Stochastic optimization - how to improve computational efficiency?

Christian Bucher

Center of Mechanics and Structural Dynamics
Vienna University of Technology
& DYNARDO GmbH, Vienna

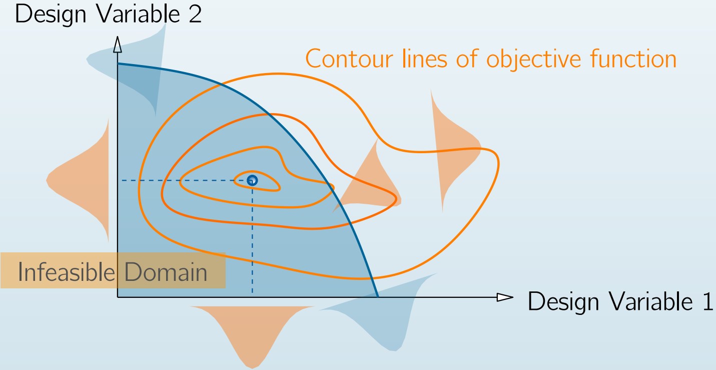
Presentation at Czech Technical University in Prague, 27 June 2014

Overview

- Introduction
- Stochastic design optimization
- Reliability analysis
- Response surface method
- Example - ten bar truss
  - Deterministic optimization
  - "Robust" optimization
- Concluding remarks

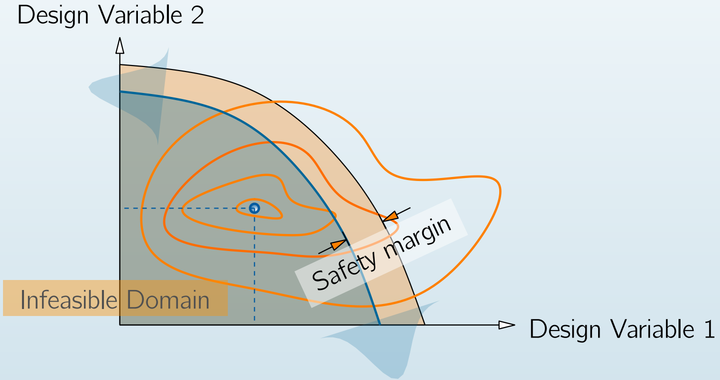
Uncertainties in optimization

- Design variables (e.g. manufacturing tolerances)
- Objective function (e.g. tolerances, external factors)
- Constraints (e.g. tolerances, external factors)



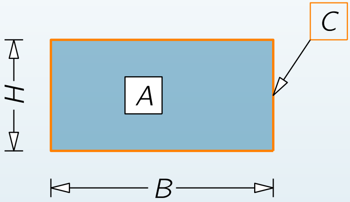
Traditional design approach

- Introduce "safety factors" into the constraints
- Leads to results satisfying safety requirement, but not necessarily optimal designs



Motivating Example 1

- Deterministic problem



  - Maximize area of a rectangle $A = B \cdot H$
  - Constraint: limited circumference $C = 2(B + H) - 4L \leq 0$
  - Solution: $B = H = L$

- Stochastic problem
  - B and H are random variables with mean values $\bar{B}$ and $\bar{H}$ and standard deviations σ_B and σ_H, respectively (e.g. normally distributed)

Motivating Example 2

- Maximize expected value of A

$$\bar{A} = \mathbf{E}[B \cdot H]$$

- or mean plus (or minus) multiple of standard deviation

$$\bar{A} \pm k \cdot \sigma_A$$

- Satisfy constraint condition with certain probability $P_C \approx 1$

$$\mathbf{Prob}[C \leq 0] = \mathbf{Prob}[2(B+H) - 4L \leq 0] \geq P_c$$

- L may be random, too
- Objective function

$$\bar{A} - k\sigma_A \rightarrow \text{Max.!}$$

- Constraint condition

$$\mathbf{Prob}[C \geq 0] = \mathbf{Prob}[2(B+H) - 4L \geq 0] \leq 1 - P_c = P_F$$

Numerical example 1

- Assume B and H correlated with ρ, L uncorrelated from B and H
- Assume all variables to be normally distributed, all have the same coefficient of variation COV
- Probability of constraint violation:

$$\bar{C} = 2(\bar{B} + \bar{H}) - 4\bar{L};\ \bar{L} = 1$$

$$\sigma_C^2 = 4(\sigma_B^2 + 2\rho\sigma_B\sigma_H + \sigma_H^2) + 16\sigma_L^2$$

