

Reliability sensitivity analysis with dependent inputs

MASCOT-NUM, 8 June 2022

Iason Papaioannou

Engineering Risk Analysis Group, TU München

E-mail: iason.papaioannou@tum.de

Joint work with Max Ehre and Daniel Straub

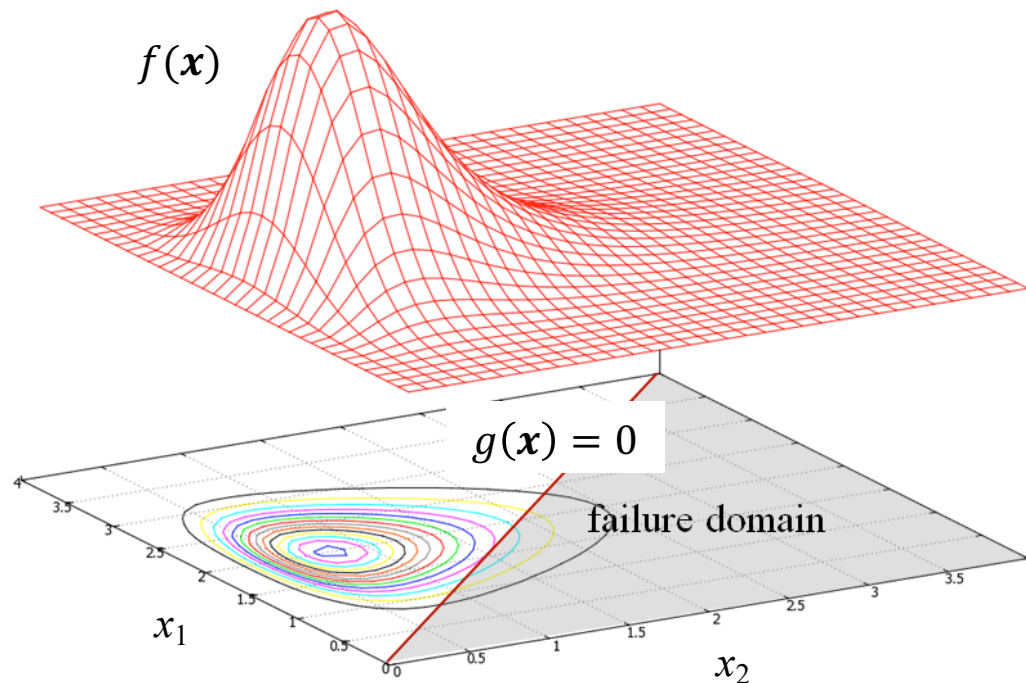
Reliability analysis

Input random variables: $\mathbf{X} = [X_1; X_2; \dots; X_d]$ with joint PDF $f(\mathbf{x})$

Limit-state function $g(\mathbf{x}) = \tilde{g} \circ \mathcal{M}(\mathbf{x})$; Failure event $F = \{\mathbf{x} \in \mathbb{R}^d : g(\mathbf{x}) \leq 0\}$

Probability of failure:

$$p_F := \mathbb{P}(F) = \int_{g(\mathbf{x}) \leq 0} f(\mathbf{x}) d\mathbf{x} = \mathbb{E}[I(g(\mathbf{X}) \leq 0)]$$



Reliability methods

- Approximation methods based on Taylor series: FORM/SORM
- Sampling methods, e.g. splitting, subset simulation, cross-entropy method, line sampling, sequential importance sampling
- Surrogate-based reliability methods, e.g. Gaussian process (kriging), polynomial chaos expansion (PCE), neural networks, low-rank tensors

Reliability sensitivity analysis

- Gradient-based sensitivity analysis: **How does a change in the (deterministic) input parameters influences p_F ?** [Wu 1994; Jensen et al. 2009; Papaioannou et al. 2013, 2018]
- Variance-based/global sensitivity analysis: **How does the variability of the input random variables influences p_F ?** [Li et al. 2012; Wei et al. 2012; Ehre et al. 2020]
- Decision-theoretic sensitivity analysis: **How does the reduction of uncertainty of the input random variables influence the optimality of an engineering decision?** [Straub et al. 2022]

Reliability sensitivity analysis

- Gradient-based sensitivity analysis: How does a change in the (deterministic) input parameters influences p_F ? [Wu 1994; Jensen et al. 2009; Papaioannou et al. 2013, 2018]
- Variance-based/global sensitivity analysis: How does the variability of the input random variables influences p_F ? [Li et al. 2012; Wei et al. 2012; Ehre et al. 2020]
- Decision-theoretic sensitivity analysis: How does the reduction of uncertainty of the input random variables influence the optimality of an engineering decision? [Straub et al. 2022]

Variance-based sensitivity analysis

Consider an n -dimensional **independent random vector** \mathbf{X} and a function $Q = h(\mathbf{X})$.

ANOVA representation:

$$h(\mathbf{X}) = h_0 + \sum_{i=1}^d h_i(X_i) + \sum_{i<j}^d h_{ij}(X_i, X_j) + \cdots + h_{1\dots n}(X_1, \dots, X_n)$$

The ANOVA representation exists and is unique provided that for any subset $\mathbf{X}_v \subseteq \mathbf{X}$ and any $i \in v$

$$\mathbb{E}[h_v(\mathbf{X}_v) | \mathbf{X}_{v \setminus i}] = \int_{-\infty}^{\infty} h_v(\mathbf{x}_v) f_i(x_i) dx_i = 0$$

Variance decomposition:

$$\mathbb{V}(Q) = \sum_{i=1}^d V_i + \sum_{i<j}^d V_{ij} + \cdots + V_{1\dots d}$$

Variance-based sensitivity indices

First-order (Sobol') indices:

$$S_i = \frac{V_i}{V(Q)} = \frac{V(\mathbb{E}[Q|X_i])}{V(Q)}$$

Total effect indices

$$S_i^T = \frac{V_i + \sum_{j \neq i}^d V_{ij} + \dots + V_{1\dots d}}{V(Q)} = 1 - \frac{V(\mathbb{E}[Q|\mathbf{X}_{\sim i}])}{V(Q)}$$

It holds: $0 \leq S_i \leq S_i^T \leq 1$

Estimation of variance-based sensitivity indices

- Double loop Monte Carlo procedures are computationally prohibitive
- Peak-freeze estimators with N samples: $N_T = N(d + 2)$ [Sobol' 1993; Jansen 1999; Saltelli et al. 2010]
- Spectral methods: FAST, RBD, polynomial chaos [Efron & Stein 1981; Plischke 2010; Sudret 2008; Ehre et al. 2020]

Variance-based reliability sensitivities

Failure event $F = \{\mathbf{x} \in \mathbb{R}^d: g(\mathbf{x}) \leq 0\}$

Quantity of interest: $Z = I(g(\mathbf{X}) \leq 0)$

It is $\mathbb{E}[Z] = p_F, \mathbb{V}[Z] = p_F(1 - p_F)$

First-order indices:

$$S_{F,i} = \frac{\mathbb{V}(\mathbb{E}[Z|X_i])}{\mathbb{V}(Z)} = \frac{\mathbb{V}(\mathbb{P}[F|X_i])}{p_F(1 - p_F)}$$

