Risk Optimization of Trusses Using a New Gradient Estimation Method

Wellison J. S. Gomes
Assistant professor, Dept. of Civil Engineering, Federal University of Santa Catarina, Florianopolis, Brazil

ABSTRACT: When dealing with structural risk optimization by means of Monte Carlo simulation methods, the total expected cost usually becomes a noisy function and it is not possible to directly calculate its derivatives. In fact, even the estimation of these derivatives becomes a challenging task. On the other hand, gradient-based optimization methods are among the most efficient ones for this kind of problem, but they require derivatives. In this paper, a new method, which allows estimating gradients of the total expected cost with respect to design variables, is used to perform risk optimization of trusses by a gradient-based optimization method. The proposed methodology is employed in the solution of a bi-dimensional and a spatial truss. The results show that this methodology presents a significantly smaller computational cost when compared to a similar procedure using finite differences and indicates that the quality of the gradients is slightly better as well.

In the context of structural optimization, the so-called risk optimization formulation has emerged as a formulation which allows finding the optimal tradeoff point between the goals of economy and safety. This is accomplished by including expected costs of failure into the total cost of the structural system and by determining the values of the design variables which minimize the so-called total expected cost. Risk optimization solutions usually require many reliability analyses and each reliability computation requires many structural analyses. Thus, unless the problem at hand is very simple, this formulation can lead to very high computational costs.

Assuming that the objective function satisfies certain conditions of convexity and smoothness, one efficient way to deal with structural risk optimization problems is by means of nonlinear programming methods. However, most of the nonlinear programming methods available in the literature require the estimation of objective function derivatives with respect to the design variables and, when failure probabilities are computed by simulation, the estimation of gradients of the total expected cost becomes a challenging task.

In this paper, a nonlinear programming method is applied to the solution of risk optimization of trusses. In order to estimate the required derivatives, a new gradient estimation method is proposed. The present paper describes the proposed gradient estimation method and explains how to apply this method for the specific case of bi and three-dimensional trusses, considering that linear structural analyses are performed and constraints are given in terms of tensile and Euler buckling stresses.

The remainder of this paper is organized as follows. First, the risk optimization problems are formulated, in a general way. Then, a brief description of the proposed gradient estimation method is presented in Section 2. Section 3 defines the limit state equations and explains how to evaluate them and how to determine their derivatives using direct differentiation. Finally, the presented methodology is employed in the solution of two problems, adapted from the literature, and some conclusions are drawn.
1. RISK OPTIMIZATION FORMULATION

1.1. Structural reliability analysis

Let \( X \) and \( d \) be vectors of structural system parameters. Vector \( X \) represents all random or uncertain system parameters and vector \( d \) contains all deterministic design variables and may also include some parameters of random variables in \( X \).

The existence of uncertainty implies possibilities of undesirable structural responses. The boundary between desirable and undesirable structural responses is defined by limit state functions \( g(X, d) \), in such a way that the failure and survival domains, \( \Omega_f \) and \( \Omega_s \), respectively, are given by:

\[
\Omega_f = \{ x \mid g(x, d) \leq 0 \} \\
\Omega_s = \{ x \mid g(x, d) > 0 \}
\]

(1)

Each limit state describes one possible failure mode of the structure. The probability of undesirable structural responses, or probability of failure, for each failure mode is defined as:

\[
P_f(x, d) = P[X \in \Omega_f] = \int_{\Omega_f} f_X(x) \, dx
\]

(2)

where \( f_X(X) \) is the joint probability density function of vector \( X \).

Probabilities of failure can be evaluated using structural reliability methods such as First and Second Order Reliability Methods (FORM and SORM) and Monte Carlo simulation (Madsen et al. 1986; Melchers 1999).

