Application of Quasi Monte Carlo methods for evaluation of derivative based global sensitivity measures

I.M. Sobol'[a], S. Kucherenko[b], D. Asotsky[a]

[a]*Institute for Mathematical Modelling,*
the Russian Academy of Sciences, Russia

[b]*Joint Research Centre of the European Commission, Italy*
Imperial College London, UK

Outline

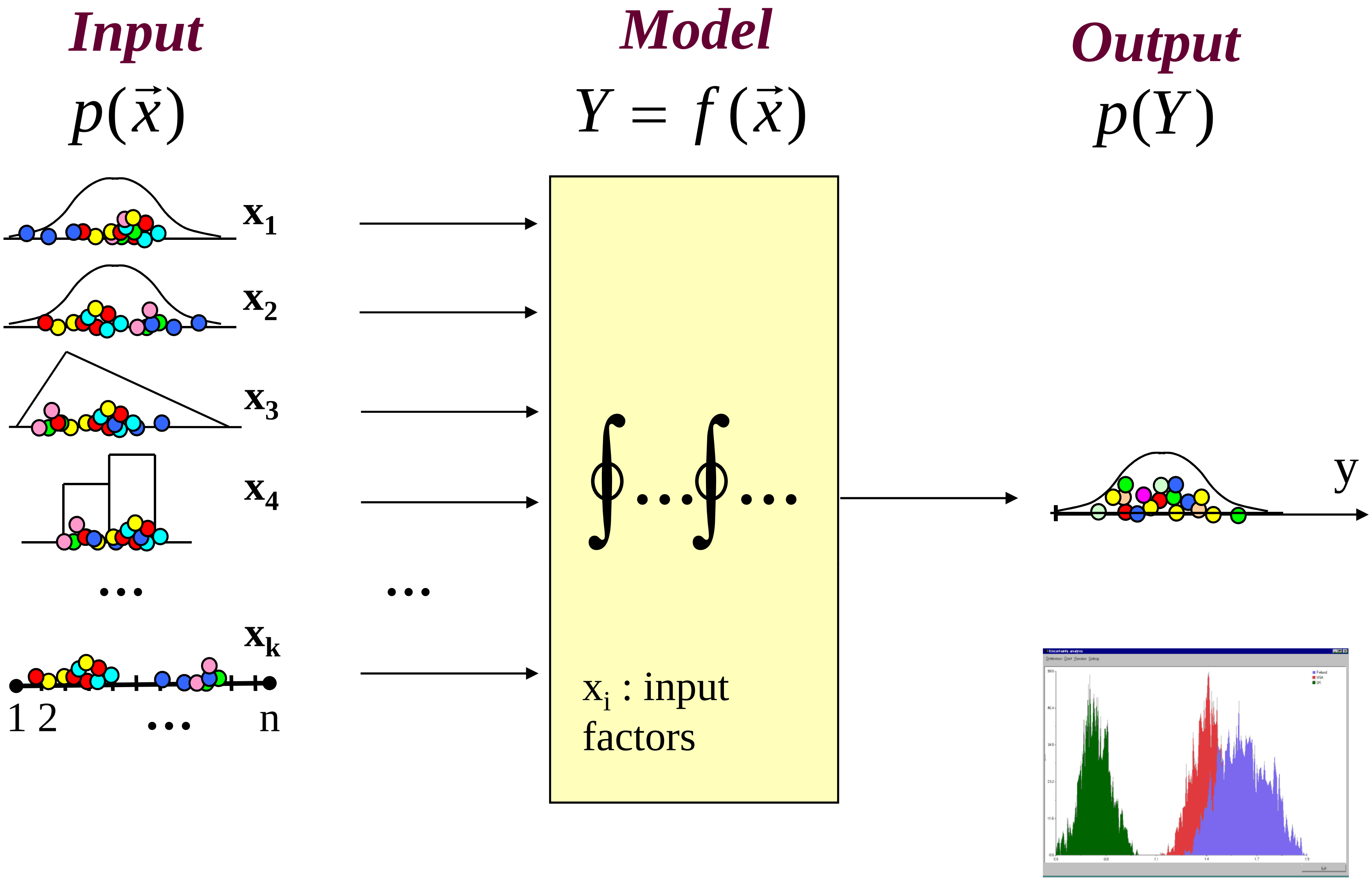
Global Sensitivity Analysis and Sobol' Sensitivity Indices

Derivative based Global Sensitivity Measures

Link between Sobol' Sensitivity Indices and Derivative based Global Sensitivity Measures

Sobol' sequence generators satisfying uniformity properties A and A' and thier comparison with other generators