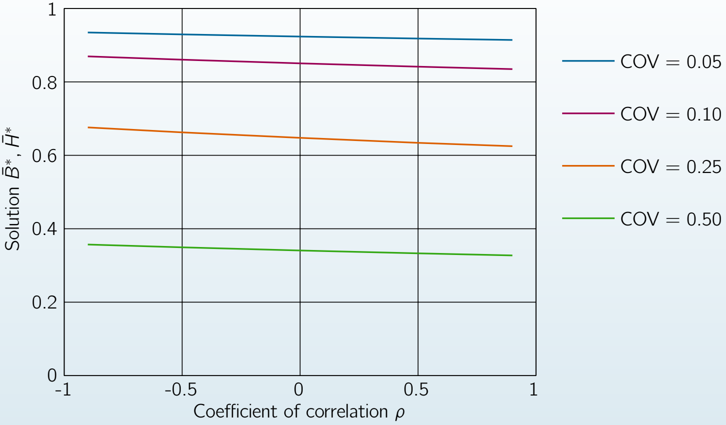
$$\textbf{Prob}[C > 0] = \Phi\left[\frac{\sigma_C}{\bar{C}}\right] \leq 0.1$$

- Expected value of area

$$\mathbf{E}[BH] = \bar{B}\bar{H} + \rho\sigma_B\sigma_H$$

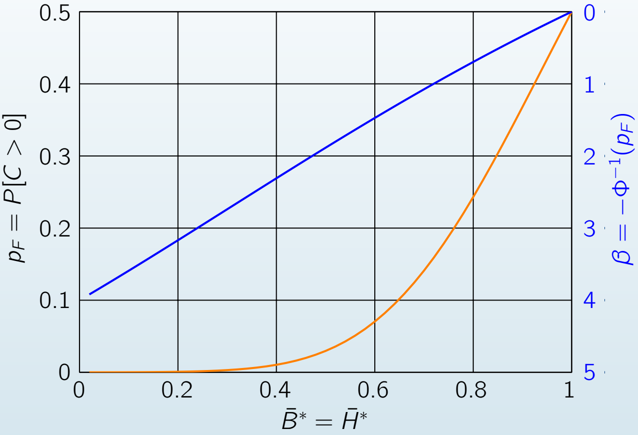
- Maximize expected area under probability constraint

Numerical example 2



Numerical example 3

- Different probability measures



- More "linear" dependence using β

Robustness

- Optimal design should not be sensitive to small variations of uncertain parameters
- For the case of stochastic uncertainty, robustness can be measured in statistical terms
- Choice of robustness measures
  - Variance-based: Global behavior, relatively frequent events
  - Probability-based: Specific behavior (often safety-related), rare events

Robustness measures 1

- Use probabilistic quantities in objective and/or constraints
- Add multiple of standard deviation of objective to the mean values to form new objective function

  $$R = \bar{f} + k\sigma_f$$

  Case $k = 0$ corresponds to game theory (large portfolio management, insurance, bank loans, fair gambling, ...)
- Use probability of constraint violation as new constraint

  $$P[f_k(\mathbf{x}^*)] > 0] \leq \varepsilon$$

  In safety related constraints, the probability level ε may be very small
- Better to use probability-related measure such as reliability index $\beta = -\Phi^{-1}(P[f_k(\mathbf{x}^*))$

Optimization strategies

- Conventional strategy
  - Optimize
  - check robustness/reliability
  - if not sufficient, optimize again with modified constraints
  - Easily implemented, may lead to non-optimal results
- Full RDO strategy
  - incorporate robustness/reliability measure into optimization problem
  - Solve directly for design satisfying robustness/reliability constraint
  - Leads directly to desired solution, may be very expensive

The need for speed ...

