Horseracing Simulation Algorithm for evaluation of small failure probabilities

Lambros Katafygiotis & Konstantin Zuev

Department of Civil Engineering
Hong Kong University of Science and Technology

zuev@ust.hk, lambros@ust.hk

23 June, 2009

Outline

Outline

- Basic Idea
- Horseracing Simulation scheme
- Implementation issues
  - Sampling
  - Construction of the empirical CDF and its updating
  - Stopping criterion
- Horseracing Simulation Algorithm
- Example
  - CAARC standard tall building model
  - Wind excitation
  - Description of the failure domain
  - Simulation results

Basic Idea

Basic Idea

Reliability Problem: to compute the failure probability $p_\Omega = \int_\Omega \pi_0(x)dx$

Basic Idea

Reliability Problem: to compute the failure probability $p_\Omega = \int_\Omega \pi_0(x)dx$

Failure domain: $\Omega = \{x \in \mathbb{R}^N | \, g(x) > z^*\}$

Basic Idea

Reliability Problem: to compute the failure probability $p_\Omega = \int_\Omega \pi_0(x)dx$

Failure domain: $\Omega = \{x \in \mathbb{R}^N | \, g(x) > z^*\}$

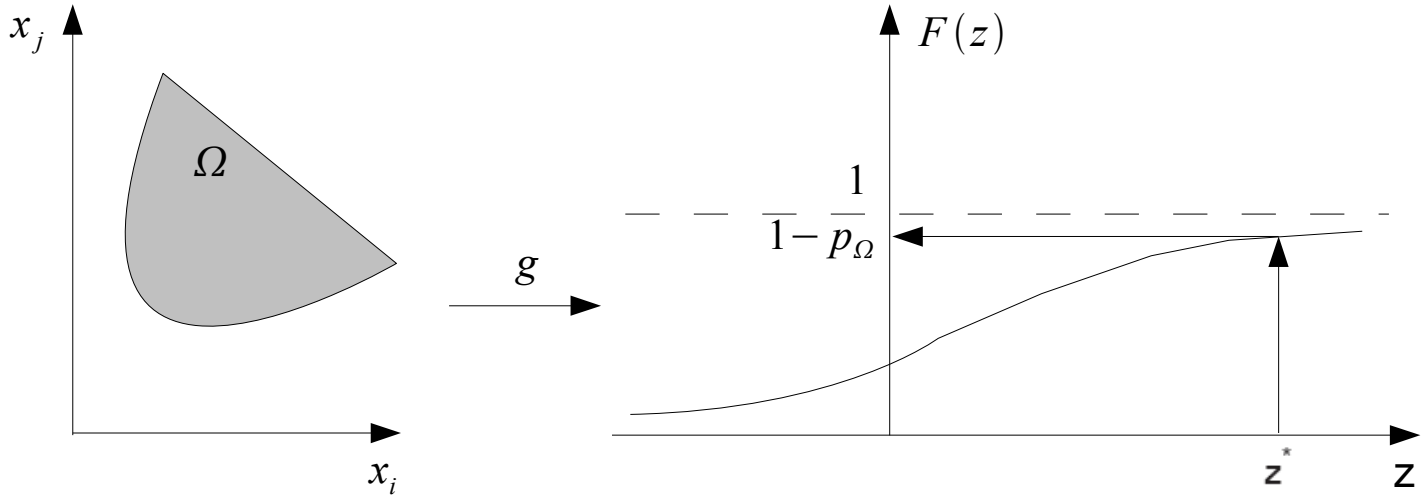
Limit-state function: $g : \mathbb{R}^N \to \mathbb{R}_+$, $z = g(x)$

Basic Idea

Reliability Problem: to compute the failure probability $p_\Omega = \int_\Omega \pi_0(x)dx$

Failure domain: $\Omega = \{x \in \mathbb{R}^N |\ g(x) > z^*\}$

Limit-state function: $g : \mathbb{R}^N \to \mathbb{R}_+,\ z = g(x)$

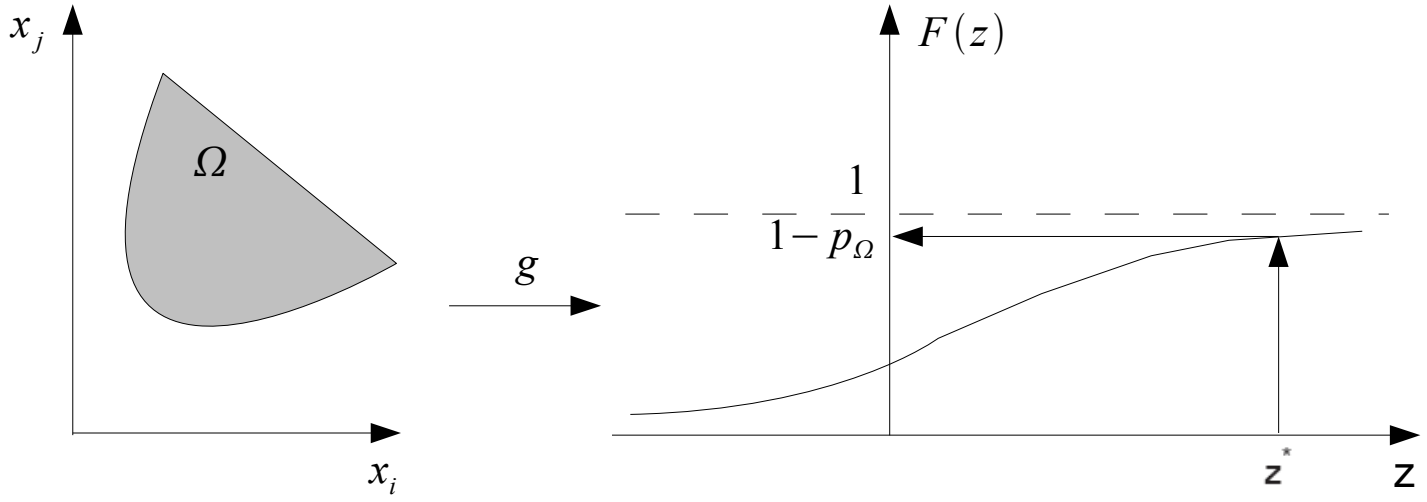


Basic Idea

Reliability Problem: to compute the failure probability $p_\Omega = \int_\Omega \pi_0(x)dx$