Propagation of uncertainty



ANOVA decomposition and Sobol' Sensitivity Indices

Consider a model

x is a vector of input variables

Y is the model output, $f(x) \in L_2$

$$Y = f(x)$$

$$x = (x_1, x_2, \ldots, x_k) \in H^k$$

ANOVA decomposition:

$$Y = f(x) = f_0 + \sum_{i=1}^{k} f_i(x_i) + \sum_i \sum_{j>i} f_{ij}(x_i, x_j) + \ldots + f_{1,2,\ldots,k}(x_1, x_2, \ldots, x_k),$$

$$\int_0^1 f_{i_1 \ldots i_s}(x_{i_1}, , \ldots, x_{i_s}) \quad dx_{i_k} = 0, \quad \forall k, \; 1 \le k \le s$$

Variance decomposition: $D^2 = \sum_i D_i^2 + \sum_{i,j} D_{ij}^2 + \ldots + D_{1,2,\ldots,n}^2$

Sobol' SI: $1 = \sum_{i=1}^{k} S_i + \sum_{i<j} S_{ij} + \sum_{i<j<l} S_{ijl} + \ldots + S_{1,2,\ldots,k}$

Sobol' Sensitivity Indices (SI)

- **Definition:** $$S_{i_1 \ldots i_s} = D^2_{i_1 \ldots i_s} / D^2$$

$$D^2_{i_1 \ldots i_s} = \int_0^1 f^2_{i_1 \ldots i_s}\left(x_{i_1}, \ldots, x_{is}\right) dx_{i_1}, \ldots, x_{is}$$ **- partial variances**

$$D^2 = \int_0^1 \left(f\left(x\right) - f_0\right)^2 dx$$ **- variance**

Requires 2^n integral evaluations for calculations $D^2_{i_1 \ldots i_s}$

- **Sensitivity indices for subsets of variables:** $x = \left(y, z\right)$

$$D_y^2 = \sum_{s=1}^{m} \sum_{(i_1 \langle \ldots \langle i_s) \in K} D^2_{i_1, \ldots, i_s}$$

Total Sensitivity indices: $$\left(D_y^{tot}\right)^2 = D_y^2 + D_{yz}^2$$

Sobol' sensitivity indices:

$$S_y = D_y^2 / D^2, \qquad S_y^{tot} = \left(D_y^{tot}\right)^2 / D^2.$$

How to use Sobol' Sensitivity Indices?

$$0 \le S_y \le S_y^{tot} \le 1$$

- $S_y^{tot} - S_y$ ***accounts for all interactions between y and z, x=(y,z).***
- ***The important indices in practice are*** S_i ***and*** S_i^{tot}

  $S_i^{tot} = 0 \to f(x)$ **does not depend on** x_i **;**

  $S_i = 1 \to f(x)$ **only depends on** x_i **;**

  $S_i = S_i^{tot}$ ***corresponds to the absence of interactions between*** x_i **and other variables**

  If $\sum_{s=1}^{n} S_i = 1,$ **then function has additive structure:** $f(x) = f_0 + \sum_i f_i(x_i)$

- ***Fixing unessential variables***

  If $S_z^{tot} << 1 \to f(x)$ **does not depend on** z **so it can be fixed**

  $f(x) \approx f(y, z_0) \to$ **complexity reduction, from** n **to** $n - n_z$ **variables**

Evaluation of Sobol' Sensitivity Indices

Straightforward use of Anova decomposition requires

2^n integral evaluations – not practical !

There are efficient formulas for evaluation of Sobol' Sensitivity Indices (Sobol' 2001,2007,2009, Saltelli 2002, Kucherenko 2002) :

$$S_y = \frac{1}{D}[\int_0^1 f(y,z) f(y,z')\, dydzdz' - f_0^2],$$

$$S_y^{tot} = \frac{1}{2D}\int_0^1 [f(y,z) - f(y',z)]^2\, dydzdz',$$

$$D = \int_0^1 f^2(y,z)\, dydz - f_0^2$$

Evaluation is reduced to high-dimensional integration by MC/QMC methods.

The number of function evaluations is N(n+2)

Pros and Cons of Variance-based Sensitivity Indices

Pros: Variance-based methods offer a comprehensive approach to the model analysis.

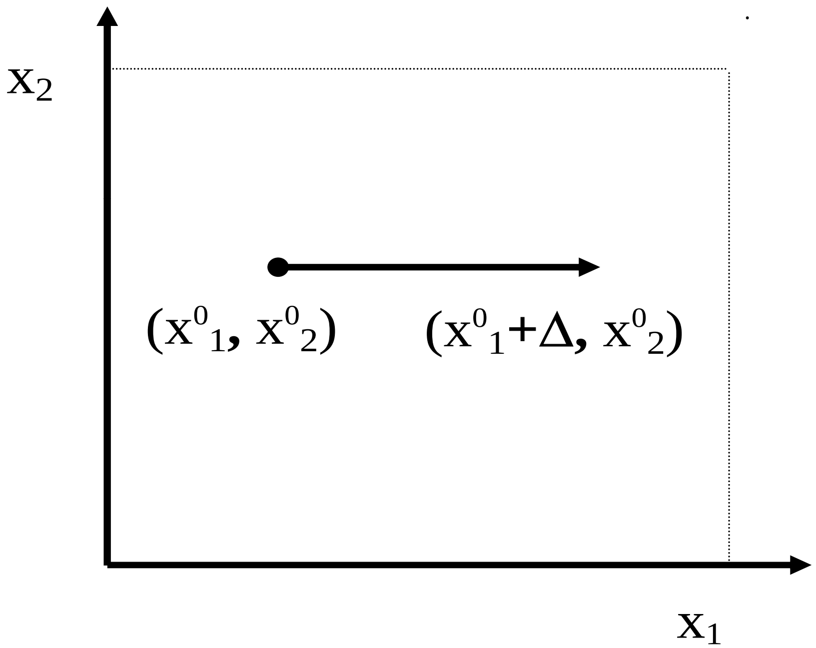
Cons: Generally require a large number of function evaluations to achieve reasonable convergence. It can become impractical for complex high dimensional problems.