When simple Monte Carlo simulation (SMC) is employed, failure probabilities can be estimated via Eq. (3), by randomly generating \( n_{\text{samp}} \) samples of \( X \) according to its joint distribution \( f_X(X) \) and by considering a so-called indicator function, \( I(x, d) \), which, for given values of \( d \), is equal to one if \( x \) belongs to the failure domain and zero otherwise.

\[
P_f(x, d) = E[I(x, d)] = \frac{1}{n_{\text{samp}}} \sum_{i=1}^{n_{\text{samp}}} I(x_i, d)
\]

(3)

SMC is, in general, less efficient than FORM or SORM, but it can be easily applied to any kind of reliability problem. In order to improve its efficiency, many versions of the Monte Carlo method have been proposed in the literature, e.g., Au and Beck (2001) and Engelund and Rackwitz (1993). In this paper, the Importance Sampling Monte Carlo method (ISMC) is adopted (Engelund and Rackwitz, 1993), since it is almost as general as, and more efficient than, the simple Monte Carlo method.

In ISMC, the samples are generated according to a sampling function, \( h_X(X) \), and the indicator function must be weighted so the probability with respect to the original joint distribution can be obtained. Eq. (3) becomes:

\[
P_f(x, d) \approx \frac{1}{n_{\text{samp}}} \sum_{i=1}^{n_{\text{samp}}} \frac{I(x_i, d) f_X(x_i)}{h_X(x_i)}
\]

(4)

The sampling function can be constructed by using information about the design points, as described in Bourgund and Bucher (1986), and the design points can be obtained by FORM. This is the procedure adopted herein.

If the failure probability is computed by Eq. (3) or Eq. (4), its analytical partial derivative with respect to, for example, some design variable, \( d_j \), involves the derivative of the indicator function, \( I(x, d) \), with respect to \( d_j \). This means that, for given values of vector \( d \), this partial derivative will always be equal to zero, unless \( g(x, d) \) is exactly equal to zero for at least one of the points defined by the \( n_{\text{samp}} \) samples. Thus, a prohibitive, infinite, number of samples is required when looking for a continuous description of the derivative over the design space or the random variables space.

This is the reason why, in this case, the gradient of the probability of failure cannot be directly obtained and must be estimated by alternative approaches, such as the method proposed herein.

1.2. Risk optimization problem

For each failure mode of the structure, an expected cost of failure, or failure risk, given by the failure probability multiplied by the respective monetary consequences of failure,
\( C_{\text{fail}} \) can be computed. By adding up the expected costs of failure and the total cost of the structure, the so-called total expected cost is obtained. The total cost can include, for example, costs of material, construction and maintenance.

Considering that all costs are null, except the initial cost, \( C_{\text{ini}} \), and the expected costs of failure, one obtains the objective function adopted herein:

\[
C_T(d) = C_{\text{ini}} + \sum_{i} C_{\text{fail}} \cdot P_f(x, d) \tag{5}
\]

For a unitary cost per weight unit ($/kg), the initial cost corresponds to the weight of the structure, which is calculate by considering a material density \( \rho = 0.00785\text{kg/cm}^3 \). The cost of failure is adopted as ten times the initial cost for a given reference configuration, \( d_0 \).

Thus, the resulting risk optimization problem can be defined as:

\[
d^* = \arg\min_{d \in S} C_T(d) : d \in S
\tag{6}
\]

where \( S = \{d_{\text{min}} \leq d \leq d_{\text{max}}\} \) is a set of constraints on the design variables.

2. GRADIENT ESTIMATION METHOD

If the failure probabilities in Eq. (5) are computed by simulation as explained in section 1.1, \( C_T(d) \) becomes a noisy function. Its derivative can be estimated by finite differences, but the results are usually affected by its discontinuities and the computational cost of the estimation highly increases as the number of design or random variables is increased.

On the other hand, if the derivatives of the failure probabilities with respect to the design variables can be computed, the chain rule can be applied to obtain the respective gradients of the total expected cost.