Total-effect indices:

$$S_{F,i}^T = 1 - \frac{\mathbb{V}(\mathbb{E}[Z|\mathbf{X}_{\sim i}])}{\mathbb{V}(Z)} = 1 - \frac{\mathbb{V}(\mathbb{P}[F|\mathbf{X}_{\sim i}])}{p_F(1 - p_F)}$$

Interpretation of reliability sensitivities

- First-order indices can be used for factor prioritization, to determine which random variable if learned will increase the accuracy of p_F the most
- Total-effect indices can be used for factor fixing, to determine the random variables with $S_{F,i}^T \approx 0$, which if fixed will not impact the prediction of p_F

Estimation of reliability sensitivities

- Peek-freeze estimators combined with importance sampling [Wei et al. 2012]
- Estimation with failure samples [Perrin & Defaux 2019, Li et al. 2019]
- Estimation with FORM [Papaioannou & Straub 2021]

Estimation with failure samples

Probability of failure with importance sampling:

$$p_F = \mathbb{E}_\psi[I(g(\mathbf{X}) \leq 0)w(\mathbf{X})]$$

Importance weight function: $w(\mathbf{x}) = \frac{f(\mathbf{x})}{\psi(\mathbf{x})}$; Importance sampling density: $\psi(\mathbf{x})$

Estimate of probability

$$\hat{p}_F = \hat{\mathbb{E}}_\psi[I(g(\mathbf{X}) \leq 0)w(\mathbf{X})] = \frac{1}{N} \sum_{k=1}^N I(g(\mathbf{x}_k) \leq 0)w(\mathbf{x}_k), \quad \mathbf{x}_k \sim \psi(\cdot)$$

Optimal (zero variance) IS density

$$\psi^*(\mathbf{x}) = f(\mathbf{x}|F) = \frac{1}{p_F} I(g(\mathbf{x}) \leq 0)f(\mathbf{x})$$

The density ψ can be estimated with adaptive sampling methods [Bucher 1988; Papaioannou et al. 2016; 2019]

Estimation with failure samples (II)

First-order indices:

$$S_{F,i} = \frac{\mathbb{V}(\mathbb{E}[Z|X_i])}{\mathbb{V}(Z)} = \frac{\mathbb{V}(\mathbb{P}[F|X_i])}{p_F(1 - p_F)}$$

Bayes' rule:

$$\mathbb{P}[F|X_i = x_i] = \frac{f_i(x_i|F)\mathbb{P}(F)}{f_i(x_i)}$$

$f_{X_i}(x_i|F)$ can be estimated with kernel density estimation using (weighted) failure samples [Perrin & Defaux 2019, Li et al. 2019]

Estimation with failure samples (III)

Total-effect indices:

$$S_{F,i}^T = 1 - \frac{\mathbb{V}(\mathbb{E}[Z|\mathbf{X}_{\sim i}])}{\mathbb{V}(Z)} = 1 - \frac{\mathbb{V}(\mathbb{P}[F|\mathbf{X}_{\sim i}])}{p_F(1 - p_F)}$$

Bayes' rule:

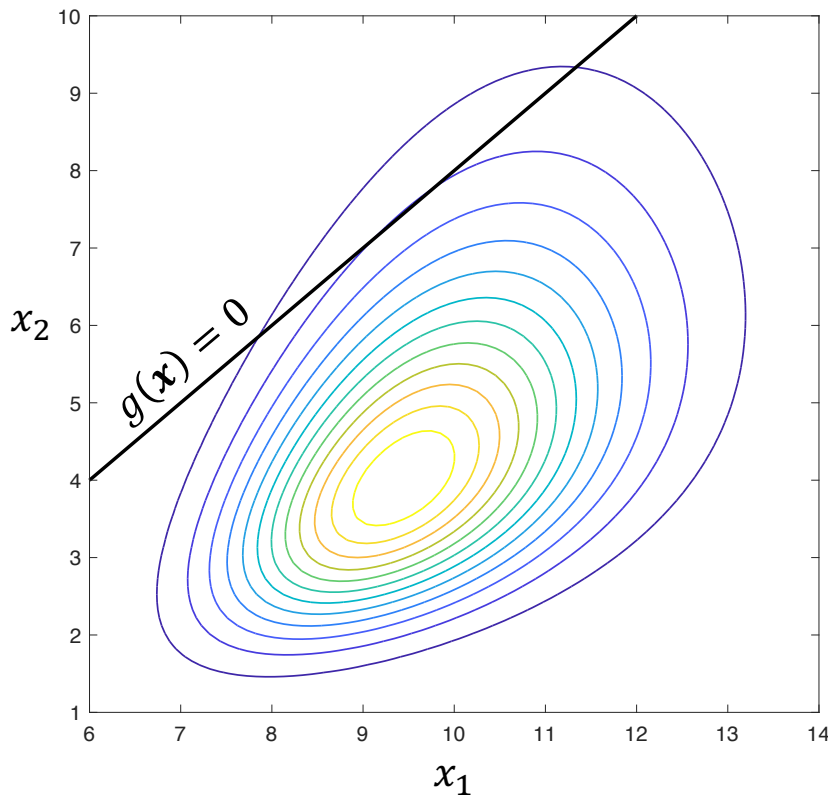
$$\mathbb{P}[F|\mathbf{X}_{\sim i} = \mathbf{x}_{\sim i}] = \frac{f_{\sim i}(\mathbf{x}_{\sim i}|F)\mathbb{P}(F)}{f_{\mathbf{x}_{\sim i}}(\mathbf{x}_{\sim i})}$$

$f_{\sim i}(\mathbf{x}_{\sim i}|F)$ can be estimated with multivariate kernel density estimation using (weighted) failure samples