Failure domain: $\Omega = \{x \in \mathbb{R}^N | \, g(x) > z^*\}$

Limit-state function: $g : \mathbb{R}^N \to \mathbb{R}_+$, $z = g(x)$



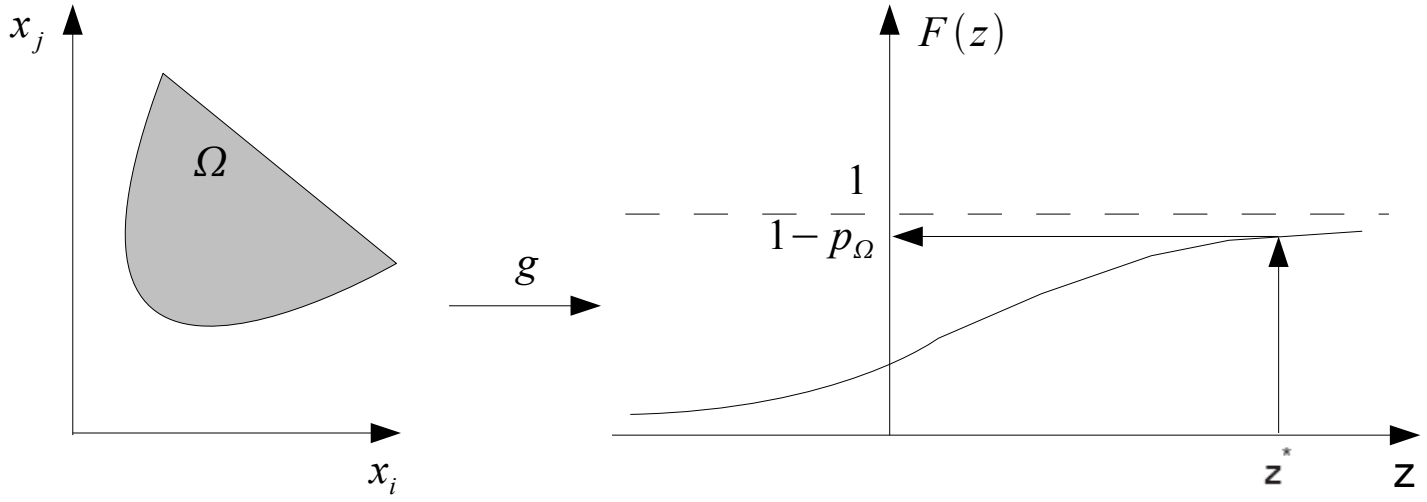
$$p_\Omega = 1 - F(z^*)$$

Basic Idea

Reliability Problem: to compute the failure probability $p_\Omega = \int_\Omega \pi_0(x)dx$

Failure domain: $\Omega = \{x \in \mathbb{R}^N |\ g(x) > z^*\}$

Limit-state function: $g : \mathbb{R}^N \to \mathbb{R}_+,\ z = g(x)$

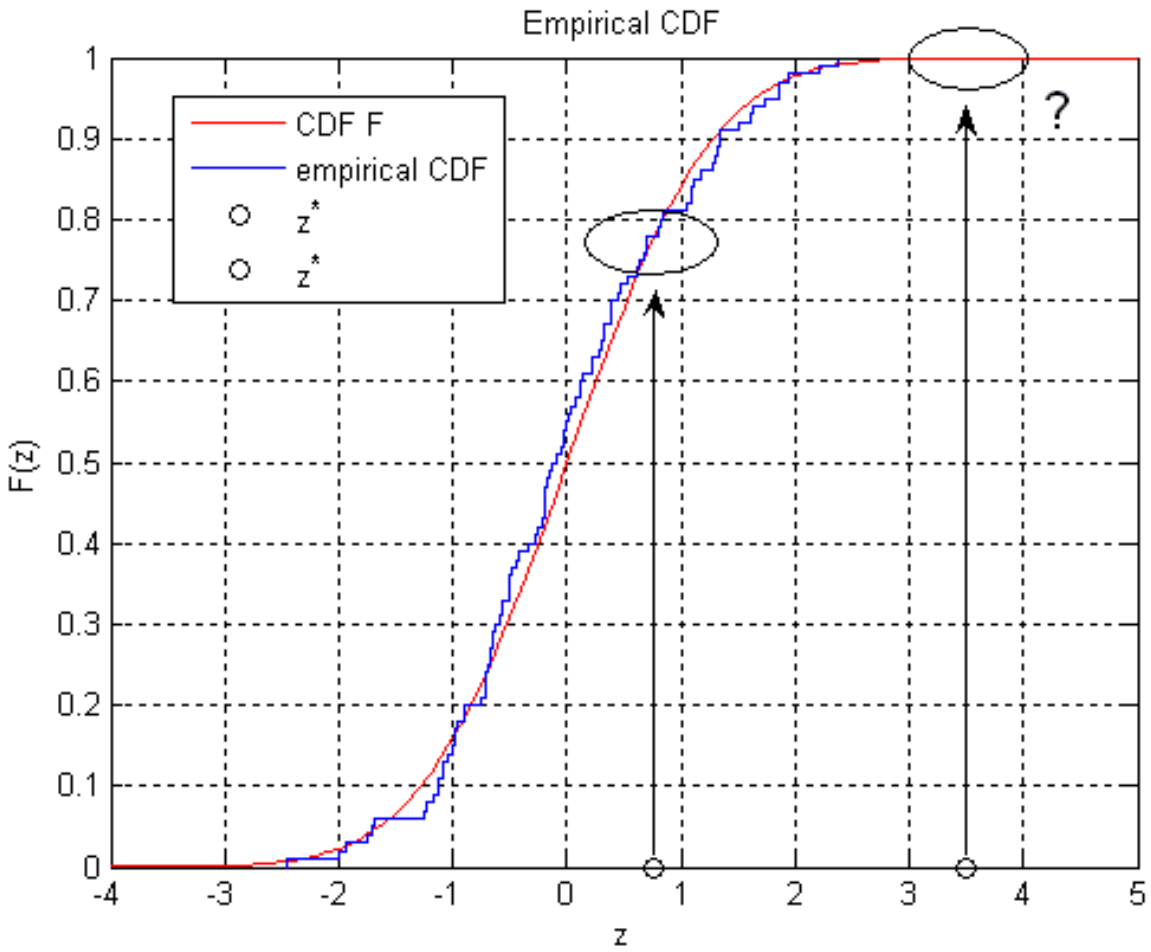


$$p_\Omega = 1 - F(z^*)$$

Basic Idea: to approximate the CDF F in the neighbourhood of z^*

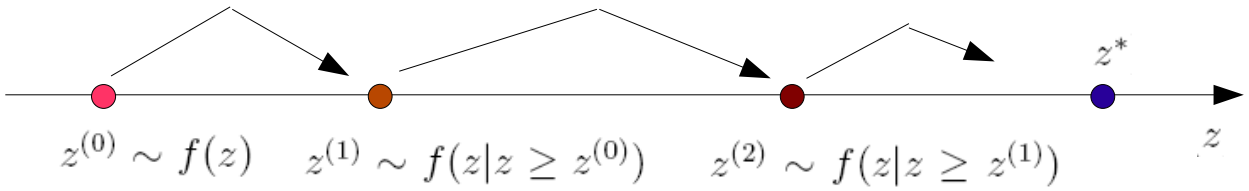
Basic Idea

If z^* is not very far from the median $\tilde{z}$, then we can use Monte Carlo Simulation



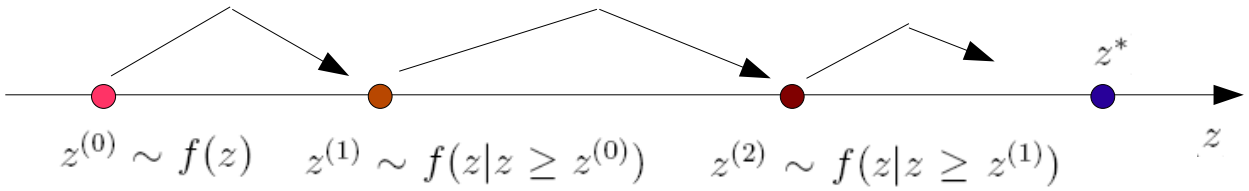
Basic Idea

Assume, we can propagate Monte Carlo samples towards the neighbourhood of z^*:



Basic Idea

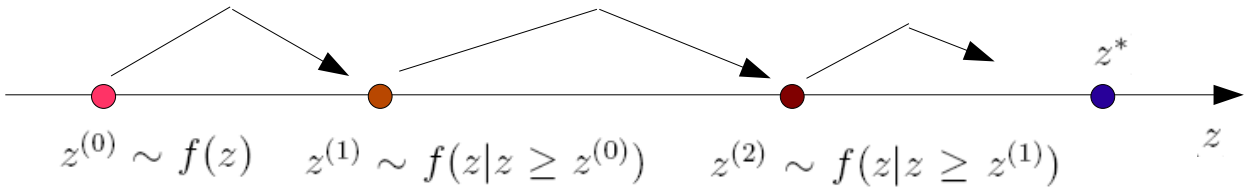
Assume, we can propagate Monte Carlo samples towards the neighbourhood of z^*:



PDF of $z^{(k)}$: $f_k(z) = \frac{(-1)^k}{k!} f(z) \left[\log(1 - F(z))\right]^k$

Basic Idea

Assume, we can propagate Monte Carlo samples towards the neighbourhood of z^*:



PDF of $z^{(k)}$: $f_k(z) = \frac{(-1)^k}{k!} f(z) \left[\log(1 - F(z))\right]^k$

Importance Sampling theory:

if $z_1^{(k)}, \ldots, z_n^{(k)} \sim f_k$, then $(z_1^{(k)}, w_1^{(k)}), \ldots, (z_n^{(k)}, w_n^{(k)}) \sim f$, where

$$w_i^{(k)} \propto \frac{f(z_i^{(k)})}{f_k(z_i^{(k)})} \propto \frac{1}{\left[\log(1-F(z_i^{(k)}))\right]^k}$$

Basic Idea

HORSERACING SIMULATION SCHEME

I. Sample $z_1^{(0)}, \ldots, z_n^{(0)}$ from $f_0 = f$,
Set $k = 0$.

II. Construct the empirical CDF $F^{(k)}$ based on $\{z_i^{(k)}\}_{i=1}^n$.

While the stopping criterion $C(z^*)$ is not fulfilled do:

III. Sample $z_i^{(k+1)}$ from $f_0(z|z \geq z_i^{(k)})$ for each $i = 1 \ldots n$.

IV. Construct the empirical CDF $G^{(k+1)}$ based on $\{(z_i^{(k+1)}, w_i^{(k+1)})\}_{i=1}^n$.

V. Update CDF $F^{(k)}$ to $F^{(k+1)}$, $(F^{(k)}, G^{(k+1)}) \rightsquigarrow F^{(k+1)}$,
Set k=k+1.

Basic Idea

HORSERACING SIMULATION SCHEME

I. Sample $z_1^{(0)}, \ldots, z_n^{(0)}$ from $f_0 = f$,
Set $k = 0$.

II. Construct the empirical CDF $F^{(k)}$ based on $\{z_i^{(k)}\}_{i=1}^n$.

While the stopping criterion $C(z^*)$ is not fulfilled do:

III. Sample $z_i^{(k+1)}$ from $f_0(z|z \geq z_i^{(k)})$ for each $i = 1 \ldots n$.

IV. Construct the empirical CDF $G^{(k+1)}$ based on $\{(z_i^{(k+1)}, w_i^{(k+1)})\}_{i=1}^n$.

V. Update CDF $F^{(k)}$ to $F^{(k+1)}$, $(F^{(k)}, G^{(k+1)}) \rightsquigarrow F^{(k+1)}$,
Set k=k+1.

In some neighbourhood of z^* : $F^{(k)} \approx F$

Implementation issues: Sampling

Step I. Sample from f_0 is easy:

Implementation issues: Sampling

Step I. Sample from f_0 is easy:

$$\begin{array}{c} x_1^{(0)}, \ldots, x_n^{(0)} \sim \pi_0 \\ \text{Monte Carlo samples} \end{array} \quad \Rightarrow \quad \begin{array}{l} z_1^{(0)} = g(x_1^{(0)}) \\ \ldots \\ z_n^{(0)} = g(x_n^{(0)}) \end{array} \sim f_0$$

Implementation issues: Sampling

Step I. Sample from f_0 is easy:

$$\begin{array}{c} x_1^{(0)}, \ldots, x_n^{(0)} \sim \pi_0 \\ \text{Monte Carlo samples} \end{array} \quad \Rightarrow \quad \begin{array}{l} z_1^{(0)} = g(x_1^{(0)}) \\ \ldots \\ z_n^{(0)} = g(x_n^{(0)}) \end{array} \sim f_0$$

Step III. Given $z_i^{(k)}$, sample $z_i^{(k+1)}$ from $f_0(z|z \geq z_i^{(k)})$

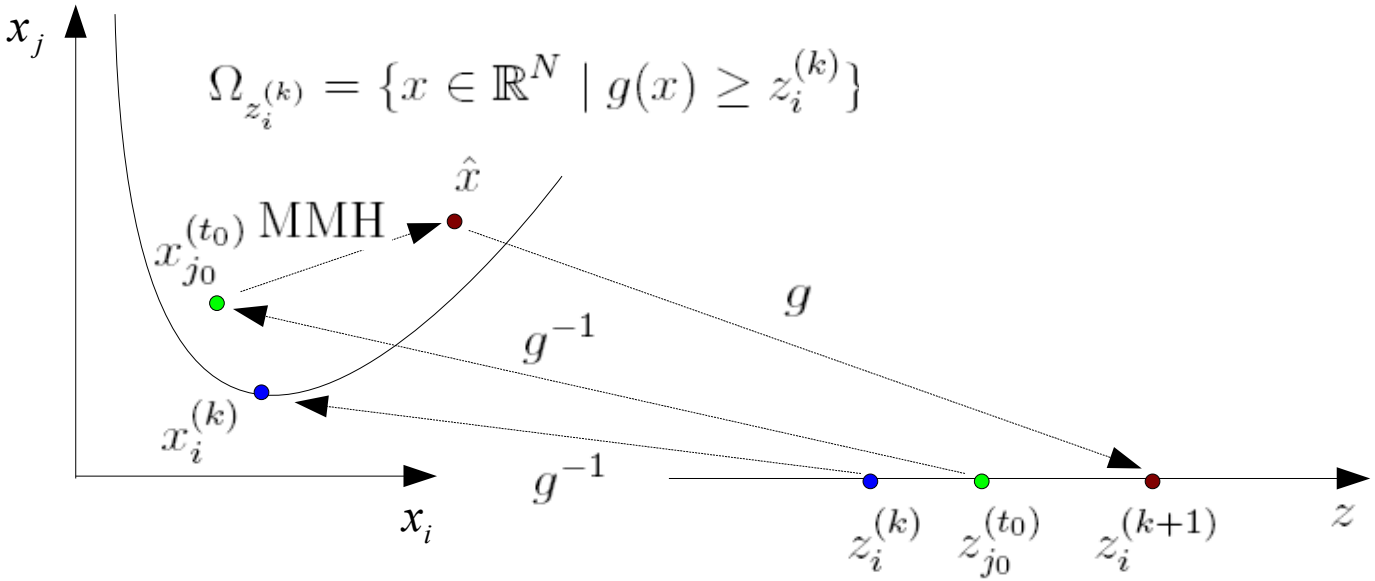
Implementation issues: Sampling

Step I. Sample from f_0 is easy:

$$x_1^{(0)}, \ldots, x_n^{(0)} \sim \pi_0 \quad \Rightarrow \quad \begin{array}{l} z_1^{(0)} = g(x_1^{(0)}) \\ \ldots \\ z_n^{(0)} = g(x_n^{(0)}) \end{array} \sim f_0$$

