^{(l)} \cup \mathcal{A}_m^{(l)}$. Keep doing this until some pre-specified stopping rule is satisfied. One reasonable choice is $\mathcal{S}_{m-1}^{(l)} = \mathcal{S}_m^{(l)}$. The set of bases that are selected when it stops is denoted by $\mathcal{S}^{(l)}$.

We connect two adjacent layers by incorporating two versions of the effect heredity principle [8] [25]. For an interaction to be active, at least one of its parents effects is required to be active in *weak heredity*; while all of its parents effects have to be active in *strong heredity*. Generally, there is no guideline on which version should be used in each specific application [24]. In the USA, the two versions are jointly utilized in a natural way for the purpose of simplifying screening without losing much accuracy. The move between two layers consists of two parts. First, at the beginning of each layer, the set of candidate bases $\mathcal{C}^{(l)}$ is generated by expanding $\mathcal{M}^{(l-1)}$ following weak heredity, which is given by

$$\mathcal{C}^{(l)} = \{\boldsymbol{\alpha}_k : \exists\ \boldsymbol{\alpha}' \in \mathcal{M}^{(l-1)}, \boldsymbol{\alpha}'' \in \mathcal{C} \text{ s.t. } \boldsymbol{\alpha}' + \boldsymbol{\alpha}'' = \boldsymbol{\alpha}_k; \text{ and } |\boldsymbol{\alpha}_k| \leq l\}. \tag{10}$$

The reason is that it is very likely that some of the parent effects of significant interactions are not significant in lower layers. Second, at the end of each layer, we expand $\mathcal{S}^{(l)}$ following strong heredity. Denote the expanded set by

$$\mathcal{D}^{(l)} = \{\boldsymbol{\alpha}_k : \exists\ \boldsymbol{\alpha}', \boldsymbol{\alpha}'' \in \mathcal{S}^{(l)} \text{ s.t. } \boldsymbol{\alpha}' + \boldsymbol{\alpha}'' = \boldsymbol{\alpha}_k; \text{ and } |\boldsymbol{\alpha}_k| \leq l\}. \tag{11}$$

Then use the lasso on $\mathcal{D}^{(l)}$ to screen and fit with the response $\boldsymbol{y}$. By doing so, the interactions between significant bases, which are likely to be significant but have been screened out by a quick variable selection method like the SIS can be re-examined by a more elaborate and accurate variable selection method like the lasso. Denote the set of the final selected bases by $\mathcal{M}^{(l)}$, which will be used to generate $\mathcal{C}^{(l+1)}$ in the next layer. The USA stops when it finishes the highest layer L.

The choice of L is straightforward. Note that the performance of the USA increases as L increases, so does the computational burden. Thus, one can keep increasing the number of layers until enough data is explained by the selected model (which can be assessed by some criterion, such as the coefficient of determination, the root-mean-square error, and so on) or the computational resource runs out. Based on our experience, $L = 3, 4, 5$ are good choices.

The USA is summarized by the following pseudo-code.

Set $\mathcal{M}^{(0)} = \varnothing$
for $l = 1, 2, \ldots, L$
  Expand $\mathcal{M}^{(l-1)}$ by (10) to obtain $\mathcal{C}^{(l)}$
  Set $\boldsymbol{r}_0 = \boldsymbol{y}$ and $\mathcal{S}_0^{(l)} = \varnothing$
  for $m = 1, 2, \ldots$
    Set $\boldsymbol{r}_{m-1}$ to be the response
    Apply the SIS on $\mathcal{C}^{(l)} \setminus \mathcal{S}_{m-1}^{(l)}$ to obtain $\mathcal{A}_m^{(l)}$
    Apply the lasso on $\mathcal{S}_{m-1}^{(l)} \cup \mathcal{A}_m^{(l)}$ to obtain $\mathcal{S}_m^{(l)}$
  endfor if converge; denote the last $\mathcal{S}_m^{(l)}$ by $\mathcal{S}^{(l)}$
  Expand $\mathcal{S}^{(l)}$ by (11) to obtain $\mathcal{D}^{(l)}$
  Set $\boldsymbol{y}$ to be the response
  Apply the lasso on $\mathcal{D}^{(l)}$ to obtain $\mathcal{M}^{(l)}$
endfor