The Morris screening method

Model $f = f(x_1, .., x_k)$

Elementary Effect for the i^th input factor at point X^o

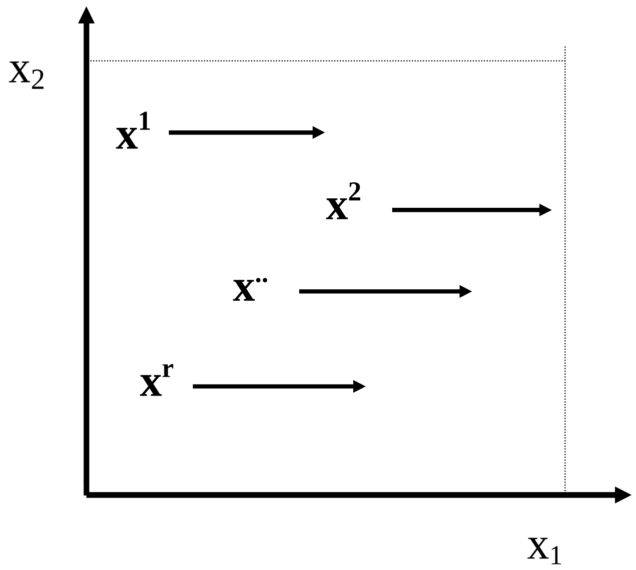
$$EE_i(x_1^0, \ldots, x_k^0) = \frac{f(x_1^0, x_2^0, .., x_{i-1}^0, x_i^0 + \Delta, x_{i+1}^0, .., x_k^0) - f(x_1^0, \ldots, x_k^0)}{\Delta}$$



Δ is large (~1)

The Elementary Effect EE_i is still a local measure

Solution: take the average of several EE



r elem. effects EE^1_i EE^2_i … EE^r_i are computed at $\mathbf{X}^1$, … , $\mathbf{X}^r$ and then averaged.

Average of $|EE_i|$ → $\mu^*(x_i)$

Derivative based Global Sensitivity Measures

Morris measure in the limit $\Delta \to 0$

$$E_i(x_1^0, \ldots, x_k^0) = \lim_{\Delta \to 0} EE_i(x_1^0, \ldots, x_k^0).$$

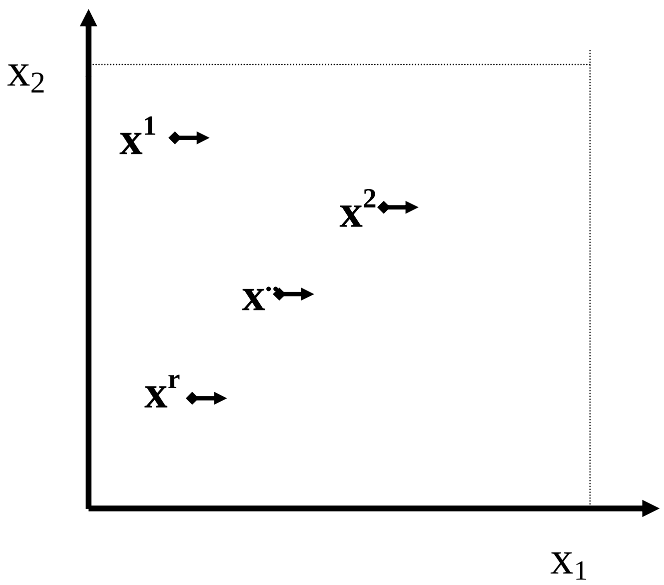
$$E_i(x_1^0, \ldots, x_k^0) = \frac{\partial f}{\partial x_i}.$$

$$M_i^* = \int_{H^n} |E_i| \, dx.$$

Consider new measure

$$\nu_i = \int_{H^n} \left(\frac{\partial f}{\partial x_i} \right)^2 dx$$

Implemention:



Sample $\mathbf{X}_1, \ldots, \mathbf{X}_r$ Sobol (or random) points, estimate $\frac{\partial f}{\partial x_i}$ numerically or using automated differentiation; $(E^1_i)^2$ and average.