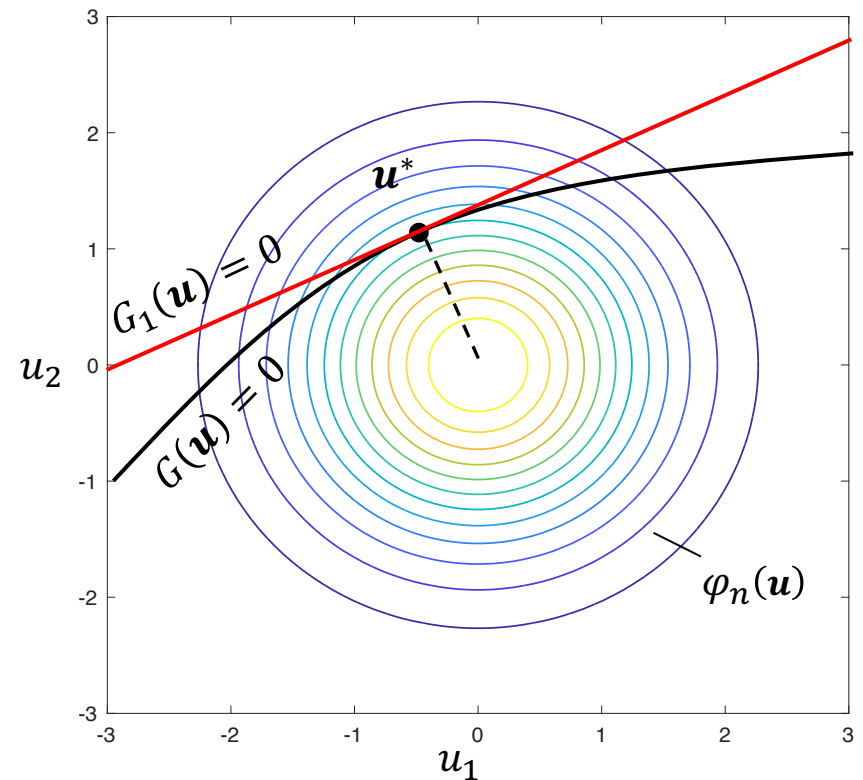
Estimation with FORM

- Transform \mathbf{X} to an equivalent space $\mathbf{U} \sim N(\mathbf{0}, \mathbf{I})$, $\mathbf{U} = T(\mathbf{X})$ with $T: U_i = \Phi^{-1} \circ F_i(X_i)$
- Transformed limit-state function: $G(\mathbf{u}) = g \circ T^{-1}(\mathbf{u})$
- Choose \mathbf{u}^* as the most likely (probable) point on the hypersurface $G(\mathbf{u}) = 0$

Original space

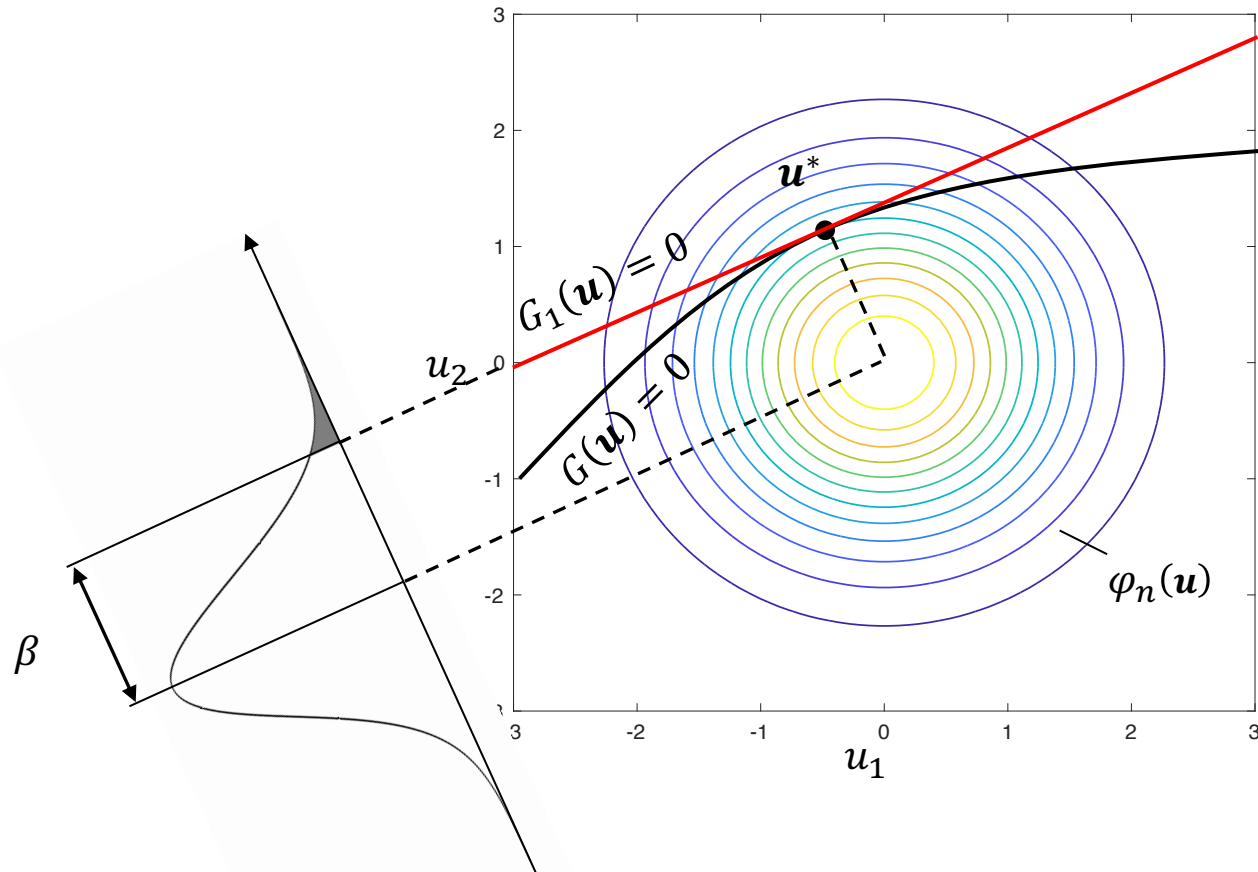


Transformed space



FORM approximation of p_F

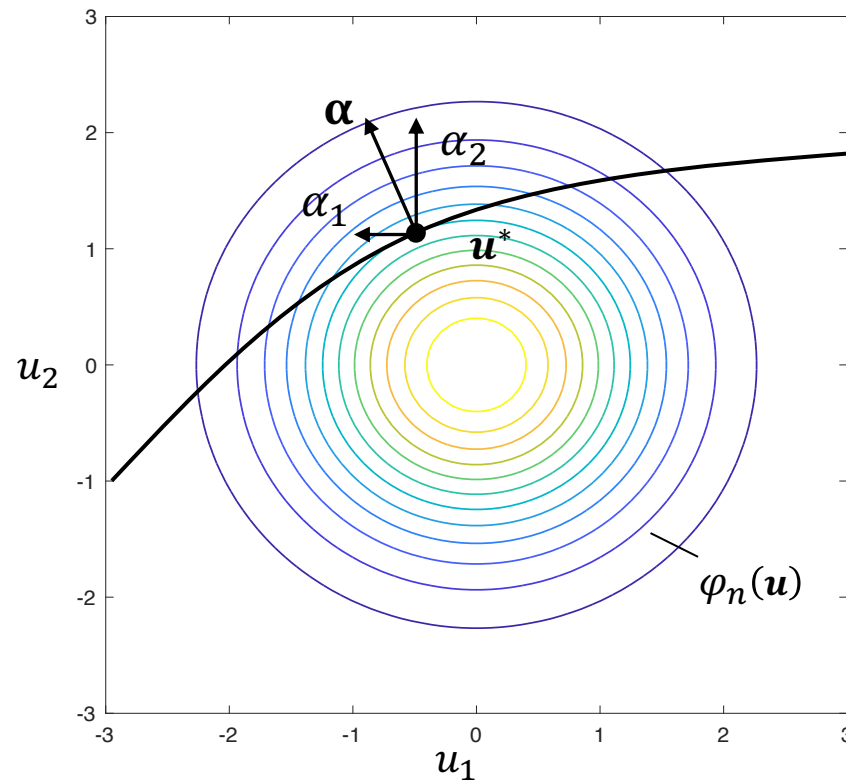
Probability approximation: $p_F \approx p_{F_1} = \Phi(-\beta)$