- Complex system (many parameters, computationally expensive, slow, ...)
- Needed: Fast and reasonably accurate response prediction (e.g. for real-time applications such as control systems)
- Possible choices:
  - Reduce model complexity based on essential physical features (reduced order model)
  - Replace model based on mathematical simplicity (metamodel)
- Stochastic analysis needs to be very efficient

Reduced order model

- Need to understand and represent physics
- May be applicable for many different load cases
- Very suitable for time dependent phenomena (structural dynamics, convection-diffusion processes)
- Can be difficult in the presence of strong nonlinearity
- Typical examples
  - Modal analysis
  - Proper orthogonal decomposition (POD)

Metamodel

- Mathematically formulated black box
- Suitable for arbitrarily nonlinear input-output relations
- Requires extensive training data
- Very difficult to extrapolate
- Time-dependent problems may be tricky
- Typical example: Linear response surface model

Model Correction factor method

- Hybrid between reduced order model and meta-model
- Replaces detailed mechanical model by simple mechanical model
- Adjust parameters of simple model by calibration to the detailed model
- Calibration based on probabilistic criteria (e.g. FORM)

Common properties

- Based on previous experience
  - Knowledge of physical processes
  - Acquired experience through training
- Limited range of applicability
  - Nonlinearities
  - Number of input variables
- NOTE: Approaches complement each other $\rightarrow$ Combination may be better than the sum of the individual parts!

Efficient reliability estimation procedures

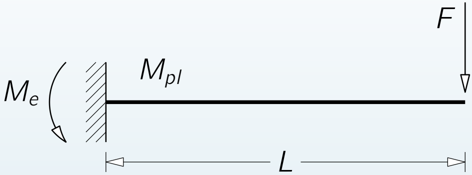
- FORM/SORM (approximation of the limit state)
- Monte Carlo methods
  - Crude MC - typically used as reference solution when developing a method
  - Importance sampling - requires identification of probabilistically relevant regions in the space of random variables
  - Asymptotic sampling - obtain reliability from results generated with artificially increased standard deviations of the basic variables

Essential Steps

- Sensitivity Analysis to determine the most relevant uncertain parameters
- Model order reduction for reducing the number of degrees of freedom and complexity
- Efficient probabilistic analysis to reduce number of samples
- Validation procedures to detect omissions and errors (includes expert judgement)

Reliability analysis

- Mechanical system



- Failure condition

$$\mathcal{F} = \{(F, L, M_{pl}) : FL \geq M_{pl}\} = \{(F, L, M_{pl}) : 1 - \frac{FL}{M_{pl}} \leq 0$$

- Failure probability

$$P(\mathcal{F}) = P[\{\mathbf{X} : g(\mathbf{X}) \leq 0\}$$

$$P(\mathcal{F}) = \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} I_g(x) f_{X_1 \ldots X_n} \mathrm{d}x_1 \ldots \mathrm{d}x_n$$

$I_g(x_1 \ldots x_n) = 1$ if $g(x_1 \ldots x_n) \leq 0$ and $I_g(.) = 0$ else

First-order reliability method (FORM)

- Rosenblatt-Transformation, e.g. for Nataf model

$$Y_i = \Phi^{-1}[F_{X_i}(X_i)]; \quad i = 1n$$

$$\mathbf{U} = \mathbf{L}^{-1}\mathbf{Y}; \quad \mathbf{C}_{\mathrm{YY}} = \mathbf{L}\mathbf{L}^T$$

- Inverse transformation

$$X_i = F_{X_i}^{-1}\left[\Phi\left(\sum_{k=1}^{n} L_{ik}U_k\right)\right]$$

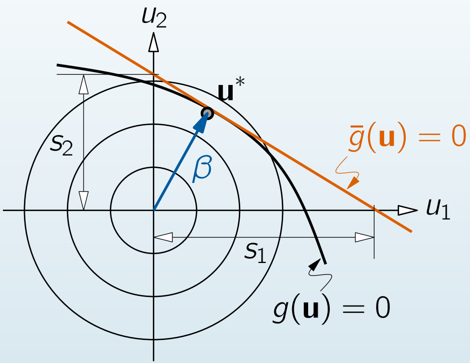
- Computation of design point

$$\mathbf{u}^* : \mathbf{u}^T\mathbf{u} \to \mathrm{Min.}; \quad \text{subject to: } g[\mathbf{x}(\mathbf{u})] = 0$$

- Linearize at the design point (in standard Gaussian space)

FORM procedure

- Find point $\mathbf{u}^*$ with minimal distance β from origin in standard Gaussian space