 ISBN: 978-1-61208-371-1

V. A SIMULATION STUDY

We use the Morris function [11] to study the performance of the USA.

$$\begin{aligned} y &= \beta_0 + \sum_{i=1}^{20} \beta_i w_i + \sum_{i<j}^{20} \beta_{ij} w_i w_j \\ &\quad + \sum_{i<j<l}^{20} \beta_{ijl} w_i w_j w_l + \sum_{i<j<l<s}^{20} \beta_{ijls} w_i w_j w_l w_s, \end{aligned} \tag{12}$$

where

$$w_i = \begin{cases} 2(1.1X_i/(X_i + 0.1) - 0.5), & \text{for } i = 3, 5, 7, \\ 2(X_i - 0.5), & \text{otherwise,} \end{cases} \tag{13}$$

and $X_i \sim U(0, 1)$. The β_i are assigned as

$$\begin{array}{ll} \beta_i = 20, & \text{for } i = 1, \ldots, 10, \\ \beta_{ij} = -15, & \text{for } i, j = 1, \ldots, 6, \\ \beta_{ijl} = -10, & \text{for } i, j, l = 1, \ldots, 5, \\ \beta_{ijls} = 5, & \text{for } i, j, l, s = 1, \ldots, 4. \end{array} \tag{14}$$

The remaining first- and second-order coefficients are independently generated from a standard Gaussian distribution; the remaining third- and fourth-order coefficients are zero.

The true model has 20 X_i input parameters, out of which the first ten are significant. 980 dummy parameters were added. So we had $p = 1000$ parameters in total. n simulations were conducted each time. The design matrix was an arbitrary Latin hypercube design with n rows and p columns.

For the SIS, $[\gamma n]$ was set to be $[n/\log(n)] = 81$. For the lasso, the number of selected bases was chosen by cross-validation. We use one simulation with $n = 500$ to illustrate how the USA works. Step-by-step details are given in Table I. For simplicity, only the results in the first and the final iterations in each layer are listed. From Table I, all significant input parameters had been already selected in the third layer. Although we can still try more layers to get better performance of the PC expansions, the PC expansions with order three can explain more than 90% variation of the data [13]. Thus, we chose three as the highest layer L in the example. Another important point is that strong heredity at the end of the second and third layers played a key role in selecting significant bases and removing insignificant bases.

The results of 20 replications with $n = 300, 500, 1000$ are listed in Table II, which show the average values of the numbers of true positives and false positives. Standard deviations are given in parentheses. The sensitivities indices can be computed analytically after all significant input parameters are found. See [13].

TABLE II. RESULTS FOR MORRIS FUNCTION WITH $p = 1000$

n	*True Positives*	*False Positives*
300	7.11(1.66)	9.58(4.05)
500	9.26(0.87)	2.95(2.83)
1000	10(0)	1.36(0.56)

By comparing the results with different n, it is clearly seen that as n increases, the performance became better. With $n = 500$ runs, the USA can identify almost all significant input parameters. The false positives, at the same time, were very low. Moreover, the USA exactly identified all significant variables each time when $n = 1000$, which is still a light computation.

We compare the performance of the USA with that of the Morris method [11]. The Morris method requires at least $r(p+1)$ function evaluations, where r is the number of levels of each input parameter. Based on the results of 20 replications, the Morris method identified 99.5% significant variables when $r = 4$, $n = 4004$. When $r = 2$, $n = 2002$ (the minimal number of runs required), only 75.5% significant variables were identified. In contrast, the USA with $n = 1000$ performed perfectly. Considering the design used in the Morris method is a special design, which cannot be used for neither UA nor SA, the USA is much better than the Morris method in terms of saving functional evaluations.

VI. CONCLUSION

In this paper, we have proposed a hierarchical variable selection dubbed the USA, for computational models with many input parameters. The major feature of the USA is that it unifies UA and SA with the use of one design, which significantly reduces the number of simulations required. By incorporating the effect hierarchy principle and the effect heredity principle, the number of candidate bases in each layer is reduced to an acceptable level for variable selection. By applying the SIS and the lasso alternately, the USA balances the computation and the performance. The performance was shown with a numerical example.