Recall: Formula for evaluation total sensitivity indices

$x = (x_1, \ldots, x_n)$. Consider an arbitrary subset

$y = (x_{i_1}, \ldots, x_{i_s}), 1 \le s < n,$ the decomposition $\rightarrow x = (y, z)$

$$S_y^{tot} = \frac{1}{2D} \int_0^1 [f(y,z) - f(y',z)]^2 \, dy dz dz',$$

$$D = \int_0^1 f^2(y,z) \, dy dz - f_0^2$$

if $y = x_i$

$$S_i^{tot} = \frac{1}{2D} \int_{H^n} \int_0^1 \left[f(x) - f\left(\overset{\circ}{x}\right) \right]^2 dx dx_i'$$

$$\overset{\circ}{x} = (x_1, \ldots, x_{i-1}, x_i', x_{i+1}, \ldots, x_n)$$

Evaluation is reduced to high-dimensional integration

Derivative based global sensitivity measures. Theorems

Theorem 1. Assume that $c \le \left| \frac{\partial f}{\partial x_i} \right| \le C$. Then

$$\frac{c^2}{12D} \le S_i^{tot} \le \frac{C^2}{12D}$$

Proof:

$$D_i^{tot} = \frac{1}{2} \int_{H^n} \int_0^1 \left[f(x) - f\left(\overset{\circ}{x}\right) \right]^2 dx dx_i'$$

$$f(x) - f\left(\overset{\circ}{x}\right) = \frac{\partial f(\hat{x})}{\partial x_i}(x_i - x_i')$$

where is $\hat{x}$ a point between x and $\overset{\circ}{x}$

Derivative based global sensitivity measures ν_i. Theorems

Theorem 2. $S_i^{tot} \le \dfrac{\nu_i}{\pi^2 D}$

Pr oof:

Consider

$$f(x) = f_0 + \sum_i f_i(x_i) + \sum_{i<j} f_{i,j}(x_i, x_j) + \dots + f_{1,2,\dots,n}(x_1, x_2, \dots, x_n)$$

denote

$$u(x) = \sum_{<i>} f_{i_1,\dots,i_s}(x_{i_1}, \dots, x_{i_s})$$

the sum of all terms in ANOVA that depend on x_i :

Then $D_i^{tot} = \int\limits_{H^n} u^2(x)\,dx$ and $\dfrac{\partial f}{\partial x_i} = \dfrac{\partial u}{\partial x_i}$

Using $\int\limits_0^1 u^2(x)\,dx_i \le \dfrac{1}{\pi^2} \int\limits_0^1 \left(\dfrac{\partial u}{\partial x_i} \right)^2 dx_i \rightarrow \square$

Functions with separated variables

Consider $f(x)=\prod_{i=1}^{n}\varphi_i(x_i)$

Then $\dfrac{\nu_i}{D_i^{tot}}=\dfrac{\int_0^1\left[\varphi_i'(t)\right]^2 dt}{D_i}$

Example 1. g – function $f=\prod_{i=1}^{n}\dfrac{|4x_i-2|+a_i}{1+a_i}$

$$\frac{\nu_i}{D_i^{tot}}=48.$$

It does not depend on the parameter a_i !

Example 2. Assume that $\varphi_i(t)=t^m$

$$\frac{\int_0^1\left[\varphi_i'(t)\right]^2 dt}{D_i}=(m+1)^2\,\frac{2m+1}{2m-1},$$

At $m=1$ the ratio $\dfrac{\nu_i}{D_i^{tot}}$ is 12, for large m it will be $\approx(m+1)^2$.

Counterexample

$$f = \sum_{i=1}^{4} c_i \left(x_i - \frac{1}{2} \right) + c_{12} \left(x_1 - \frac{1}{2} \right) \left(x_2 - \frac{1}{2} \right)^5$$

$-$ANOVA decomposition

Here $c_i = 1,\ 1 \le i \le 4,\ c_{12} = 50$

Then $S_i = 0.237,\ 1 \le i \le 4$

$S_{12} = 0.0523$, and

$S_1^{tot} = S_2^{tot} = 0.289$ – equally important

$S_3^{tot} = S_4^{tot} = 0.237$ – equally important

$\nu_1 = 1.22,\ \nu_2 = 3.26,\ \nu_3 = \nu_4 = 1.0$

$\rightarrow \nu_2 > \nu_1 + \nu_3 + \nu_4$

ranking of influential variables based on DGSM

can be very different from that based on S_i^{tot}

Derivative based global sensitivity measures for groups of variables

$x = (x_1, \ldots, x_n)$

Consider an arbitrary subset

$y = (x_{i_1}, \ldots, x_{i_s}), 1 \le s < n,$ the decomposition $\rightarrow x = (y, z)$

The total variance: $D = D_y + D_z + D_{yz}$

$D_y^{tot} = D_y + D_{yz}$

$D_z^{tot} = D_z + D_{yz}$

$$D_y^{tot} = \frac{1}{2} \int \left[f(y', z) - f(x) \right]^2 dx dy'$$

Consider theTaylor expansion

$$f(y', z) - f(y, z) = \sum_{p=1}^{s} \frac{\partial f(x)}{\partial x_{i_p}} (x'_{i_p} - x_{i_p}) + \ldots$$