The FORM α -factors

Normalized negative gradient of the limit-state function at the design point \mathbf{u}^*

$$\boldsymbol{\alpha} = -\frac{\nabla G(\mathbf{u}^*)}{\|\nabla G(\mathbf{u}^*)\|} = \frac{\mathbf{u}^*}{\beta}$$



FORM α -factors as variance-based sensitivities

Approximation of limit-state function in \mathbf{U} -space

$$G(\mathbf{U}) \approx G_1(\mathbf{U}) = \nabla G(\mathbf{u}^*)(\mathbf{U} - \mathbf{u}^*) = \|\nabla G(\mathbf{u}^*)\|(\beta - \boldsymbol{\alpha}\mathbf{U})$$

where $\beta = \boldsymbol{\alpha}\mathbf{u}^*$ is the FORM reliability index

Variance decomposition of FORM approximation of the linearized limit-state function $G_1(\mathbf{U})$:

$$\mathbb{V}(G_1) = \|\nabla G(\mathbf{u}^*)\|^2 \sum_{i=1}^n \alpha_i^2 = \|\nabla G(\mathbf{u}^*)\|^2$$

First-order indices of G_1

$$S_{G_1,i} = \frac{\mathbb{V}(E[G_1|U_i])}{\mathbb{V}(G_1)} = \frac{\|\nabla G(\mathbf{u}^*)\|^2 \alpha_i^2}{\|\nabla G(\mathbf{u}^*)\|^2} = \alpha_i^2$$

Variance-based SA with FORM [Papaioannou & Straub 2021]

Quantity of interest: $Z_1 = I(G_1(\mathbf{U}) \leq 0)$

It is $\mathbb{E}[Z_1] = p_{F_1}, \mathbb{V}[Z_1] = p_{F_1}(1 - p_{F_1})$

First-order index:

$$S_{F_1,i} = \frac{\mathbb{V}(\mathbb{E}[Z_1|U_i])}{\mathbb{V}(Z_1)} = \frac{1}{p_{F_1}(1 - p_{F_1})} \int_0^{\alpha_i^2} \varphi_2(-\beta, -\beta, r) dr$$

Total effect index:

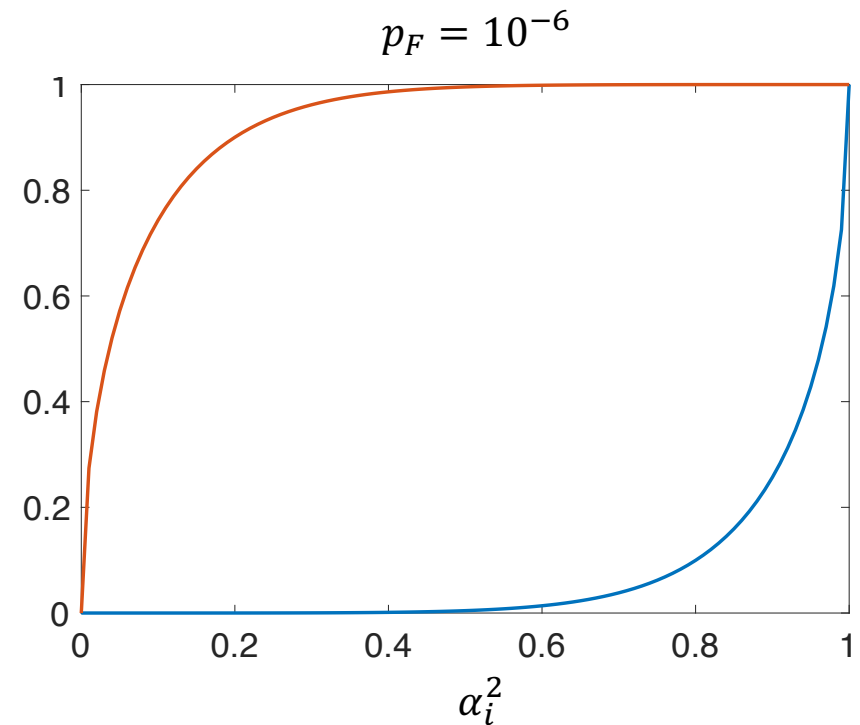
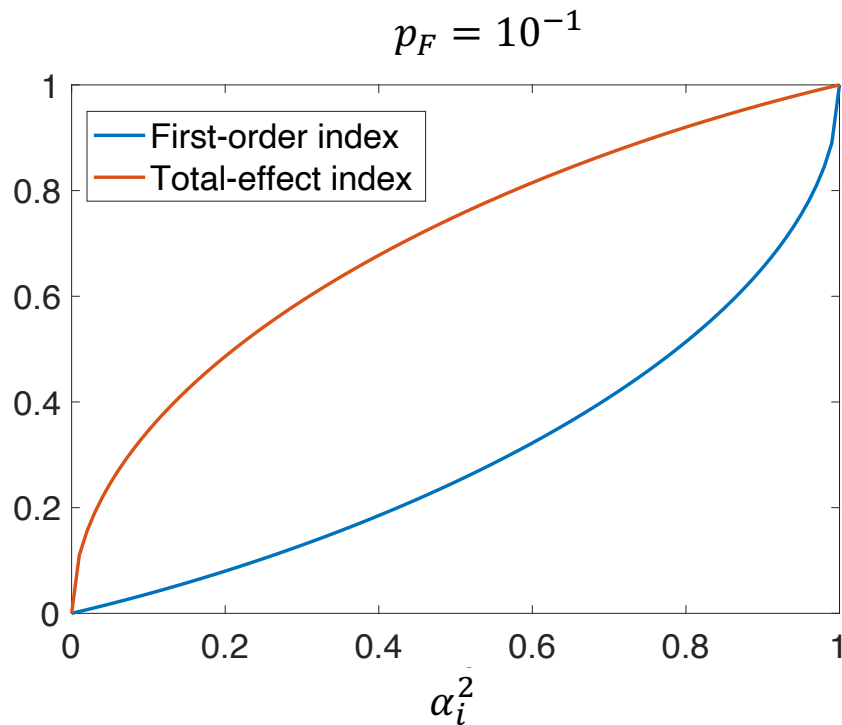
$$S_{F_1,i}^T = 1 - \frac{\mathbb{V}(\mathbb{E}[Z_1|\mathbf{U}_{\sim i}])}{\mathbb{V}(Z_1)} = \frac{1}{p_{F_1}(1 - p_{F_1})} \int_{1-\alpha_i^2}^1 \varphi_2(-\beta, -\beta, r) dr$$

with

$$\varphi_2(-\beta, -\beta, r) = \frac{1}{2\pi\sqrt{1-r^2}} \exp\left(-\frac{\beta^2}{1+r}\right)$$

Illustration of FORM variance-based indices

FORM-based first-order and total effect indices vs. α_i^2 for varying p_F values



Variance based SA with dependent inputs

If X is dependent...