As we mentioned in Section II, the coefficients in the PC expansions can be estimated in several ways. Although we only consider the regression method, the USA is independent of the method for estimating coefficients, and hence can be extended to other methods, which is left for future research.

(Note: An expanded version of this paper, including an application to building energy simulation, will be forthcoming as a joint work with Yuming Sun and Godfried Augenbroe.)

ACKNOWLEDGMENT

This work is supported by NSF-EFRI-SEED Award 1038248 and DOE grant DE-SC 0010548.

REFERENCES

[1] J. C. Helton and F. J. Davis, "Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems," Reliability Engineering & System Safety, vol. 81, no. 1, 2003, pp. 23–69.

[2] R. Cukier, C. Fortuin, K. E. Shuler, A. Petschek, and J. Schaibly, "Study of the sensitivity of coupled reaction systems to uncertainties in rate coefficients. i theory," The Journal of Chemical Physics, vol. 59, no. 8, 1973, pp. 3873–3878.

[3] I. M. Sobol', "Sensitivity estimates for nonlinear mathematical models," Mathematical Modeling and Computational Experiment, vol. 1, no. 4, 1993, pp. 407–414.

[4] A. Saltelli, K. Chan, and E. M. Scott, Sensitivity analysis. Wiley New York, 2000, vol. 134.

[5] 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), vol. 66, no. 3, 2004, pp. 751–769.

[6] B. Sudret, "Global sensitivity analysis using polynomial chaos expansions," Reliability Engineering & System Safety, vol. 93, no. 7, 2008, pp. 964–979.

[7] S. Touzani and D. Busby, "Smoothing spline analysis of variance approach for global sensitivity analysis of computer codes," Reliability Engineering & System Safety, vol. 112, 2013, pp. 67–81.

[8] C. F. J. Wu and M. S. Hamada, Experiments: planning, analysis, and optimization. John Wiley & Sons, 2009, vol. 552.

 ISBN: 978-1-61208-371-1

TABLE I. ILLUSTRATION OF THE USA WITH $n = 500$ AND $p = 1000$

Layer	*Stage*	*Bases*
1	Weak heredity	N/A; $\mathcal{C}^{(1)} = 1000$ bases
1	First: SIS	$\mathcal{A}_1^{(1)} = 81$ bases
1	First: Lasso	$\mathcal{S}_1^{(1)} = 19$ bases: 3, 5, 7, 8, 9, 10, 93, 203, 284, 286, 362, 434, 510, 511, 623, 815, 896, 940, 941
1	Last: SIS, Lasso	$\mathcal{S}^{(1)} = 17$ bases: 3, 5, 7, 8, 9, 10, 93, 203, 284, 286, 362, 434, 510, 511, 556, 896, 966
1	Strong heredity	N/A; $\mathcal{M}^{(1)} = \mathcal{S}^{(1)}$
2	Weak heredity	$\mathcal{C}^{(2)} = 16711$ bases
2	First: SIS	$\mathcal{A}_1^{(2)} = 81$ bases
2	First: Lasso	$\mathcal{S}_1^{(2)} = 23$ bases: 5, 7, 8, 9, 10, 286, 1&3, 1&5, 2&5, 3&4, 3&6, 4&6 161&284, 163&556, 167&896, 198&284, 198&544, 238&284, 284&858, 284&885, 510&592, 510&966, 673&792
2	Last: SIS, Lasso	$\mathcal{S}^{(2)} = 19$ bases: 5, 7, 8, 9, 10, 286, 1&3, 1&5, 2&5, 3&4, 3&6, 93&108, 161&284, 163&556, 167&896, 198&284, 238&284, 510&592, 510&869
2	Strong heredity	$\mathcal{M}^{(2)} = 16$ bases: 7, 8, 9, 10, 1&2, 1&3, 1&5, 1&6, 2&4, 2&5, 2&6, 3&4, 3&6, 4&6, 7&7, 163&556
3	Weak heredity	$\mathcal{C}^{(3)} = 15744$ bases
3	First: SIS	$\mathcal{A}_1^{(3)} = 81$ bases
3	First: Lasso	$\mathcal{S}_1^{(3)} = 21$ bases: 7, 8, 9, 10, 1&2, 1&3, 1&4, 1&5, 1&6, 2&4, 2&5, 2&6, 3&4, 3&6, 4&5, 4&6, 5&6, 7&7, 3&4&421, 5&574, 7&32
3	Last: SIS, Lasso	$\mathcal{S}^{(3)} = 16$ bases: 7, 8, 9, 10, 1&2, 1&3, 1&4, 1&5, 1&6, 2&4, 2&5, 2&6, 3&4, 3&6, 4&6, 7&7
3	Strong heredity	$\mathcal{M}^{(3)} = 20$ bases: 7, 8, 9, 10, 1&2, 1&3, 1&4, 1&5, 1&6, 2&4, 2&5, 2&6, 3&4, 3&6, 4&6, 7&7, 3&4&5, 3&5&5, 7&7&7, 10&10&10