DGSM for groups of variables

$$\hat{\tau}_y = \frac{1}{2}\int \left[\sum_{p=1}^{s} \frac{\partial f(x)}{\partial x_{i_p}} \left(x'_{i_p} - x_{i_p} \right) \right]^2 dxdy'$$

a simplification of : $\hat{\tau}_y$ only one main term is retained

$$\tau_y = \sum_{p=1}^{s} \int \left(\frac{\partial f(x)}{\partial x_{i_p}} \right)^2 \frac{1 - 3x_{i_p} + 3x_{i_p}^2}{6} dx$$

Theorem 1. If $f(x)$ is linear with respect to $x_{i_1}, \ldots, x_{i_s}$,

then $$S_y^{tot} = \frac{\tau_y}{D}.$$

Theorem 2. A general inequality holds

$$S_y^{tot} \le \frac{24}{\pi^2} \frac{\tau_y}{D}.$$

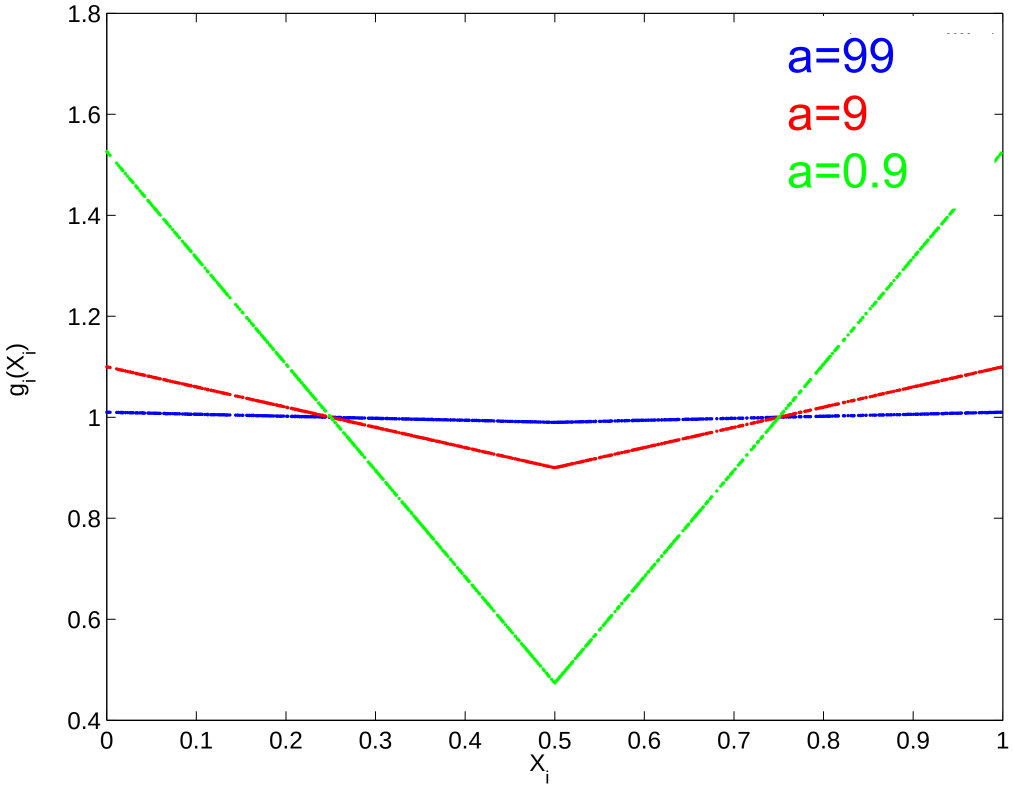
Test: g-function

$$g_i(X_i) = \frac{|4X_i - 2| + a_i}{1 + a_i}$$

$$a_i \geq 0$$

$$X_i \sim U[0,1]$$

$$y = \prod_{i=1}^{k} g_i(X_i)$$



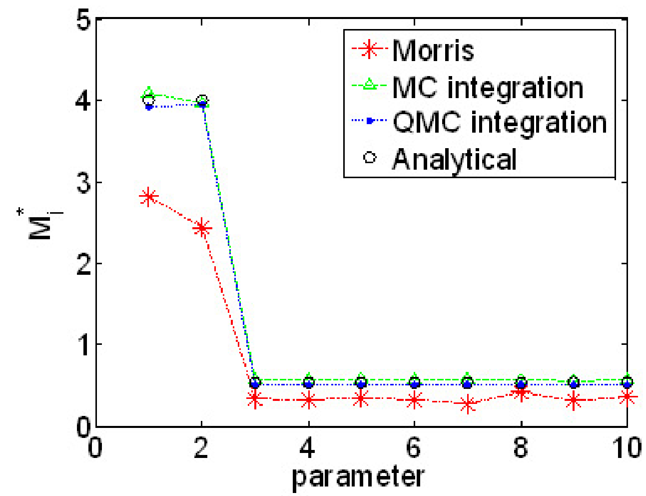
The integration error vs. N.