- ... the ANOVA representation is non-unique
- ... the summands are not orthogonal
- ... the variance decomposition includes contributions from covariance terms

Approaches for dealing with dependent inputs

- Compute terms of a generalized ANOVA decomposition [Li et al. 2010; Chastaig et al. 2012]
- Apply a transformation to decorrelate the inputs [Xu & Gertner 2008; Mara & Tarantola 2012; 2015]
- Compute Shapley effects [Owen 2014; Owen & Prieur 2017; Il Idrissi et al. 2021]

Variance based SA with dependent inputs

If X is dependent...

- ... the ANOVA representation is non-unique
- ... the summands are not orthogonal
- ... the variance decomposition includes contributions from covariance terms

Approaches for dealing with dependent inputs

- Compute terms of a generalized ANOVA decomposition [Li et al. 2010; Chastaig et al. 2012]
- **Apply a transformation to decorrelate the inputs** [Xu & Gertner 2008; Mara & Tarantola 2012; 2015]
- Compute Shapley effects [Owen 2014; Owen & Prieur 2017; Il Idrissi et al. 2021]

Iso-probabilistic transform

Consider a random vector $\mathbf{X}: \mathcal{X} \rightarrow \mathbb{R}^d$ with joint PDF $f(\mathbf{x})$. Find $T: \mathbb{R}^d \rightarrow \mathbb{R}^d$ s.t.

$$\mathbf{U} = T(\mathbf{X})$$

follows the independent standard normal distribution, i.e., $\mathbf{U} \sim N(\mathbf{0}, \mathbf{I})$.

The transformation T is called **hierarchical** if it takes the following form:

$$T: \begin{bmatrix} U_1 \\ U_2 \\ \vdots \\ U_d \end{bmatrix} = \begin{bmatrix} \eta_1(X_1) \\ \eta_2(X_1, X_2) \\ \vdots \\ \eta_d(X_1, X_2, \dots, X_d) \end{bmatrix}$$

Example: Rosenblatt transform

$$T_R: \begin{bmatrix} U_1 \\ U_2 \\ \vdots \\ U_d \end{bmatrix} = \begin{bmatrix} \Phi^{-1} \circ F_1(X_1) \\ \Phi^{-1} \circ F_{2|1}(X_2|X_1) \\ \vdots \\ \Phi^{-1} \circ F_{d|1,\dots,d-1}(X_d|X_1, \dots, X_{d-1}) \end{bmatrix}$$

F_1 : Marginal CDF of X_1

$F_{i|1,\dots,i-1}$: Conditional CDF of X_i given X_1, \dots, X_{i-1}

Example: Nataf transform

$$T_N: \begin{bmatrix} U_1 \\ U_2 \\ \vdots \\ U_d \end{bmatrix} = \mathbf{A}^{-1} \begin{bmatrix} \Phi^{-1} \circ F_1(X_1) \\ \Phi^{-1} \circ F_2(X_2) \\ \vdots \\ \Phi^{-1} \circ F_d(X_d) \end{bmatrix} = \mathbf{A}^{-1} \mathbf{Y}$$

F_i : Marginal CDF of X_i

\mathbf{A} : Lower triangular matrix of the Cholesky decomposition $\mathbf{R} = \mathbf{A}\mathbf{A}^T$

\mathbf{R} : Correlation matrix of \mathbf{Y}

Remarks:

- The Nataf transform implies that the vector \mathbf{X} follows a Gaussian copula joint distribution [Lebrun & Duftoy 2009]
- The Nataf transform is equivalent to the Rosenblatt transform of the Gaussian copula model

Variance-based SA in U -space

Transform the quantity of interest: $Q = h(\mathbf{X}) = h(T^{-1}(\mathbf{U}))$.

First-order (Sobol') indices:

$$S_{U_i} = \frac{\mathbb{V}(\mathbb{E}[Q|U_i])}{\mathbb{V}(Q)}$$

Total effect indices:

$$S_{U_i}^T = 1 - \frac{\mathbb{V}(\mathbb{E}[Q|\mathbf{U}_{\sim i}])}{\mathbb{V}(Q)}$$

Owing to the hierarchical nature of the transform:

- $S_{U_1}, S_{U_d}^T$ reflect the **full** first-order and total effect contributions of X_1 (including those due to its dependence with other variables)
- $S_{U_d}, S_{U_1}^T$ reflect the **independent** first-order and total effect contributions of X_d

Variance-based SA in U -space (II)

Consider d cyclic shifts of the ordered set $\{X_1, X_2, \dots, X_d\}$.

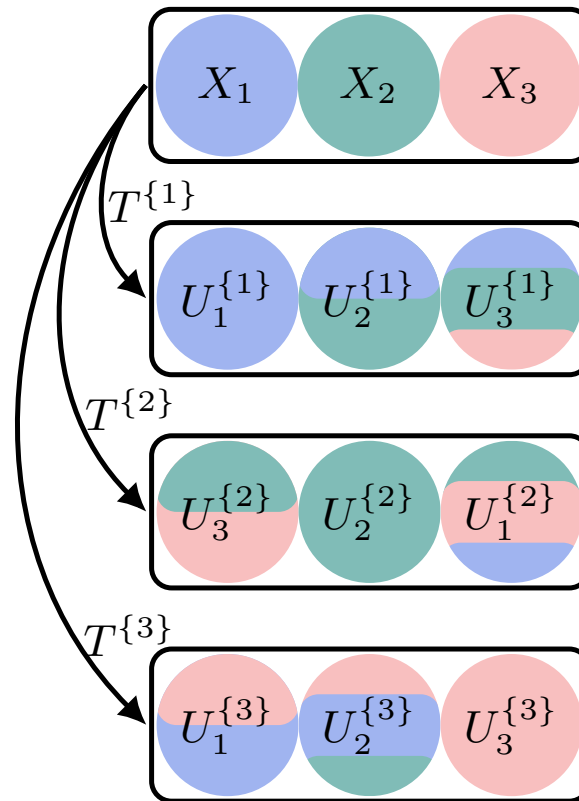
Define the set of hierarchical transforms $\{T^{\{i\}}\}_{i=1}^d$ with

$$T^{\{i\}}: \begin{bmatrix} U_i^{\{i\}} \\ U_{i+1}^{\{i\}} \\ \vdots \\ U_{i-1}^{\{i\}} \end{bmatrix} = \begin{bmatrix} \eta_1^{\{i\}}(X_i) \\ \eta_2^{\{i\}}(X_i, X_{i+1}) \\ \vdots \\ \eta_d^{\{i\}}(X_i, X_{i+1}, \dots, X_{i-1}) \end{bmatrix}$$

corresponding to the variable ordering $\{X_i, X_{i+1}, \dots, X_{i-1}\}$ and $X_0 := X_d$.