The proposed method allows estimating the derivatives of the probability of failure with respect to any design variable. It is based on using some information which is usually disregarded in reliability analysis by simulation: the derivatives of the limit state functions. Thus, it can be applied only when these derivatives are available at a suitable computational cost.

The method consists of, for each sample and each direction, obtaining an approximation of the crossing point, that is, the point at which the limit state function equals zero.

In order to understand how the proposed method works, let us look at the one-dimensional case, that is, the case with just one design variable, \( d \). Let us also consider that \( d \) assumes a certain value, \( \bar{d} \), at which the gradient must be evaluated.

For a given sample, \( i \), the approximate crossing point can be obtained by taking a first-order Taylor series expansion around \( d_0 \) and by determining the value of \( d \) for which the expansion equals zero. This value can be rewritten as \( \bar{d} + \delta \) and the crossing point for each sample \( i \) can be represented in a relative way by \( \delta_i \), as illustrated in Figure 1. To construct the expansions, both the limit state function, \( g = g(x, d) \), and its first-order derivative with respect to \( d \) must be evaluated at \( \bar{d} \), for each sample, but this information is directly obtained as a product of one simulation run, the same required to compute \( C_T(\bar{d}) \) during the optimization process.

![Figure 1 - Approximated crossing points for some samples.](image)

At each crossing point, a sample goes inside or outside the failure domain and the probability of failure is accordingly increased or decreased, which can be defined considering the sign of the limit state function derivative. This information can be used to determine relative values of the
failure probability at each crossing point. In SMC, the variation of the probability of failure estimate at this point is equal to $1/n_{\text{samp}}$; in the ISMC case, the variation is equal to $1/n_{\text{samp}}$ multiplied by the sampling weight associated with that sample.

After determining the crossing points for all samples and the respective relative failure probability values, those points which are closest to the point of interest are used to estimate the derivatives at that point. For example, the crossing points whose distance to the point of interest is less than or equal to $\delta_{\text{lim}}$ can be chosen, where $\delta_{\text{lim}}$ is a parameter, similar to the perturbation used in finite differences. A minimum number of chosen crossing points, e.g., two points, can be also specified.

The procedure continues by fitting a linear function, $P^*_f(\delta) = a + b \cdot \delta$, to the chosen crossing points via least squares method, as illustrated in Figure 2 for the SMC case. The required gradient is approximately equal to $b$ and it is noted that $b$ is affected only by the variation of the probability of failure in the vicinity of $\delta$, not by the exact values of the failure probabilities, which simplifies the computations. The estimated gradient tends to its exact value as the number of samples increases.

Figure 2: Linear fitting by least squares considering four crossing points.

This procedure requires only one simulation run to compute the gradients with respect to all design variables, in contrast to finite differences where the number of required simulations depends on the number of variables. It also leads to estimates which are smoother than those obtained by finite differences, an important characteristic when performing optimization by gradient-based methods.

Furthermore, if the limit state function at hand is highly nonlinear, the estimated crossing points can be improved by applying some iterations of the Newton's method. In this paper, only one iteration is performed, which results in the previously described procedure.

3. LIMIT STATE FUNCTIONS AND DERIVATIVES

For a given structure, considering it as a series system and describing the failure of each member by a specific limit state function, system failure occurs whenever at least one of the limit state functions assumes a value less than or equal to zero. Based on these assumptions, for a given truss structure subject to some loading conditions, the following steps are performed to evaluate its failure probability.

First, the stiffness matrix method (a linear structural analysis) is applied to determine the axial internal forces, $N_{ix}$, in each member of the structure.