Monte Carlo samples

Step III. Given $z_i^{(k)}$, sample $z_i^{(k+1)}$ from $f_0(z|z \geq z_i^{(k)})$



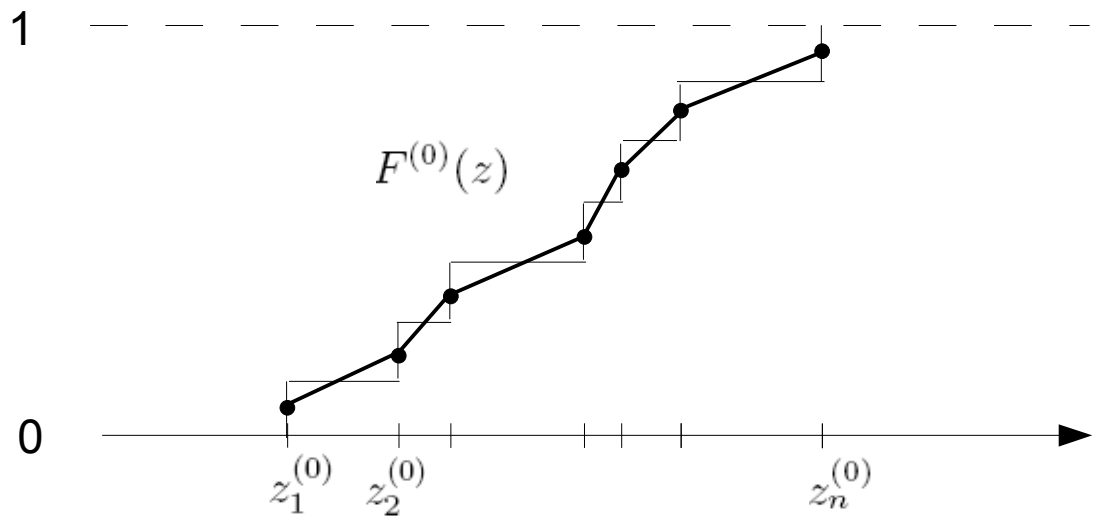
Implementation issues: Construction of the empirical CDF

Implementation issues: Construction of the empirical CDF

Step II. Construct a zeroth approximation $F^{(0)}$ based on $z_1^{(0)}, \ldots, z_n^{(0)}$

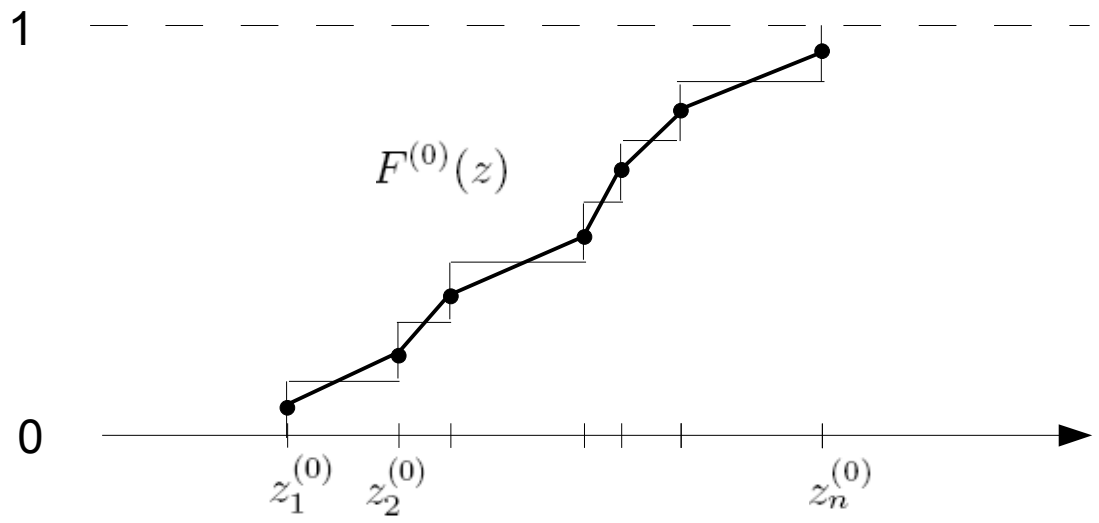
Implementation issues: Construction of the empirical CDF

Step II. Construct a zeroth approximation $F^{(0)}$ based on $z_1^{(0)}, \ldots, z_n^{(0)}$



Implementation issues: Construction of the empirical CDF

Step II. Construct a zeroth approximation $F^{(0)}$ based on $z_1^{(0)}, \ldots, z_n^{(0)}$



$$F^{(0)}(z) = \frac{z}{n\left(z_{i+1}^{(0)} - z_i^{(0)}\right)} + \frac{(2i-1)z_{i+1}^{(0)} - (2i+1)z_i^{(0)}}{2n\left(z_{i+1}^{(0)} - z_i^{(0)}\right)}, \quad \text{for } z \in [z_i^{(0)}, z_{i+1}^{(0)}]$$

Implementation issues: Construction of the empirical CDF

Step IV. Construct the empirical CDF $G^{(k+1)}$ based on the weighted samples $\{(z_i^{(k+1)}, w_i^{(k+1)})\}_{i=1}^n$

For $z \in [z_i^{(k+1)}, z_{i+1}^{(k+1)}]$:

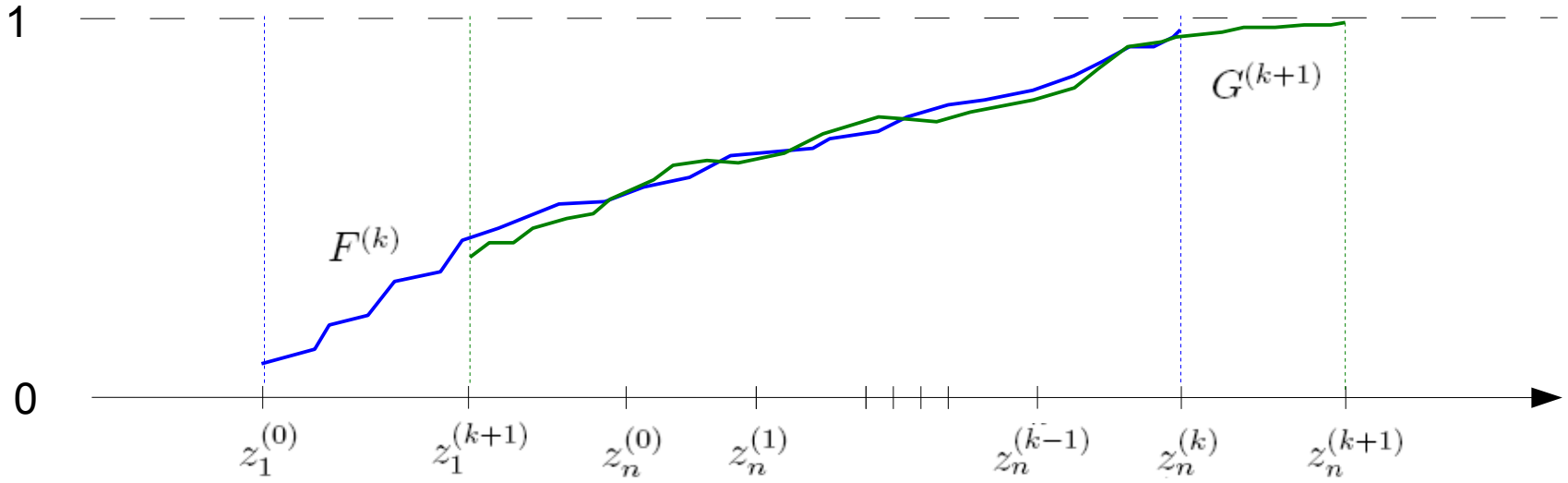
$$G^{(k+1)}(z) = \frac{w_i^{(k+1)} + w_{i+1}^{(k+1)}}{2\left(z_{i+1}^{(k+1)} - z_i^{(k+1)}\right)}\, z + \frac{\left(2\sum_{j=1}^{i} w_j^{(k+1)} - w_i^{(k+1)}\right) z_{i+1}^{(k+1)} - \left(2\sum_{j=1}^{i} w_j^{(k+1)} + w_{i+1}^{(k+1)}\right) z_i^{(k+1)}}{2\left(z_{i+1}^{(k+1)} - z_i^{(k+1)}\right)}$$

Implementation issues: Updating of the empirical CDF

Step V. Update the CDF $F^{(k)}$, using new information provided by $G^{(k+1)}$, and construct a new approximation $F^{(k+1)}$ of the CDF F.