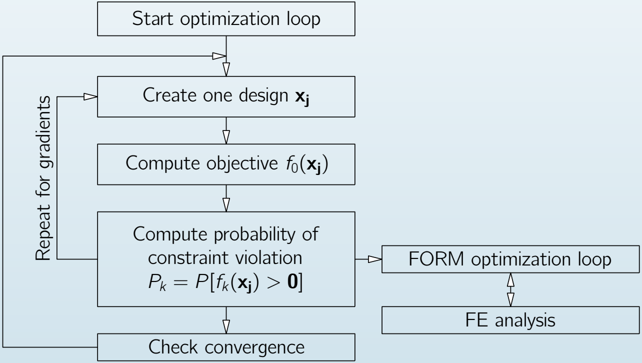


$$\bar{g} : -\sum_{i=1}^{n} \frac{u_i}{s_i} + 1 = 0; \quad \sum_{i=1}^{n} \frac{1}{s_i^2} = \frac{1}{\beta^2}$$

$$P(\mathcal{F}) = \Phi(-\beta)$$

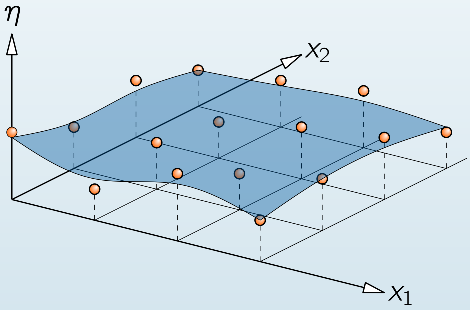
Optimization loop

- Outer optimization loop controls the structural design
- Probability of constraint violation computed by FORM
- Inner optimization driven by random variables
- Both loops need gradients



Response surface method

- Reduce computational effort by replacing expensive FE analyses
- Establish meta-models in terms of simple mathematical functions
- Fit model parameters to FE solution using regression analysis



Regression

- Adjust a model to experiments

$$Y = f(X, \mathbf{p})$$

- Set of parameters

$$\mathbf{p} = [p_1, p_2, \dots, p_n]^T$$

- Experimental values for input X and output Y

$$(X^{(k)}, Y^{(k)}), k = 1 \dots m$$

- Search for best model by minimizing the residual

$$S(\mathbf{p}) = \sum_{k=1}^{m} \left[Y^{(k)} - f(X^{(k)}, \mathbf{p}) \right]^2; \quad \mathbf{p}^* = \text{argmin } S(\mathbf{p})$$

Linear regression

- Linear dependence on parameters (not on variables!)

$$f(X, \mathbf{p}) = \sum_{i=1}^{n} p_i g_i(X)$$

- Necessary condition for a minimum

$$\frac{\partial S}{\partial p_j} = 0; \quad j = 1 \dots n$$

- Solution

$$\sum_{k=1}^{m} \left\{ g_j(X^k)[Y^k - \sum_{i=1}^{n} p_i g_i(X^k)] \right\} = 0; \quad j = 1 \dots n$$

$$\mathbf{Qp} = \mathbf{q}$$

Quality of regression

- Coeffient of determination (CoD): correlation between experimental data and model predictions

$$R^2 = \left(\frac{\mathbf{E}[Y \cdot Z]}{\sigma_Y \sigma_Z} \right)^2 ; Z = \sum_{i=1}^{n} p_i g_i(X)$$

- Adjusted (reduced) CoD for small sample sizes (penalize overfitting)

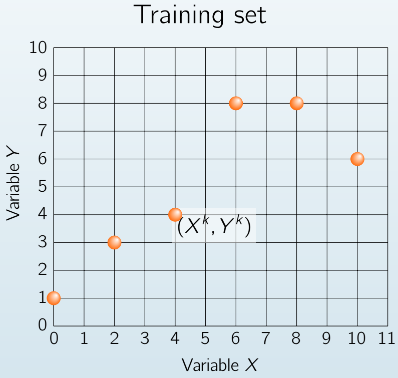
$$R^2_{adj} = R^2 - \frac{n-1}{m-n} \left(1 - R^2\right)$$

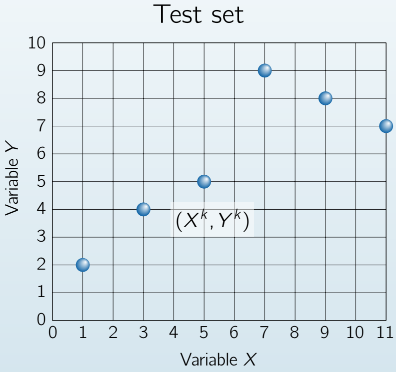
- If an additional test data set T is available: Coefficient of Prognosis (CoP)

$$\text{CoP} = \left(\frac{\mathbf{E}[T \cdot Z_T]}{\sigma_Y \sigma_Z} \right)^2 ;\; Z_T = \sum_{i=1}^{n} p_i g_i(X_T);\; 0 \leq \text{CoP} \leq 1$$