Then, for each bar, a limit state function is defined by its ultimate capacity, according to Aoues and Chateauneuf (2010). The tensile capacity is given by the yield stress $f_{yt}$ and the compression capacity is given by the Euler buckling stress $f_{yc} = \frac{\pi^2 E}{\lambda^2}$, where $E$ is the elasticity modulus of the material and $\lambda$ is the slenderness ratio. The limit state function for the $i$th bar is given by the minimum value between the following:

$$g_i^*(x, d) = f_{yt}^i(x) \cdot A_i(d) - N_{ix}^i(x, d)$$

$$g_i^*(x, d) = f_{yc}^i(x) \cdot A_i(d) - N_{ix}^i(x, d)$$

where: $A_i$ is the area of the $i$th bar, which is a function of the design variables or a design variable itself; $f_{yt}$ and $f_{yc}$ are related to properties of the material, thus they depend on and they are random variables; and the axial internal force can
be a function of random loads and design variables.

The limit state functions are evaluated for all samples within the ISMC procedure, and the results are combined to estimate the probability of failure of the system.

As previously explained, the gradient estimation method requires derivatives of the limit state functions with respect to design and random variables. The determination of the design points via FORM, within the ISMC, requires derivatives of the limit state functions with respect to random variables. These derivatives are obtained via chain rule as explained in Enevoldsen and Sorensen (1994). In the literature, the implementation of the chain rule within the stiffness matrix method or some other methods, e.g., finite element method, is often called direct differentiation method (DDM), see Zhang and Der Kiureghian (1993) or Der Kiureghian et al. (2006). DDM has been successfully applied in many cases, mainly within the context of optimization, which is usually a computationally demanding task.

4. EXAMPLES

The presented methodology is used within the optimization of two cases adapted from the literature: a bi-dimensional five-bar truss and a twenty-five-bar spatial truss. The results are compared with those obtained by using a similar procedure, but with derivatives of CT obtained by finite differences. As the gradient estimation method is based on crossing points, it is identified as CP from now on. Solutions obtained by finite differences are referred to as reference solutions.

The optimization is performed by the BFGS method and a linear search based on the Davies, Swann and Campey algorithm (Himmelblau, 1972). In all cases, a maximum of ten iterations of the BFGS algorithm is allowed.

The comparisons are made according to the following procedure. First, the convergence of CT and of its derivative with respect to some design variables is studied by varying the number of simulations used within the ISMC. Then, a number of samples is chosen, considering the competing goals of precision and computational effort. Finally, the optimization is carried out using the chosen number of samples.

All computational codes were developed in MATLAB and the computational time refers to conventional single-thread computation using an Intel® Core™ i7 CPU 860 @ 2.80GHz processor.

4.1. Five-bar plane truss

This case consists of a five-bar bi-dimensional truss, previously studied by Aoues and Chateauneuf (2010) and illustrated in Figure 3. Five design variables are considered, which correspond to the cross-sectional areas of each bar. The random variables are the material tensile strength, \( f_{yt} \), and the elasticity modulus, \( E \), of each bar, as well as the applied forces \( F \) and \( P \), leading to twelve random variables in total. All random variables are assumed to follow normal distributions whose parameters are described in Table 1.

![Figure 3: Five-bar bi-dimensional truss.](image)

<table>
<thead>
<tr>
<th>Random Variable</th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>( f_{yt} ) (kN/cm²)</td>
<td>17.2</td>
<td>2.58</td>
</tr>
<tr>
<td>( E ) (kN/cm²)</td>
<td>6895</td>
<td>689.5</td>
</tr>
<tr>
<td>( F ) (kN)</td>
<td>20</td>
<td>6</td>
</tr>
<tr>
<td>( P ) (kN)</td>
<td>15</td>
<td>3</td>
</tr>
</tbody>
</table>

The starting point, \( d_{ini} = \{A_1, A_2, \ldots, A_5\} \), for the optimization algorithm, which is also the
reference configuration, $d_0$, is fixed to $A_i=25\text{cm}^2$, with an initial cost $C_{ini}=348.64$ and a cost of failure $C_{fail}=3486.39$. Convergence of the total expected costs and of its derivatives with respect to the area of the fourth bar, one of the most critical elements in the reference configuration, is studied in Figure 4 and Figure 5. Note that both methods use the same approach to calculate the failure probability. Thus, there are no differences between the total expected costs computed by each method.