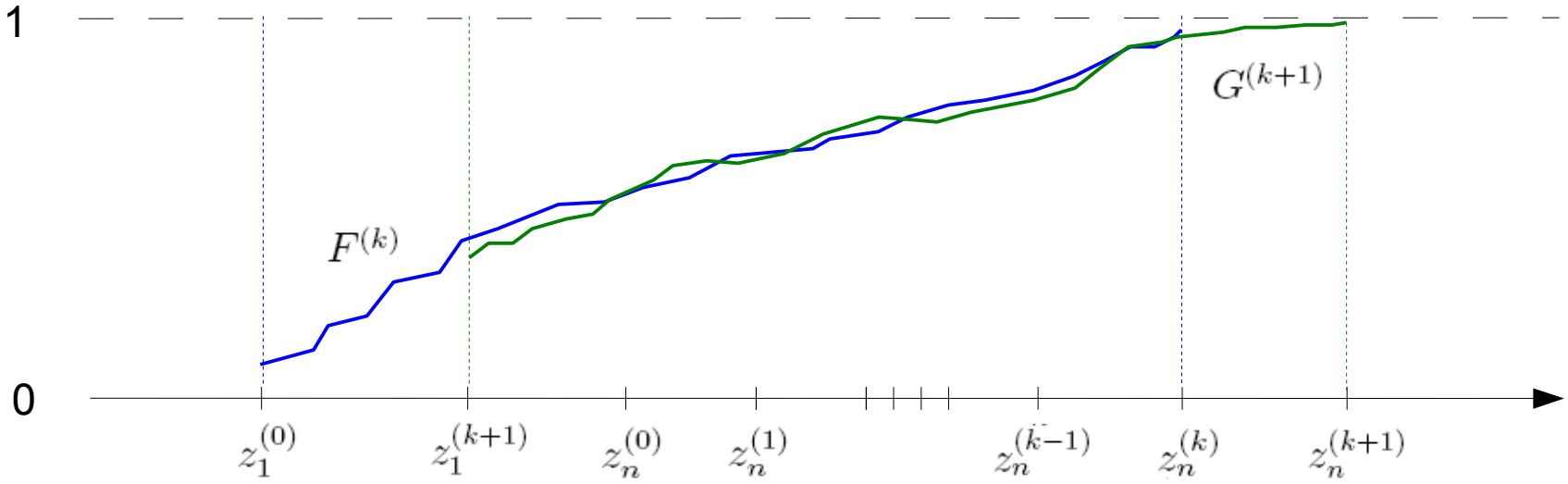
Implementation issues: Updating of the empirical CDF

Step V. Update the CDF $F^{(k)}$, using new information provided by $G^{(k+1)}$, and construct a new approximation $F^{(k+1)}$ of the CDF F.



Implementation issues: Updating of the empirical CDF

Step V. Update the CDF $F^{(k)}$, using new information provided by $G^{(k+1)}$, and construct a new approximation $F^{(k+1)}$ of the CDF F.



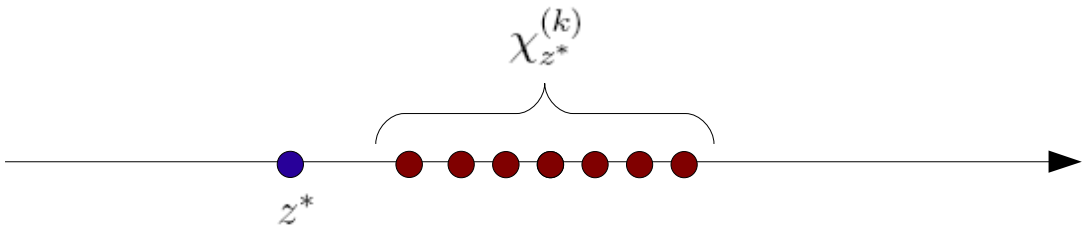
$$F^{(k+1)}(z) = \begin{cases} F^{(k)}(z), & \text{for } z \in [z_1^{(0)}, z_1^{(k+1)}); \\ \frac{(k+1)F^{(k)}(z)+G^{(k+1)}(z)}{k+2}, & \text{for } z \in [z_1^{(k+1)}, z_n^{(0)}]; \\ \frac{kF^{(k)}(z)+G^{(k+1)}(z)}{k+1}, & \text{for } z \in (z_n^{(0)}, z_n^{(1)}]; \\ \dots, & \dots \\ \frac{F^{(k)}(z)+G^{(k+1)}(z)}{2}, & \text{for } z \in (z_n^{(k-1)}, z_n^{(k)}]; \\ G^{(k+1)}(z), & \text{for } z \in (z_n^{(k)}, z_n^{(k+1)}]. \end{cases}$$