Example

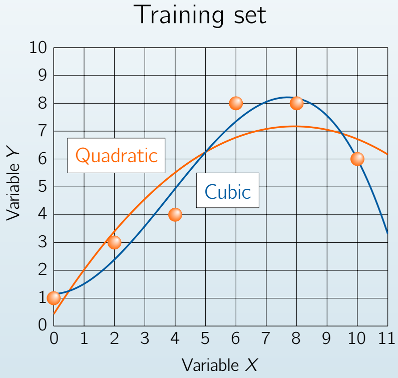
- Adapt polynomial function to 6 data points

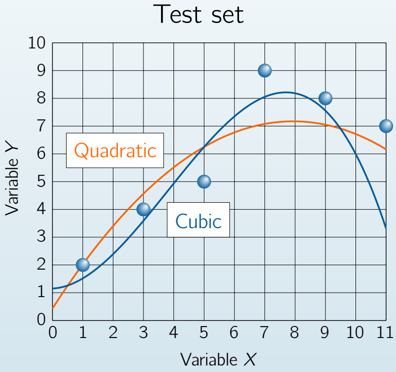




Regression result

- Adjust model to training data, cross-check with test data

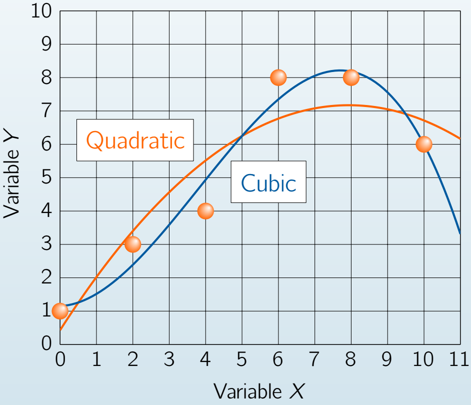




Quality depending on model order

- Repeat regression and test for different polynomial orders
- Quadratic model has best prediction capability

n	R^2	CoP	R^2_{adj}
1	0.83	0.83	0.79
2	0.93	0.92	0.88
3	0.98	0.81	0.94
4	0.98	0.73	0.90
5	1.00	0.70	-



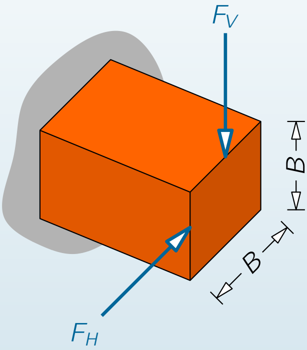
What to approximate?

- For reliability analysis, the limit state function $g(\mathbf{X})$ is needed
- Immediate approximation of g may introduce unwanted nonlinear dependencies on input variables
- Example: shear stresses in a console with square cross section

$$\tau_{xy} = \frac{3F_H}{2B^2};\ \tau_{xz} = \frac{3F_V}{2B^2}$$

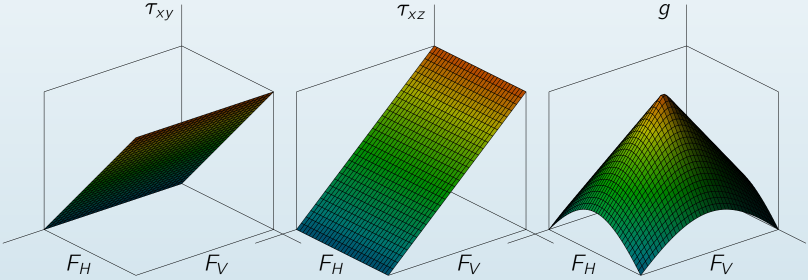
- Failure criterion (v. Mises)

$$g(F_H, F_V) = \beta_F - \sqrt{3(\tau_{xy}^2 + \tau_{xz}^2)} \leq 0$$



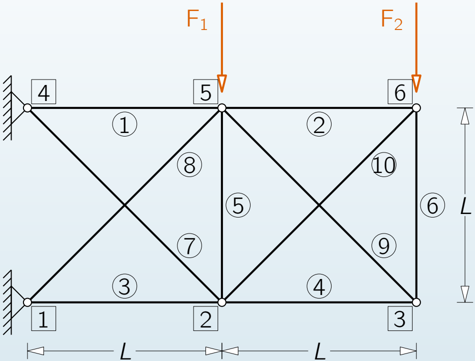
Comparison

- Stresses τ_{xy} and τ_{xz} are linear functions of F_H and F_v
- Limit state function $g(F_H, F_V)$ is highly nonlinear and contains a singularity
- $\rightarrow$ Try to use "easy" formulation for constraint functions