![Figure 4: Convergence of the total expected cost for the five-bar truss problem.](image)

For both methods, the maximum number of allowed iterations was reached, which indicates that better configurations could still be obtained by the optimization procedure. However, these results are sufficient for comparison purposes. Table 2 shows that the CP method, in this example, requires less than half the NCLS required by the reference solution. This is reflected in the total computational times, which indicates that the procedure for gradient estimation does not add a significant computational cost to the solution and allows avoiding many simulation runs during the optimization. Also, it is seen that the optimal design provided by the CP method is slightly better than by finite differences, which indicates higher accuracy of the estimated gradients.

### 4.2. 25-bar spatial truss

This case consists of a 25-bar spatial truss which has been studied in the literature from a deterministic point of view, for example, in Kaveh and Talatahari (2009). The structure is illustrated in Figure 6.

The 25 bars are divided in eight groups, as described in Table 3, and each group has its own properties, resulting in a total of eight design variables, one cross-sectional area per group, and eighteen random variables. The random variables are the material tensile strength, $f_{yt}$, the elasticity modulus, $E$, of each group and the applied forces $P_y$ and $P_z$. The concentrated force $P_y$ is applied at node (1), in the positive $y$ direction, and at node

| Table 2: Numerical results for the five-bar truss problem. |
|---------------------------------|-----------------|-----------------|
| $A_1$ (cm²) | Finite Diff. (reference) | CP Method |
| 20.80 | 22.57 |
| $A_2$ (cm²) | 4.62 | 4.01 |
| $A_3$ (cm²) | 4.57 | 1.51 |
| $A_4$ (cm²) | 26.44 | 28.99 |
| $A_5$ (cm²) | 28.72 | 28.03 |
| $C_T$ | 227.67 | 212.70 |
| NCLS | 1911298 | 821278 |
| Time (min) | 36.98 | 16.59 |
| Optimum $P_f$ | $4.2506 \times 10^{-3}$ | $1.0261 \times 10^{-3}$ |
(2), in the negative y direction. $P_z$ is applied at nodes (1) and (2), in the negative direction. This loading condition corresponds to the case 1 presented in Kaveh and Talatahari (2009).

Figure 6: 25-bar spatial truss (Kaveh and Talatahari, 2009).

Table 3: Groups of elements with common properties.

<table>
<thead>
<tr>
<th>Group</th>
<th>Bars</th>
<th>Group</th>
<th>Bars</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>5</td>
<td>12-13</td>
</tr>
<tr>
<td>2</td>
<td>2-5</td>
<td>6</td>
<td>14-17</td>
</tr>
<tr>
<td>3</td>
<td>6-9</td>
<td>7</td>
<td>18-21</td>
</tr>
<tr>
<td>4</td>
<td>10-11</td>
<td>8</td>
<td>22-25</td>
</tr>
</tbody>
</table>

All random variables are assumed to follow normal distributions, as described in Table 4.

Table 4: Random variables for the 25-bar truss problem.

<table>
<thead>
<tr>
<th>Random Variable</th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>$f_{yt}$ (kN/cm²)</td>
<td>27.58</td>
<td>2.758</td>
</tr>
<tr>
<td>$E$ (kN/cm²)</td>
<td>6895</td>
<td>689.5</td>
</tr>
<tr>
<td>$P_y$ (kN)</td>
<td>89</td>
<td>8.9</td>
</tr>
<tr>
<td>$P_z$ (kN)</td>
<td>22.25</td>
<td>4.45</td>
</tr>
</tbody>
</table>

The starting point, $d_{ini} = \{A_1, A_2, \ldots, A_8\}$, for the optimization algorithm, is fixed to $A_i = 50\, \text{cm}^2$, while $d_0$ is given by $A_i = 10\, \text{cm}^2$, with $C_{ini} = 659.42$ and $C_{fail} = 6594.24$. Convergence of the total expected costs and of its derivatives with respect to the area of the seventh group, one of the most critical groups for the initial configuration, is shown in Figure 7 and Figure 8.