Implementation issues: Stopping Criterion

Implementation issues: Stopping Criterion

$C(z^*)$: the race is over when $r\%$ of horses reach z^*,

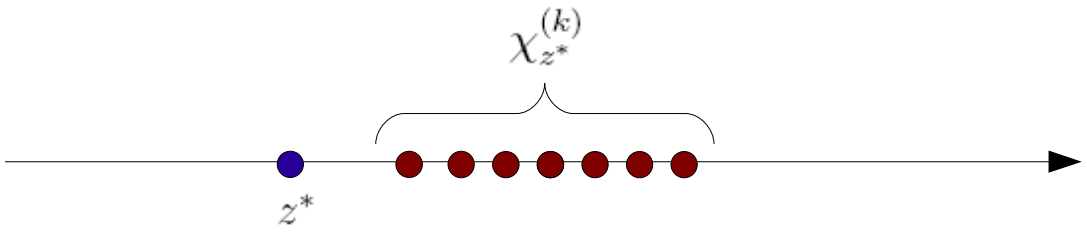
$$\chi_{z^*}^{(k)} \geq r\%$$



Implementation issues: Stopping Criterion

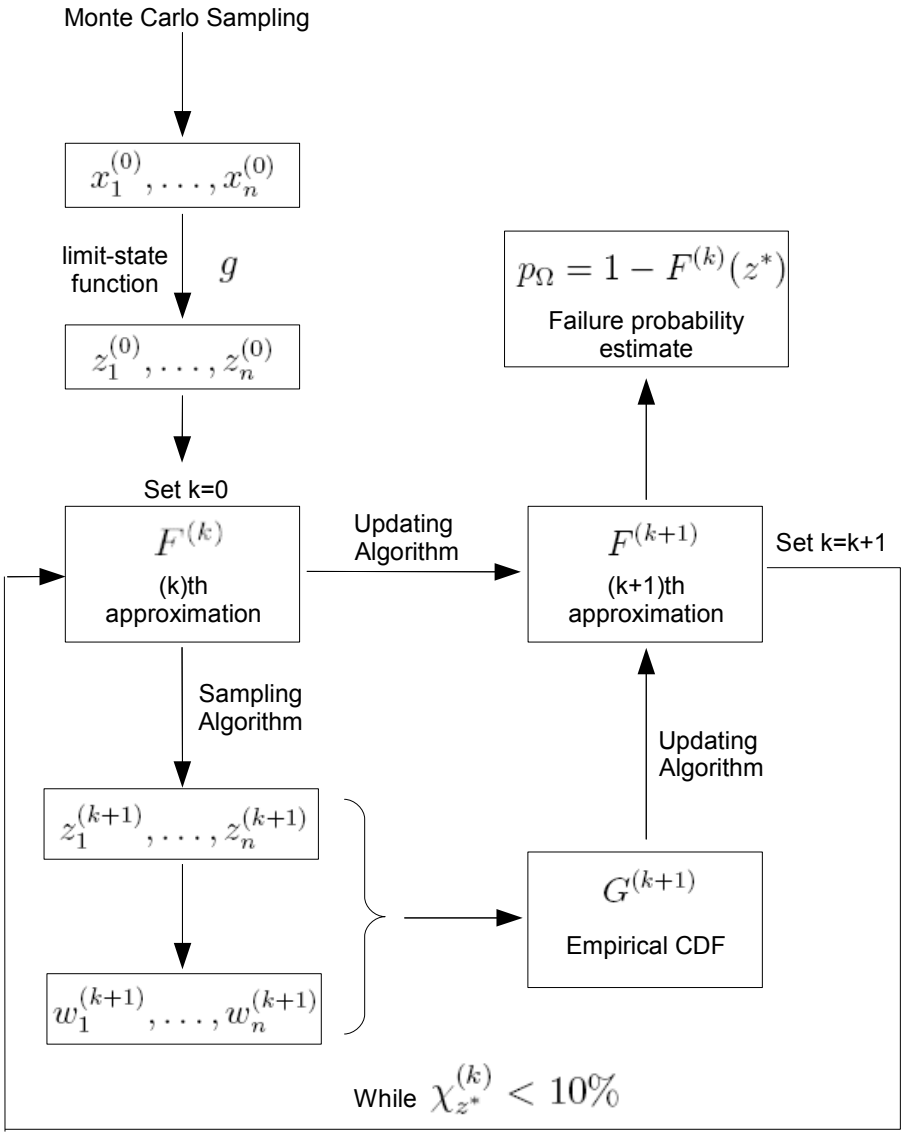
$C(z^*)$: the race is over when $r\%$ of horses reach z^*,

$$\chi_{z^*}^{(k)} \geq r\%$$



- For a target probability level of 10^{-2} to 10^{-5}, choosing $r = 10\%$ is found to yield good efficiency

Horseracing Simulation Algorithm



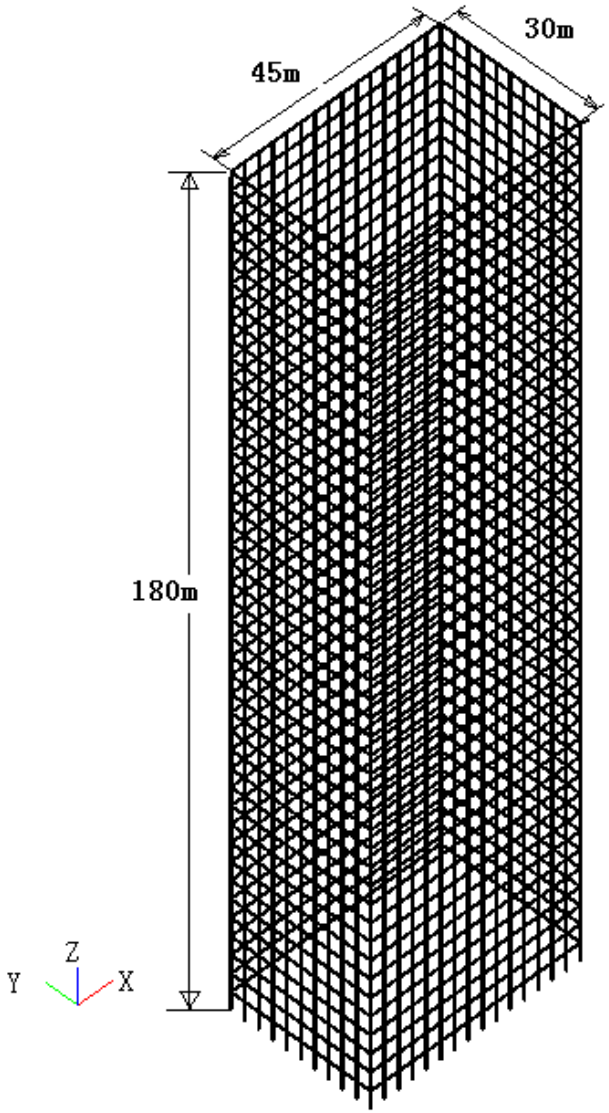
Example

CAARC standard tall building model

Table 1: Design of colum members and beam members

Floor zone	Column members	Beam members
1 ~9F	W14X550	W30X357
10~18F	W14X500	W30X326
19~27F	W14X370	W30X292
28~36F	W14X257	W30X261
37~45F	W14X159	W30X221

- $180\text{m} \times 30\text{m} \times 45\text{m}$
- 45-story, 10-bay by 15-bay
- Story height 4 m
- Bay width 3 m
- Floor
  - mass 6.75×10^5 kg
  - rotational moment of inertia 1.645×10^8 kg.m^2



Example

Wind excitation:

Example

Wind excitation: The along-wind excitation in the Y-direction of the building

Example

Wind excitation: The along-wind excitation in the Y-direction of the building

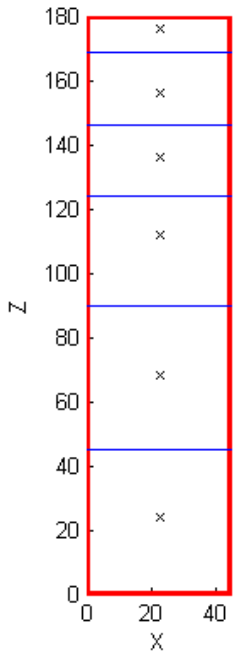


Table 2: Acting heights and acting areas of 6 excitation forces in the discretization scheme

Excitation	Acting height (m)	Acting area (m^2)
$U_1(t)$	24	45×45
$U_2(t)$	68	45×45
$U_3(t)$	112	33.75×45
$U_4(t)$	136	22.5×45
$U_5(t)$	156	22.5×45
$U_6(t)$	176	11.25×45

Example

Wind excitation: The along-wind excitation in the Y-direction of the building

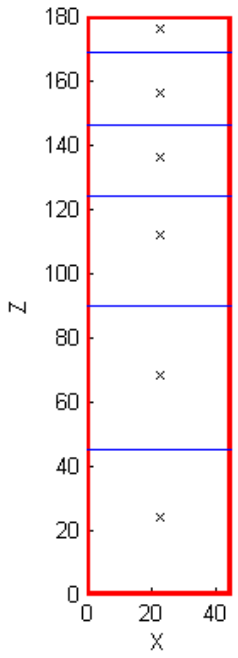


Table 2: Acting heights and acting areas of 6 excitation forces in the discretization scheme