Example - ten bar truss

- Configuration: $L = 360$, $F_1 = F_2 = 100000$



- Objective: Minimize structural mass
- Constraints:
  All member stresses < 25000
  All nodal displacements < 2

Deterministic optimization

- Gradient-based solver (Conmin)

DEMO

Optimal design

- Objective function: $m = 5065$
- Cross sectional areas

Member	Area	Member	Area
1	30.23	6	1.12
2	0.10	7	7.46
3	23.77	8	20.82
4	15.03	9	21.38
5	0.10	10	0.10

- Active constraints: no member stress, displacements in nodes 3 and 6
- Effort: $\approx$ 200 FE analyses

Robustness evaluation

- Take into account randomness of loads (F_1 and F_2 independent, log-normally distributed, coefficient of variation = 0.3)
- Take into account randomness of cross sections (all independent, normally distributed, coefficient of variation = 0.15)
- Compute variability of constraints
- Estimate probability of constraint violation(s) using a distribution hypothesis (Gaussian)
- High probability of violating active constraints, but also 3 inactive ones
- Effort: 100 FE analyses

Member	P_σ	Member	P_σ
1	0.00	6	0.00
2	0.00	7	0.17
3	0.00	8	0.00
4	0.00	9	0.00
5	0.51	10	0.00

Node	P_u	Node	P_v
2	0.00	2	0.00
3	0.00	3	0.53
5	0.00	5	0.27
6	0.00	6	0.53

Upscaled deterministic solution

- Increase cross sectional areas of design uniformly to match reliability constraint
- Use repeated FORM analysis

COV	0.10	0.2	0.30
m_{opt}	7167	8966	11347

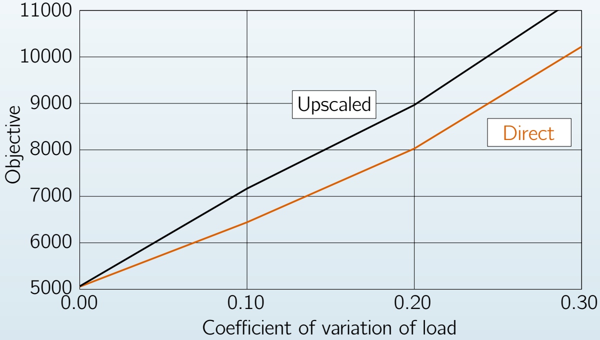
- Effort: $\approx$ 300 FE analyses for each β

Reliability-based optimization

- Model randomness of loads (F_1 and F_2 independent, log-normally distributed)
- Model randomness of cross sections (all independent, normally distributed)
- Accept constraint violation(s) with a probability corresponding to a reliability index $\beta = 3$
- Use FORM to obtain β for all designs during the optimization process
- Best designs depend on the coefficient of variation of F_1 and F_2
- Effort: $\approx$ 60.000 FE analyses (factor 300 vs. deterministic case)

Robust optima

COV	0.10	0.2	0.30
m_{opt}	6682	8532	11092



- Direct: Computational effort: 60.000 FE runs for 1 COV
- Upscaled: Computational effort: 1500 FE runs for 1 COV

Response surface

- Use 500 support points in 12-dimensional space
- Assess quality of regression
- Cross validation by splitting samples 2/1
- Compute coefficient of prognosis (CoP) from applying regression model to unused samples

Quantity	σ_1	σ_2	σ_3	σ_4	σ_5
CoP	0.999	0.873	0.999	0.999	0.955

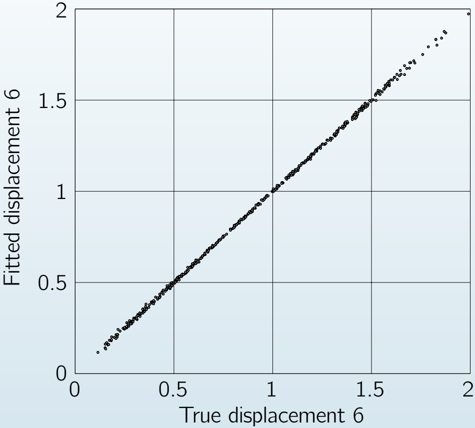
Quantity	σ_6	σ_7	σ_8	σ_9	σ_{10}
CoP	0.870	0.999	0.999	0.999	0.879