g-function in 10 dimnensions

$$f(x)=\prod_{i=1}^{n}\frac{|4x_i-2|+a_i}{1+a_i},$$

$$a_1=a_2=0,\ \ a_3=\ldots=a_n=6.52$$

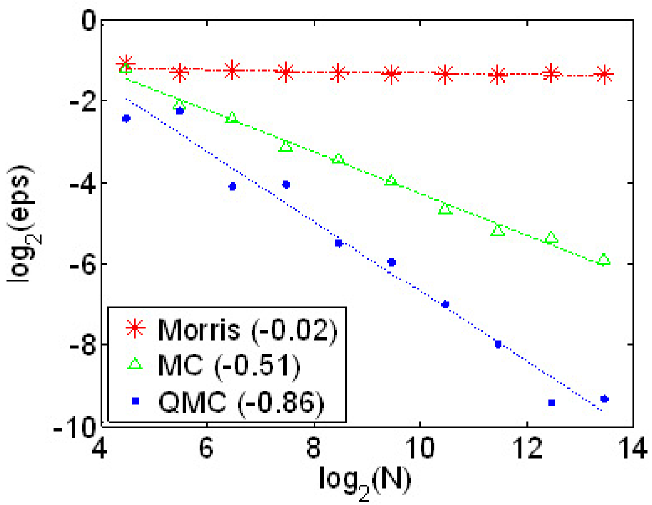
(a)



$$\varepsilon=\left(\frac{1}{K}\sum_{k=1}^{K}(I-I_N^k)^2\right)^{1/2}$$

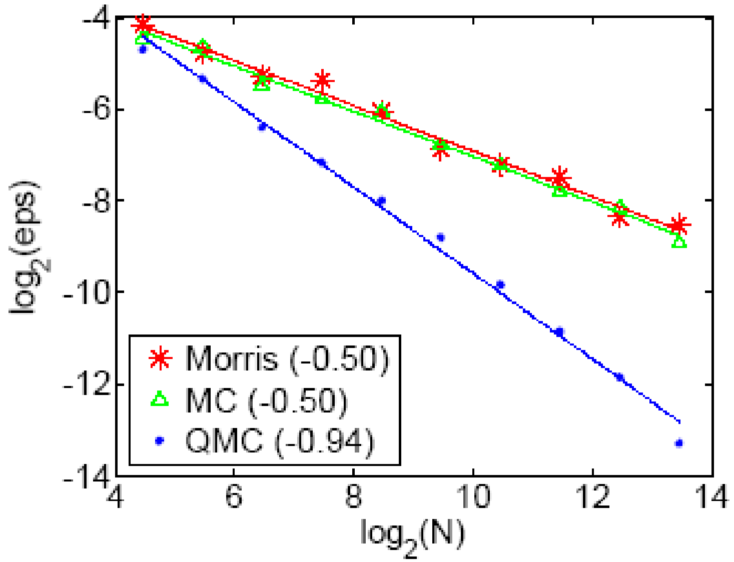
$$\varepsilon \sim N^{-\alpha},\ \ 0<\alpha<1$$

(b)



The integration error vs. N
and CPU time

$$\prod_{i=1}^{n} \frac{n - x_i}{n - 0.5}$$



$\varepsilon \sim N^{-\alpha},\ 0 < \alpha < 1$

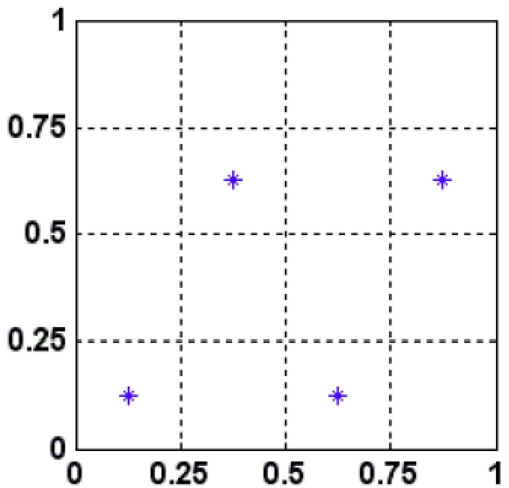
CPU time and number of sampled points required for achieving 1% accuracy

	MORRIS		DGSM(MC)		DGSM(QMC)	
Function	CPU time	*N*	CPU time	*N*	CPU time	*N*
A1	1.33	11264	0.55	5632	0.10	704
B1	0.12	704	0.10	704	0.02	88
C1	1398.10	1441792	139.06	360448	36.59	180224
C2	-	-	439.34	720869	154.33	360448