![Figure 7: Convergence of the total expected cost for the 25-bar truss problem.](image)

In this case, the analysis of the derivative of $C_T$ could lead to choose a very small number of samples. However, by verifying the convergence of the derivative of the failure probability with respect to the same design variable, Figure 9, it is decided to adopt $n_{samp} = 5000$.

Figure 8: Convergence of the total expected cost w.r.t. the seventh design variable.

![Figure 8: Convergence of the total expected cost w.r.t. the seventh design variable.](image)

Figure 9: Convergence of the total expected cost for the 25-bar truss problem.

Table 5 shows the optimal design and other results obtained by the optimization.
Table 5: Numerical results for the 25-bar truss problem.

<table>
<thead>
<tr>
<th></th>
<th>Finite Diff. (reference)</th>
<th>CP Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A_1 (\text{cm}^2)$</td>
<td>47.00</td>
<td>43.16</td>
</tr>
<tr>
<td>$A_2-A_5 (\text{cm}^2)$</td>
<td>43.18</td>
<td>44.14</td>
</tr>
<tr>
<td>$A_6-A_9 (\text{cm}^2)$</td>
<td>39.09</td>
<td>41.08</td>
</tr>
<tr>
<td>$A_{10}-A_{11} (\text{cm}^2)$</td>
<td>44.01</td>
<td>36.11</td>
</tr>
<tr>
<td>$A_{12}-A_{13} (\text{cm}^2)$</td>
<td>44.01</td>
<td>36.08</td>
</tr>
<tr>
<td>$A_{14}-A_{17} (\text{cm}^2)$</td>
<td>21.36</td>
<td>22.87</td>
</tr>
<tr>
<td>$A_{18}-A_{21} (\text{cm}^2)$</td>
<td>54.68</td>
<td>54.63</td>
</tr>
<tr>
<td>$A_{22}-A_{25} (\text{cm}^2)$</td>
<td>30.07</td>
<td>22.07</td>
</tr>
<tr>
<td>$C_T$</td>
<td>2637.97</td>
<td>2489.42114</td>
</tr>
<tr>
<td>NCLS</td>
<td>1412572</td>
<td>388018</td>
</tr>
<tr>
<td>Time (min)</td>
<td>113.94</td>
<td>31.07</td>
</tr>
<tr>
<td>Optimum $P_f$</td>
<td>$1.5674 \times 10^{-3}$</td>
<td>$9.8444 \times 10^{-3}$</td>
</tr>
</tbody>
</table>

In this case, the performance of the CP method was still better than in the first case, when compared to the reference solution, due to the higher number of design and random variables. The computational time for the CP method was about one third of that required by finite differences. The final solution, in terms of total expected costs, was also slightly better than the reference.

5. CONCLUSIONS
In this paper, a gradient estimation method, previously proposed by the author, was applied within a gradient-based risk optimization scheme in order to determine optimal cross sectional areas of trusses. The presented methodology was applied in the solution of one bi-dimensional and one spatial truss, and the results were compared with those obtained by using finite differences.

In terms of computational effort, the results obtained by the presented methodology were considerably better than those obtained by finite differences. The results in terms of optimal design were slightly better.

From the author’s knowledge, this is the only gradient estimation method, apart from finite differences, which can be applied in the case where total expected costs are computed by simulation. Focusing on risk optimization of real structures, efforts should be concentrated in further improving this method as well as the optimization procedure itself.

ACKNOWLEDGEMENTS
Sponsorship of this research project by the Brazilian Council for Research and Development (CNPq) through grant number 442183/2014-3 is greatly acknowledged.

6. REFERENCES