Quantity	u_2	u_3	u_5	u_6
CoP	0.999	0.998	0.999	0.999

Quantity	v_2	v_3	v_5	v_6
CoP	0.999	0.999	0.999	0.999

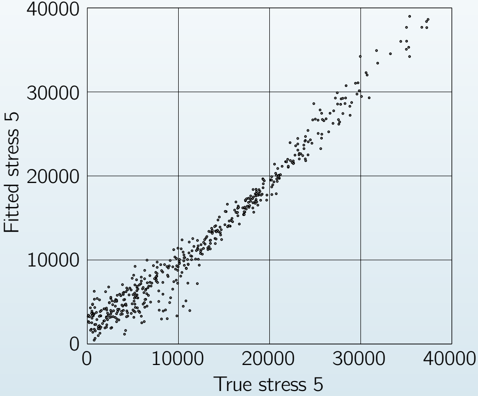
Approximation of displacements

- Vertical displacement of node 6
- Compare displacement computed from FE analysis to response surface results (500 random samples)



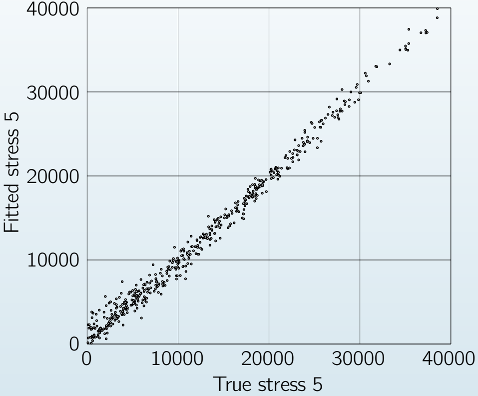
Approximation of stresses

- Stress in member 5
- Compare stress computed from FE analysis to response surface results (full quadratic, 500 random samples)



Alternative approximation of stresses

- Stress in member 5
- Compare stress computed from FE analysis to response surface results (thin plate spline, 500 random samples)



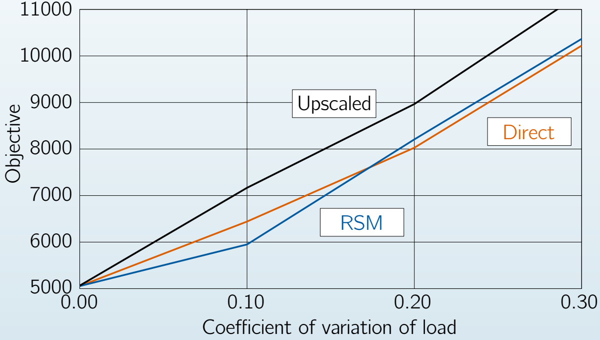
Reliability-bsed optimization using RSM

- Gradient-based optimizer (Conmin), FORM based on response surfaces

DEMO

Robust optima from RSM

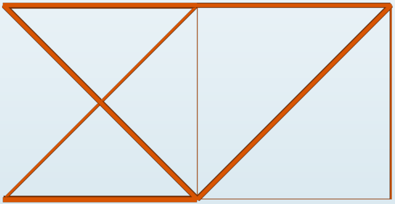
COV	0.10	0.2	0.30
m_{opt}	6108	8741	11261



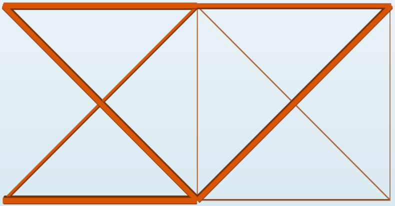
- Computational effort reduced by a factor of 100 for one COV (300 for 3 COVs)

Compare robust designs

- All elements are strengthened with larger COV of load
- Relative member cross section ratios remain similar (not too different from upscaled deterministic solution)



COV = 0.1



COV = 0.3

Benefits of robust optimization

- Avoids highly specialized designs, reduces imperfection sensitivity
- Naturally includes statistical uncertainties into the design optimization process
- Allows the inclusion of quality control measures (manufacturing, maintenance) into the design process
- BUT: computationally very expensive unless based on approximations such as response surface models