[9] S. C. Cotter, "A screening design for factorial experiments with interactions," Biometrika, vol. 66, no. 2, 1979, pp. 317–320.

[10] B. Bettonvil, Detection of important factors by sequential bifurcation. Tilburg University Press, 1990.

[11] M. D. Morris, "Factorial sampling plans for preliminary computational experiments," Technometrics, vol. 33, no. 2, 1991, pp. 161–174.

[12] M. D. McKay, R. J. Beckman, and W. J. Conover, "Comparison of three methods for selecting values of input variables in the analysis of output from a computer code," Technometrics, vol. 21, no. 2, 1979, pp. 239–245.

[13] G. Blatman and B. Sudret, "Efficient computation of global sensitivity indices using sparse polynomial chaos expansions," Reliability Engineering & System Safety, vol. 95, no. 11, 2010, pp. 1216–1229.

[14] D. Xiu and G. E. Karniadakis, "The wiener–askey polynomial chaos for stochastic differential equations," SIAM Journal on Scientific Computing, vol. 24, no. 2, 2002, pp. 619–644.

[15] C. Soize and R. Ghanem, "Physical systems with random uncertainties: chaos representations with arbitrary probability measure," SIAM Journal on Scientific Computing, vol. 26, no. 2, 2004, pp. 395–410.

[16] R. G. Ghanem and P. D. Spanos, Stochastic finite elements: a spectral approach. Springer, 1991, vol. 41.

[17] T. Crestaux, O. Le Maître, and J.-M. Martinez, "Polynomial chaos expansion for sensitivity analysis," Reliability Engineering & System Safety, vol. 94, no. 7, 2009, pp. 1161–1172.

[18] G. Blatman and B. Sudret, "Adaptive sparse polynomial chaos expansion based on least angle regression," Journal of Computational Physics, vol. 230, no. 6, 2011, pp. 2345–2367.

[19] J. Fan and R. Li, "Variable selection via nonconcave penalized likelihood and its oracle properties," Journal of the American Statistical Association, vol. 96, no. 456, 2001, pp. 1348–1360.

[20] H. Zou, "The adaptive lasso and its oracle properties," Journal of the American Statistical Association, vol. 101, no. 476, 2006, pp. 1418–1429.

[21] E. Candes and T. Tao, "The dantzig selector: statistical estimation when p is much larger than n," The Annals of Statistics, 2007, pp. 2313–2351.

[22] J. Fan and J. Lv, "Sure independence screening for ultrahigh dimensional feature space," Journal of the Royal Statistical Society: Series B (Statistical Methodology), vol. 70, no. 5, 2008, pp. 849–911.

[23] R. Tibshirani, "Regression shrinkage and selection via the lasso," Journal of the Royal Statistical Society: Series B (Methodological), vol. 58, no. 1, 1996, pp. 267–288.

[24] M. Yuan, V. R. Joseph, and Y. Lin, "An efficient variable selection approach for analyzing designed experiments," Technometrics, vol. 49, no. 4, 2007, pp. 430–439.

[25] H. Chipman, "Bayesian variable selection with related predictors," Canadian Journal of Statistics, vol. 24, no. 1, 1996, pp. 17–36.

 ISBN: 978-1-61208-371-1