Excitation	Acting height (m)	Acting area (m^2)
$U_1(t)$	24	45×45
$U_2(t)$	68	45×45
$U_3(t)$	112	33.75×45
$U_4(t)$	136	22.5×45
$U_5(t)$	156	22.5×45
$U_6(t)$	176	11.25×45

Wind excitation forces:

$$U_j(t) = \frac{1}{2}\rho A_j(\bar{V}_j + v_j(t))^2$$

- Air density $\rho = 1.2$ kg/m^3
- Acting aria A_j
- Mean wind speed
  $\bar{V}_j = 41\left(\frac{h_j}{180}\right)^{0.25}$
- Fluctuating components
  $v_j(t) \leftarrow$ SR method

Example

Structure response

Example

Structure response

Displacement response $Y(t)$ at the top floor of the building is of interest

$$Y(t) = \sum_{j=1}^{6} \int_{0}^{t} q_j(t,\tau) U_j(\tau)\, d\tau$$

$q_j(t,\tau)$ is the response function for $Y(t)$

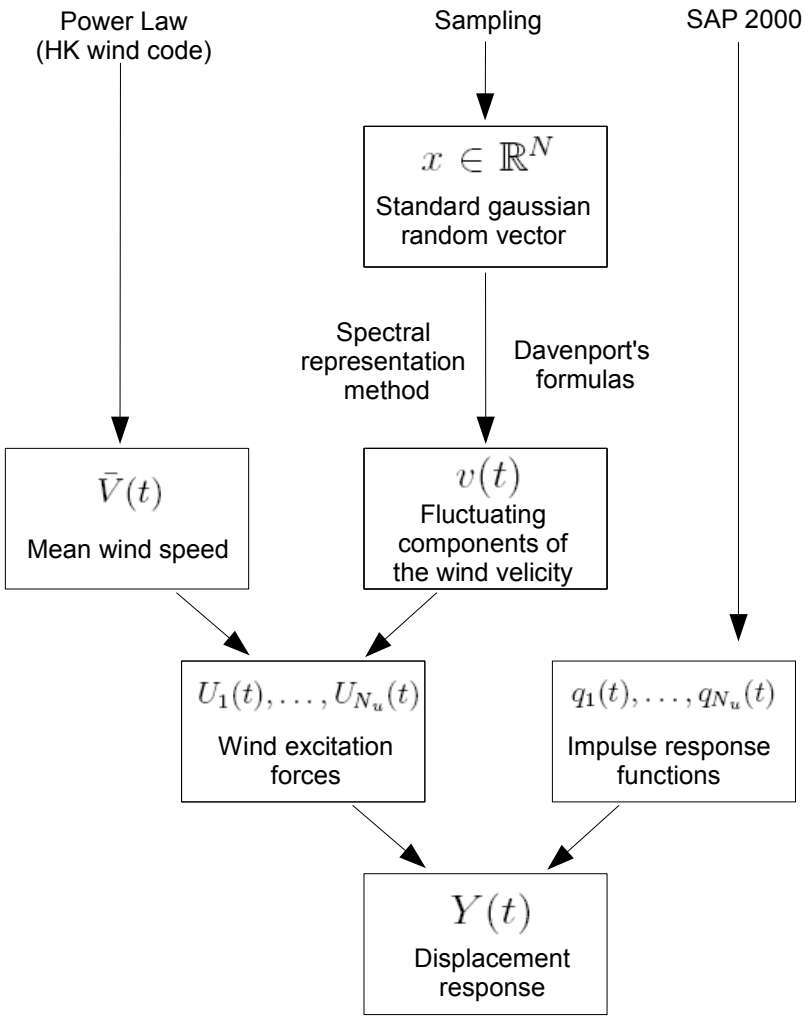
Example

Structure response

Displacement response $Y(t)$ at the top floor of the building is of interest

$$Y(t) = \sum_{j=1}^{6} \int_0^t q_j(t,\tau) U_j(\tau)\, d\tau$$

$q_j(t,\tau)$ is the response function for $Y(t)$



Example

Failure domain

Example

Failure domain

Ω: $|Y(t)|$ exceeds a specified threshold z^* within one hour.

Example

Failure domain

Ω: $|Y(t)|$ exceeds a specified threshold z^* within one hour.

- Discrete time formulation
  - Duration time $T = 3600$ s
  - Sampling time interval $\Delta t = 0.01$ s
  - Number of time instants $N_t = T/\triangle t = 3.6 \times 10^5$

Example

Failure domain

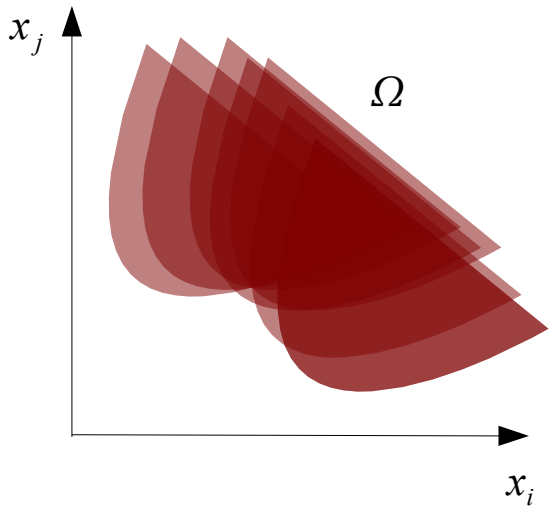
Ω: $|Y(t)|$ exceeds a specified threshold z^* within one hour.

- Discrete time formulation
  - Duration time $T = 3600$ s
  - Sampling time interval $\Delta t = 0.01$ s
  - Number of time instants $N_t = T/\triangle t = 3.6 \times 10^5$

$$\Omega = \bigcup_{i=1}^{N_t} \left\{ x \in \mathbb{R}^N \ : \ |Y(i)| > z^* \right\}$$

$$N = 2 \times N_u \times N_\omega = 8640$$

$$g(x) = \max\{|Y(i)|, \ i = 1, \ldots, N_t\}$$



Example

Simulation results

Example

Simulation results

1. $z_1^* = 1.25$ m $\Rightarrow \quad p_{\Omega_1} = 3.3 \times 10^{-2}, \quad \delta_1 = 5.4\%, \quad n_{MC} = 10^4$
2. $z_2^* = 1.35$ m $\Rightarrow \quad p_{\Omega_2} = 6.8 \times 10^{-3}, \quad \delta_2 = 12.1\%, \quad n_{MC} = 10^4$
3. $z_3^* = 1.45$ m $\Rightarrow \quad p_{\Omega_3} = 1.5 \times 10^{-3}, \quad \delta_3 = 25.8\%, \quad n_{MC} = 10^4$

Example

Simulation results

1. $z_1^* = 1.25$ m $\Rightarrow \quad p_{\Omega_1} = 3.3 \times 10^{-2}, \quad \delta_1 = 5.4\%, \quad n_{MC} = 10^4$
2. $z_2^* = 1.35$ m $\Rightarrow \quad p_{\Omega_2} = 6.8 \times 10^{-3}, \quad \delta_2 = 12.1\%, \quad n_{MC} = 10^4$
3. $z_3^* = 1.45$ m $\Rightarrow \quad p_{\Omega_3} = 1.5 \times 10^{-3}, \quad \delta_3 = 25.8\%, \quad n_{MC} = 10^4$