Variance-based SA in U -space (III)

The d cyclic shifts and corresponding hierarchical transforms allow isolating the independent contributions of each input to $V(Q)$ [Mara & Tarantola 2012]



Variance-based sensitivity indices

First-order (Sobol') indices:

$$S_i = S_{U_i^{\{i\}}} = \frac{\mathbb{V}(\mathbb{E}[Q|U_i^{\{i\}}])}{\mathbb{V}(Q)}$$

$$S_{i,\text{ind}} = S_{U_i^{\{i+1\}}} = \frac{\mathbb{V}(\mathbb{E}[Q|U_i^{\{i+1\}}])}{\mathbb{V}(Q)}$$

Total effect indices:

$$S_i^T = S_{U_i^{\{i+1\}}}^T = 1 - \frac{\mathbb{V}(\mathbb{E}[Q|\mathbf{U}_{\sim i}^{\{i+1\}}])}{\mathbb{V}(Q)}$$

$$S_{i,\text{ind}}^T = S_{U_i^{\{i\}}}^T = 1 - \frac{\mathbb{V}(\mathbb{E}[Q|\mathbf{U}_{\sim i}^{\{i\}}])}{\mathbb{V}(Q)}$$

Peak-freeze estimators with N samples: $N_T = 0.5Nd(d + 2)$ [Mara & Tarantola 2015]

Variance-based RS with dependent inputs

Failure event $F = \{\mathbf{x} \in \mathbb{R}^d: g(\mathbf{x}) \leq 0\}$

Quantity of interest: $Z = I(g(\mathbf{X}) \leq 0)$

First-order indices:

$$S_{F,i} = \frac{\mathbb{V}\left(\mathbb{E}\left[Z|U_i^{\{i\}}\right]\right)}{\mathbb{V}(Z)} = \frac{\mathbb{V}\left(\mathbb{P}\left[F|U_i^{\{i\}}\right]\right)}{p_F(1-p_F)}$$

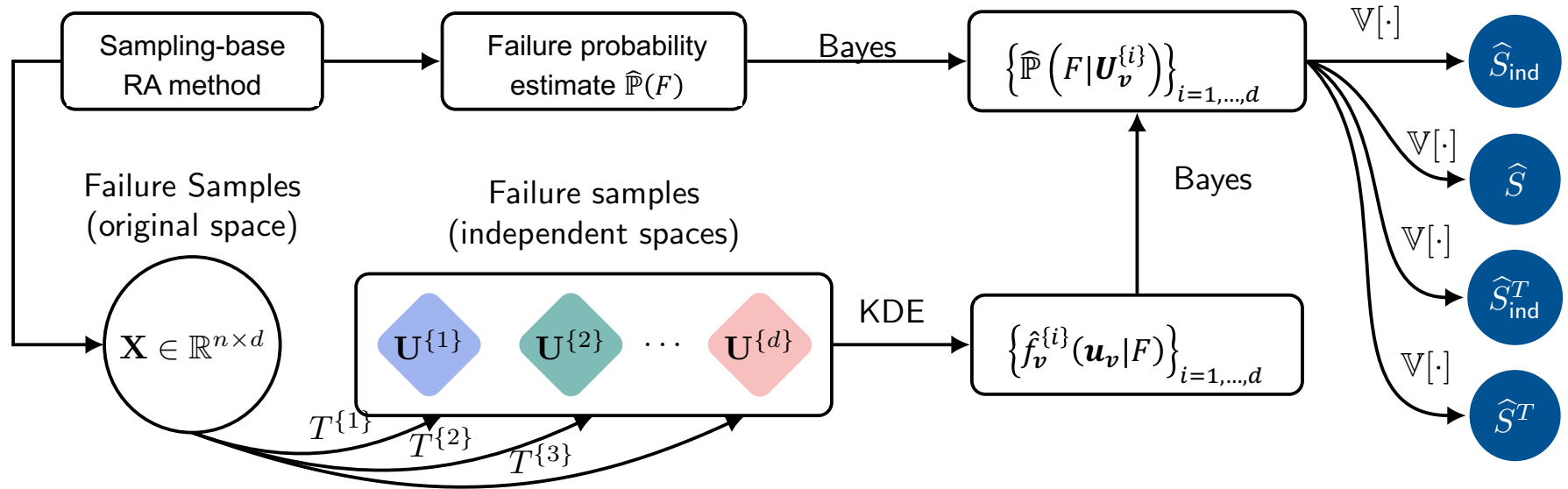
$$S_{F,i,\text{ind}} = \frac{\mathbb{V}\left(\mathbb{E}\left[Z|U_i^{\{i+1\}}\right]\right)}{\mathbb{V}(Z)} = \frac{\mathbb{V}\left(\mathbb{P}\left[F|U_i^{\{i+1\}}\right]\right)}{p_F(1-p_F)}$$

Total-effect indices:

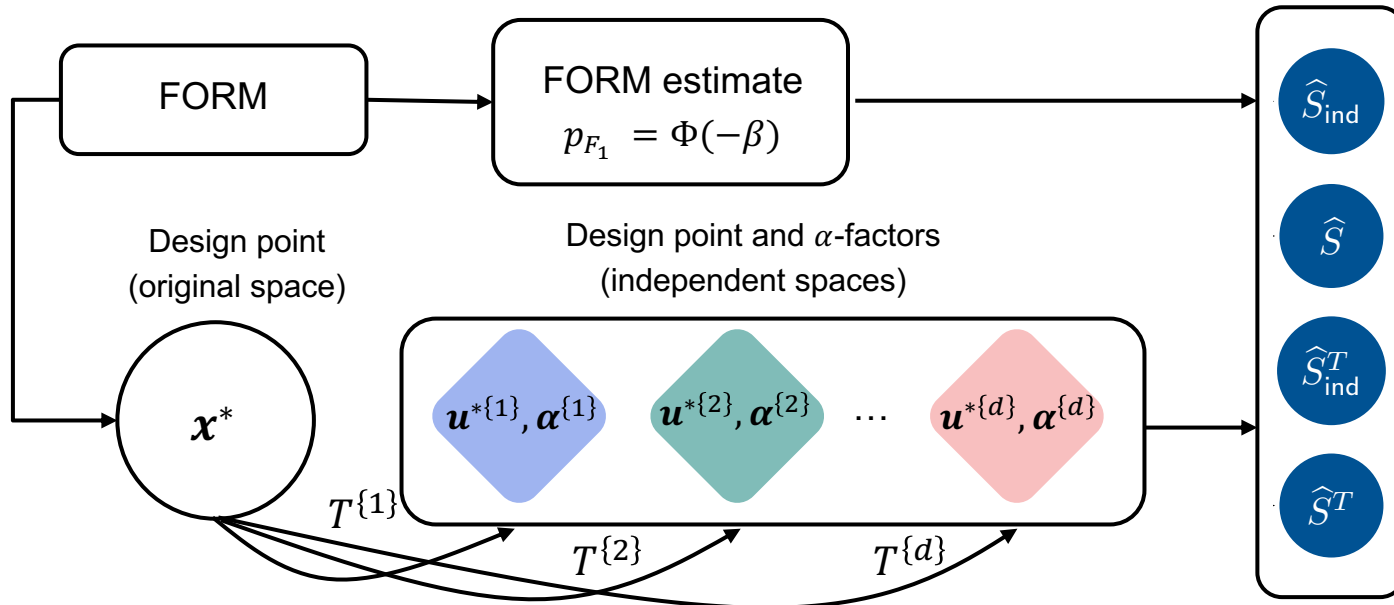
$$S_{F,i}^T = 1 - \frac{\mathbb{V}\left(\mathbb{E}\left[Z|\mathbf{U}_{\sim i}^{\{i+1\}}\right]\right)}{\mathbb{V}(Z)} = 1 - \frac{\mathbb{V}\left(\mathbb{P}\left[F|\mathbf{U}_{\sim i}^{\{i+1\}}\right]\right)}{p_F(1-p_F)}$$