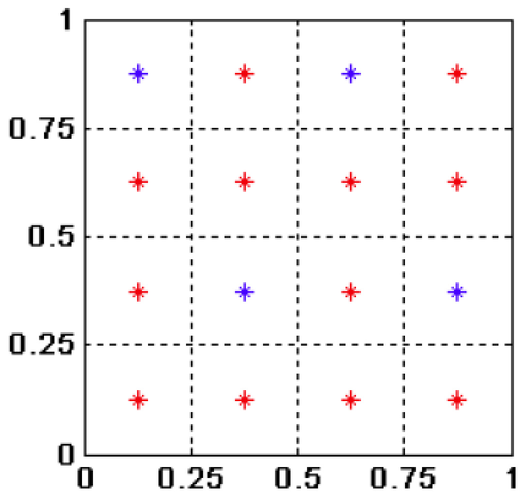
Sobol LDS. Property A and Property A’

Property A. Consider *n*-dimensional hypercube which is cut by plains x_j=1/2 into 2^n subcubes. Sequence of Sobol points satisfies Property A, if after dividing the sequence into binary (!) segments of 2^n points, each one of the points in any one segment belong to a different subcube

Property A’. Consider *n*-dimensional hypercube which is cut by plains $x_j=k/4$, j=1,..,*n*, *k*=1,2,3 into 4^n subcubes. Sequence of Sobol points satisfies Property A’, if after dividing the sequence into into binary (!) segments of 4^n points, each one of the points in any one segment belong to a different subcube

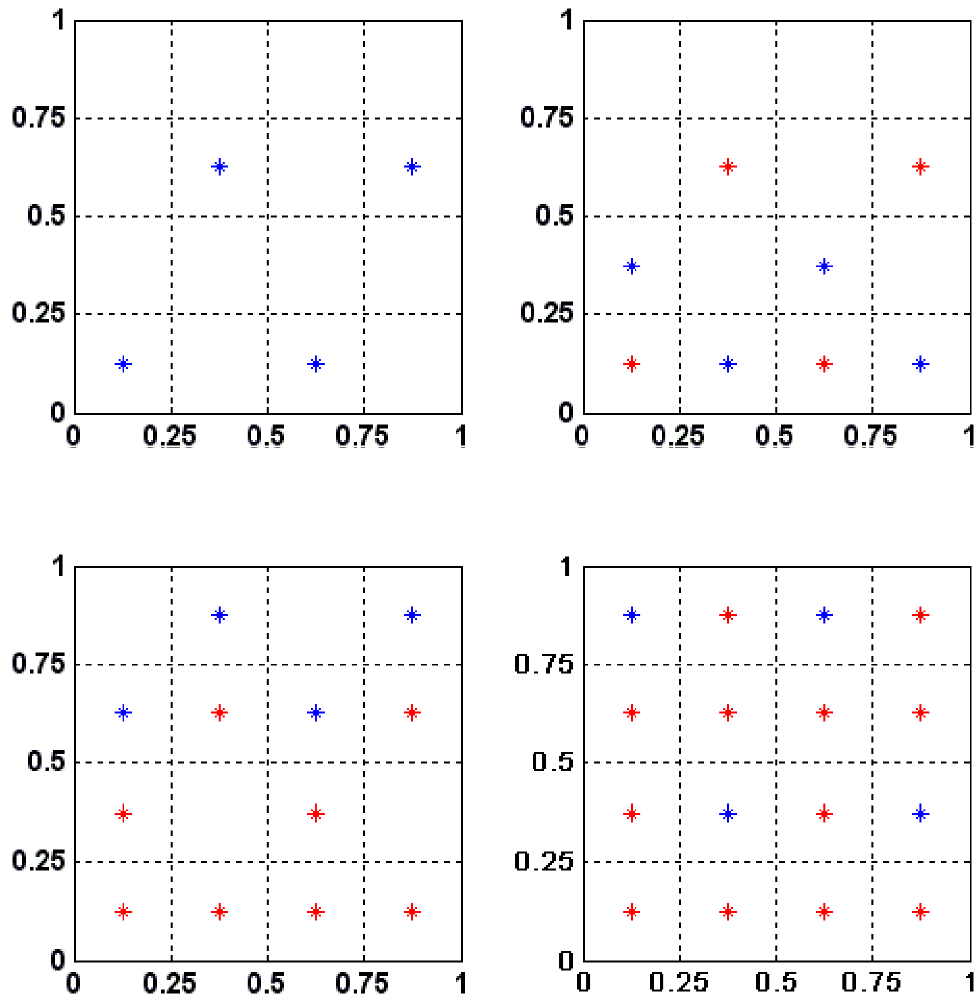


Property A



Property A’