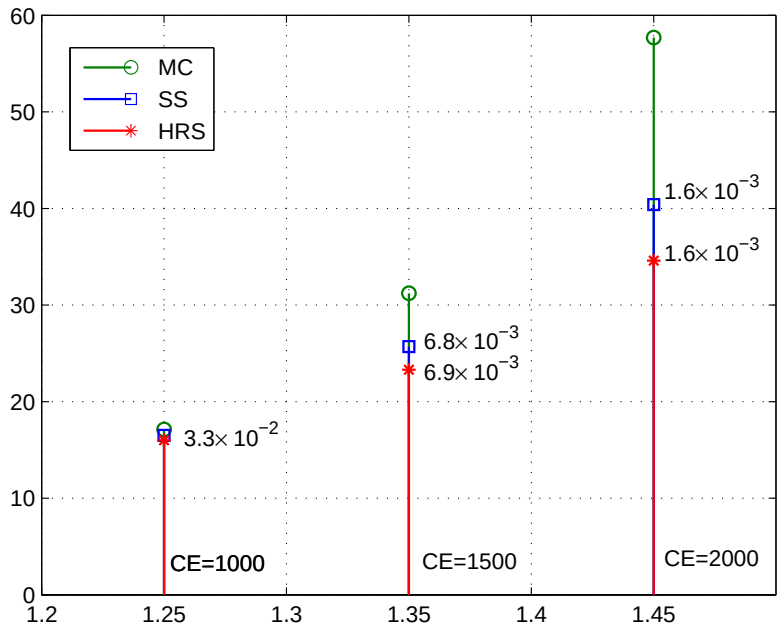
Horseracing Simulation algorithm is applied with $n = 500$ initial samples

Example

Simulation results

1. $z_1^* = 1.25$ m $\Rightarrow \quad p_{\Omega_1} = 3.3 \times 10^{-2}, \quad \delta_1 = 5.4\%, \quad n_{MC} = 10^4$
2. $z_2^* = 1.35$ m $\Rightarrow \quad p_{\Omega_2} = 6.8 \times 10^{-3}, \quad \delta_2 = 12.1\%, \quad n_{MC} = 10^4$
3. $z_3^* = 1.45$ m $\Rightarrow \quad p_{\Omega_3} = 1.5 \times 10^{-3}, \quad \delta_3 = 25.8\%, \quad n_{MC} = 10^4$

Horseracing Simulation algorithm is applied with $n = 500$ initial samples

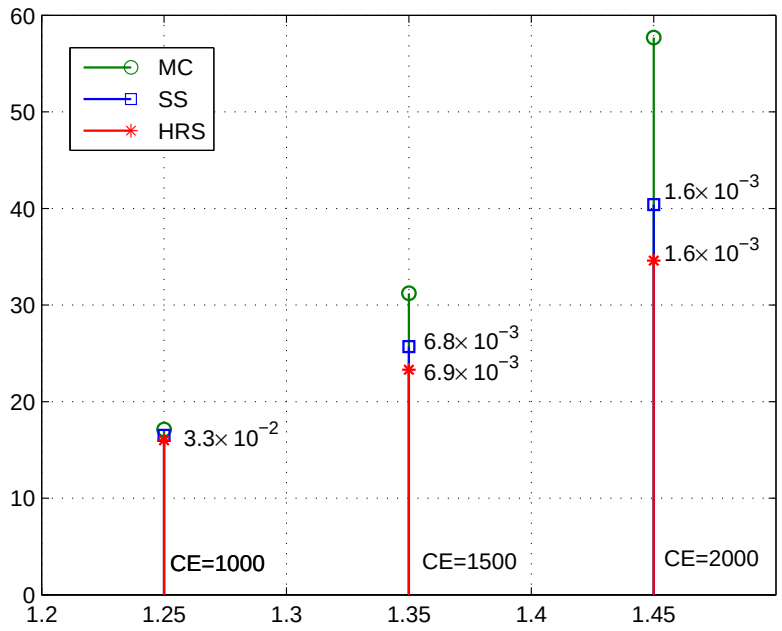


Example

Simulation results

1. $z_1^* = 1.25$ m $\Rightarrow \quad p_{\Omega_1} = 3.3 \times 10^{-2}, \quad \delta_1 = 5.4\%, \quad n_{MC} = 10^4$
2. $z_2^* = 1.35$ m $\Rightarrow \quad p_{\Omega_2} = 6.8 \times 10^{-3}, \quad \delta_2 = 12.1\%, \quad n_{MC} = 10^4$
3. $z_3^* = 1.45$ m $\Rightarrow \quad p_{\Omega_3} = 1.5 \times 10^{-3}, \quad \delta_3 = 25.8\%, \quad n_{MC} = 10^4$

Horseracing Simulation algorithm is applied with $n = 500$ initial samples



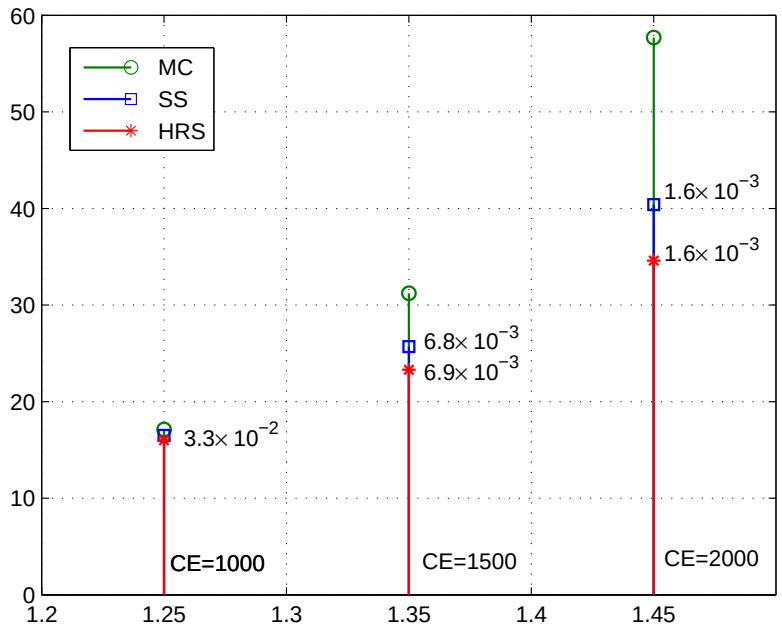
1. HRS≈SS≈MC

Example

Simulation results

1. $z_1^* = 1.25$ m $\Rightarrow \quad p_{\Omega_1} = 3.3 \times 10^{-2}, \quad \delta_1 = 5.4\%, \quad n_{MC} = 10^4$
2. $z_2^* = 1.35$ m $\Rightarrow \quad p_{\Omega_2} = 6.8 \times 10^{-3}, \quad \delta_2 = 12.1\%, \quad n_{MC} = 10^4$
3. $z_3^* = 1.45$ m $\Rightarrow \quad p_{\Omega_3} = 1.5 \times 10^{-3}, \quad \delta_3 = 25.8\%, \quad n_{MC} = 10^4$

Horseracing Simulation algorithm is applied with $n = 500$ initial samples



1. HRS≈SS≈MC
2. Reductions in CV
   $(\delta_2^{SS} - \delta_2^{HRS})/\delta_2^{SS} = 9.3\%$
3. Reductions in CV
   $(\delta_3^{Ss} - \delta_3^{HRS})/\delta_3^{SS} = 14.4\%$

Conclusion

- A new advanced stochastic simulation algorithm, called Horseracing Simulation (HRS), is proposed for solving high-dimensional reliability problems
- The key idea behind HRS is to approximate the CDF of the random variable of interest by empirical CDFs constructed from specially designed samples
- The accuracy and efficiency of the new method is demonstrated with real-life example

Thank You for attention!