$$S_{F,i,\text{ind}}^T = 1 - \frac{\mathbb{V}\left(\mathbb{E}\left[Z|\mathbf{U}_{\sim i}^{\{i\}}\right]\right)}{\mathbb{V}(Z)} = 1 - \frac{\mathbb{V}\left(\mathbb{P}\left[F|\mathbf{U}_{\sim i}^{\{i\}}\right]\right)}{p_F(1-p_F)}$$

Estimation with failure samples [Ehre et al. In prep]



Estimation with FORM [Papaioannou et al. In prep]



Modified α -factors

Define:

$$\begin{aligned}\tilde{\alpha}_i &= \alpha_i^{\{i\}} \\ \tilde{\alpha}_{i,\text{ind}} &= \alpha_i^{\{i+1\}}\end{aligned}$$

The modified squared α -factors can be viewed as variance-based sensitivities in the corresponding transformed spaces:

$$\begin{aligned}\tilde{\alpha}_i^2 &= S_{G_1, U_i^{\{i\}}} = \frac{\mathbb{V}\left(\mathbb{E}\left[G_1 \mid U_i^{\{i\}}\right]\right)}{\mathbb{V}(Q)} \\ \tilde{\alpha}_{i,\text{ind}}^2 &= S_{G_1, U_i^{\{i+1\}}} = \frac{\mathbb{V}\left(\mathbb{E}\left[G_1 \mid U_i^{\{i+1\}}\right]\right)}{\mathbb{V}(Q)}\end{aligned}$$

Interpretation of modified α -factors

- $\tilde{\alpha}_i^2$ is the the **full** first-order index of the linearized LSF transformed back to \mathbf{X} -space $G_1 \circ T(\mathbf{X})$
- $\tilde{\alpha}_{i,\text{ind}}^2$ is the the **independent** first-order index of $G_1 \circ T(\mathbf{X})$

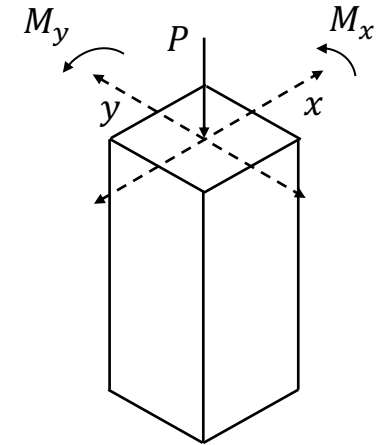
Remark: The γ -factors proposed in [Der Kiureghian 2005] attempt to approximate the independent factors $\tilde{\alpha}_{i,\text{ind}}$ by linearizing $T(\mathbf{X})$.

Example: Short column

Column subjected to biaxial bending and axial force.

Limit-state function:

$$g(\mathbf{x}) = 1 - \frac{M_x}{s_x Y} - \frac{M_y}{s_y Y} - \left(\frac{P}{AY}\right)^2$$



Deterministic parameters:

$$s_x = 0.03\text{m}^3, s_y = 0.015\text{m}^3, A = 0.19\text{m}^2$$

Parameter	Distribution	Mean	CV	\mathbf{R}_x			
				M_x	M_y	P	Y
M_x [KNm]	Normal	250	0.3	1.0	0.5	0.3	0.0
M_y [KNm]	Normal	125	0.3	0.5	1.0	0.3	0.0
P [kN]	Gumbel	2500	0.2	0.3	0.3	1.0	0.0
Y [MPa]	Weibull	40	0.1	0.0	0.0	0.0	1.0

Example: Short column

Column subjected to biaxial bending and axial force.

Reference probability of failure:

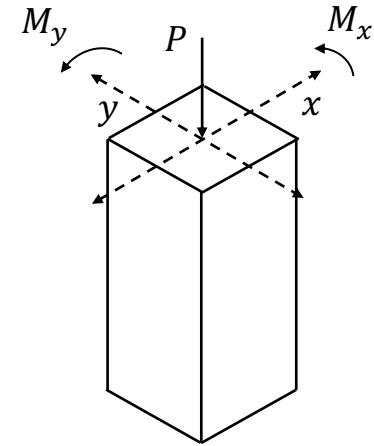
$$p_F = 9.29 \times 10^{-3}$$

FORM estimate:

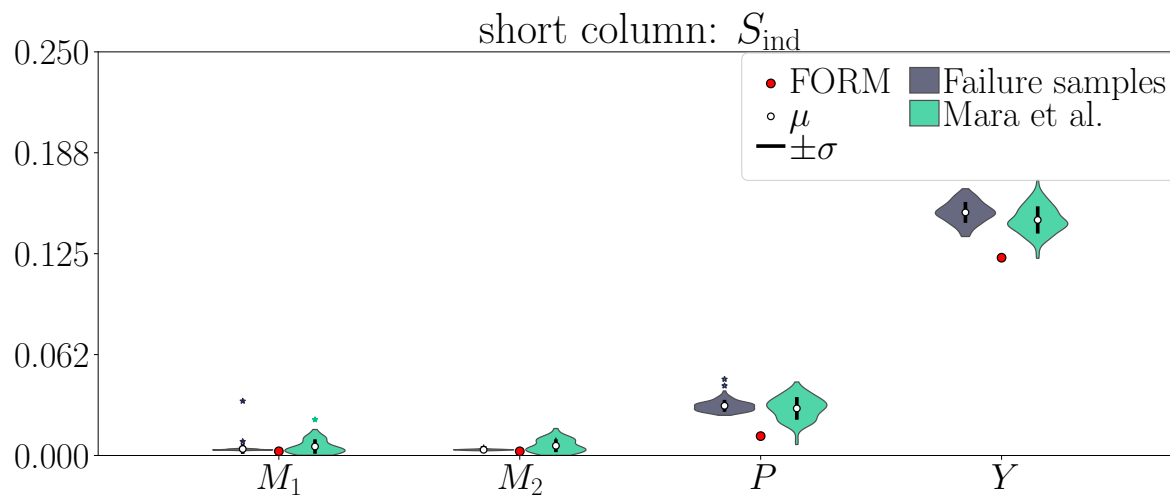
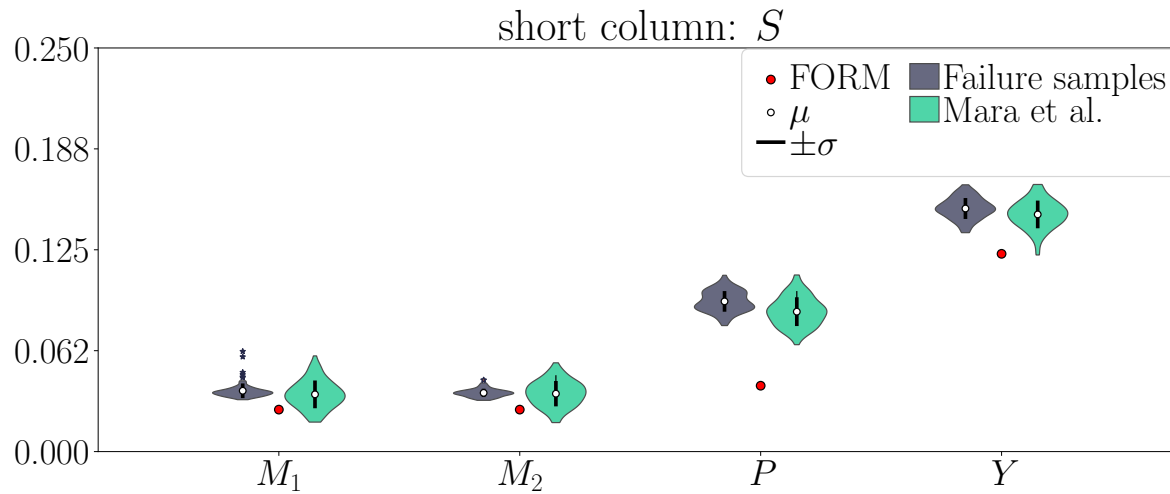
$$p_F = 6.82 \times 10^{-3}$$

Estimates with failure samples using improved CE method [Papaioannou et al. 2019] with $N = 5 \cdot 10^4$ samples per level

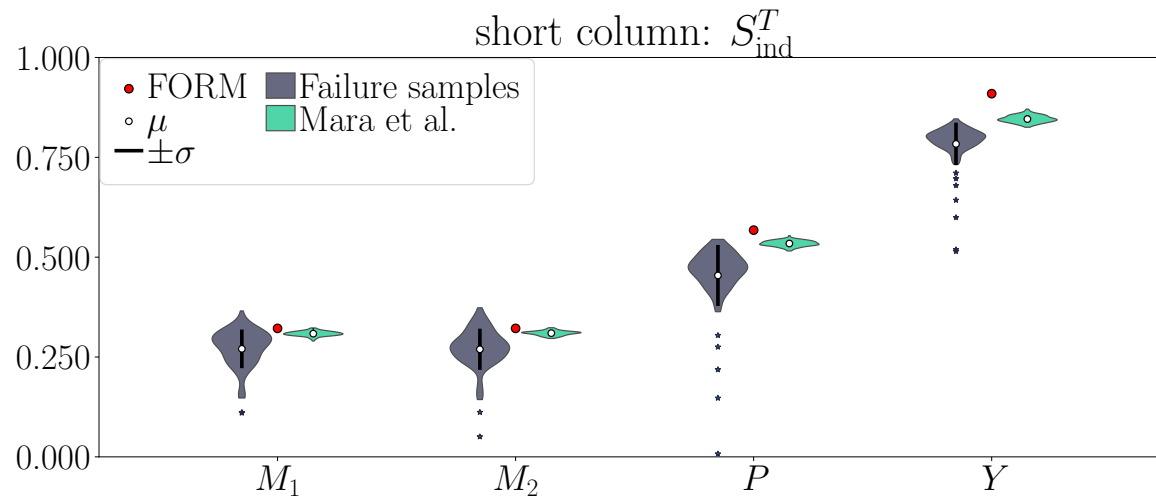
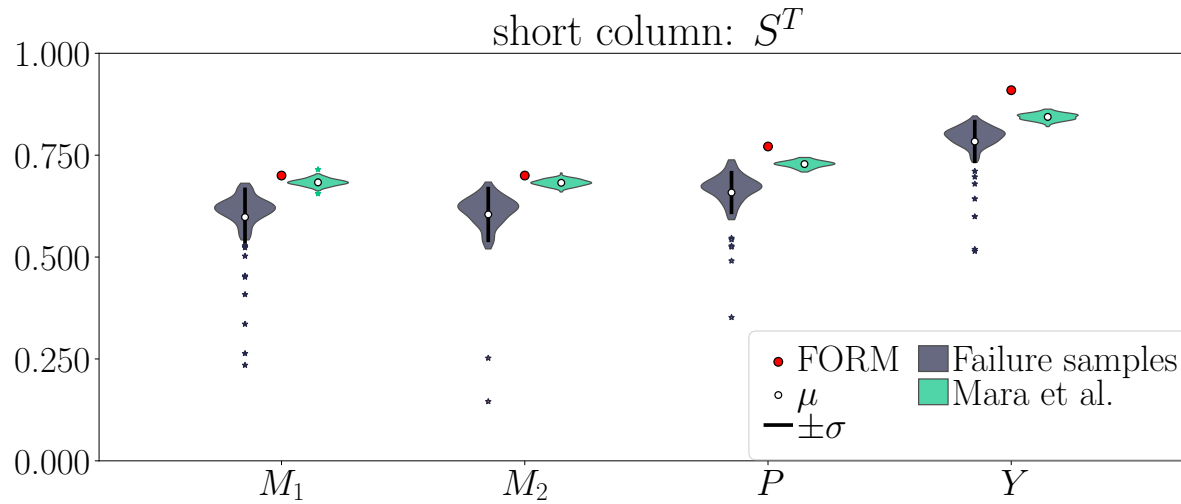
Reference solution: Peak-freeze estimator with $N = 10^6$ samples [Mara & Tarantola 2015]



First-order indices



Total-effect indices



Modified α -factors

Parameter	Normalized $\tilde{\alpha}^2$	Normalized $\tilde{\alpha}_{\text{ind}}^2$	γ^2
M_x	18%	6%	7%
M_y	18%	6%	7%
P	24%	19%	20%
Y	40%	69%	65%

Conclusion

- Extension of variance-based reliability sensitivity indices for dependent inputs
- Estimates of the sensitivity indices as byproducts of standard reliability methods
- Extensions of the FORM α -factors for problems with dependent inputs
- The γ -factors of [Der Kiureghian 2005] can be viewed as approximations of the modified factors $\tilde{\alpha}_{\text{ind}}$