Property A’ but not Property A

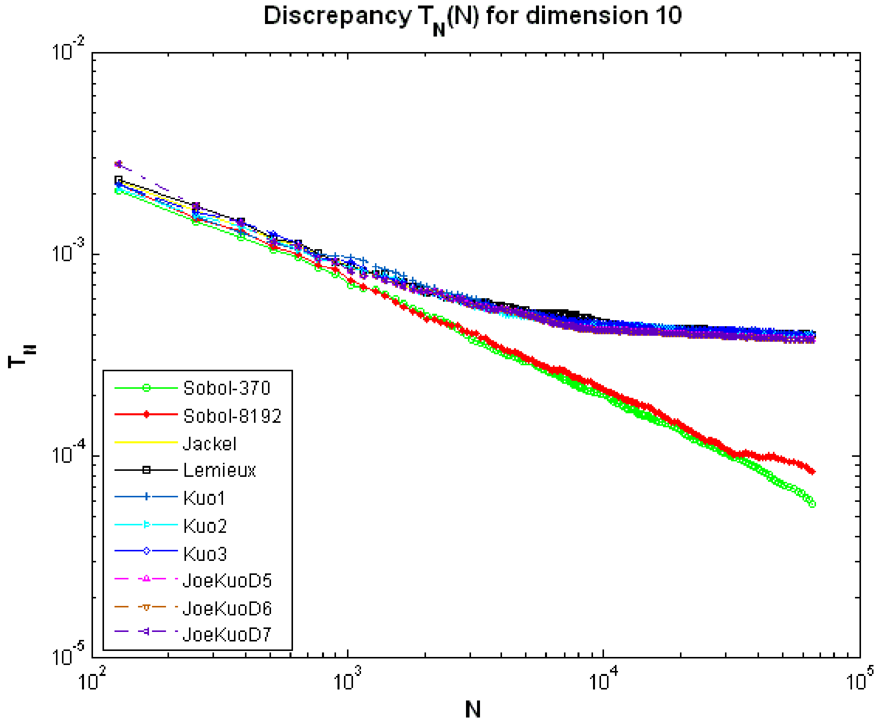


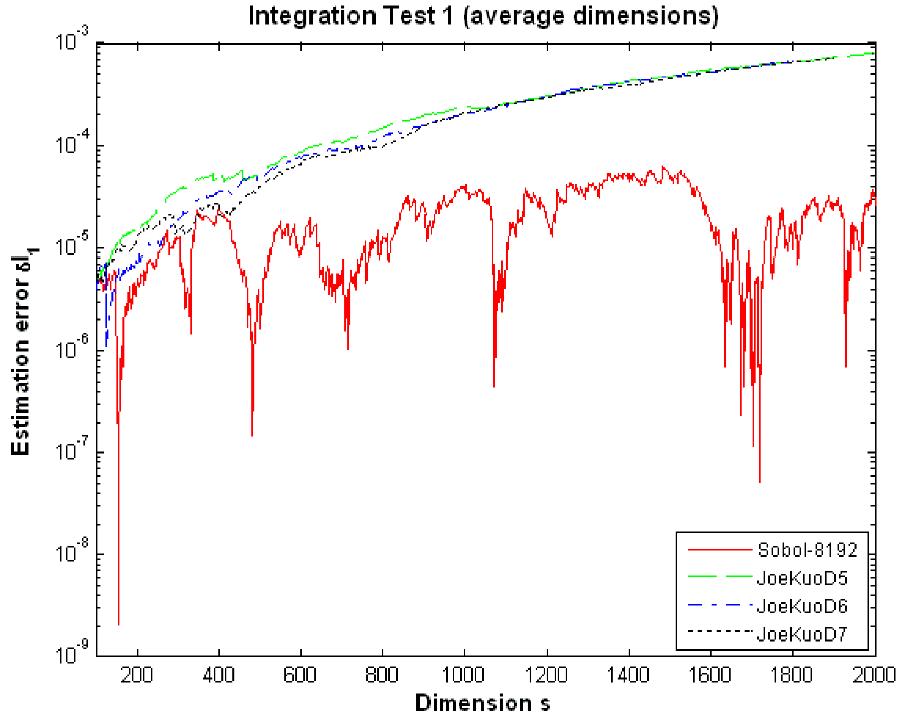
Comparison of Sobol' sequence generators

Generator	Maximum Dimension
Sobol-370	370
Sobol-8192	8192
Jackel	32
Lemieux	360
Kuo1	4925
Kuo2	3946
Kuo3	4586
JoeKuoD5	1999
JoeKuoD6	1799
JoeKuoD7	1899

BRODA's SobolSeq (max dimensionality 16394, www,broda.co.uk). Sobol' sequences produced by SobolSeq satisfy two additional uniformity properties: Property A for all 16394 dimensions and Property A' for adjacent dimensions

$$I = \int_{[0,1]^n} \prod_{i=1}^{s} \left(1 + c_i\left(x_i - 0.5\right)\right) dx_i$$





Summary

1. We introduced new derivative based global sensitivity measures (DGSM) and established a link between DGSM and Sobol' global sensitivity indices

2. For functions linear with respect to a group of variables the performance of DGSM for the group is equivalent to the performance of S^{tot} for the same group.

3. Small values of DGSM always imply small values of S^{tot}.

4. DGSM can be orders of magnitude more efficient than Sobol' SI in terms of CPU time. It can be further improved by using automatic differentiation.

5. The SobolSeq generators satisfying uniformity properties A and A' demonstrate superior performance over other generators