Targeted Random Sampling for Reliability Assessment: A Demonstration of Concept

Michael D. Shields
Assistant Professor
Dept. of Civil Engineering
Johns Hopkins University

V.S. Sundar
Postdoctoral Fellow
Dept. of Civil Engineering
Johns Hopkins University

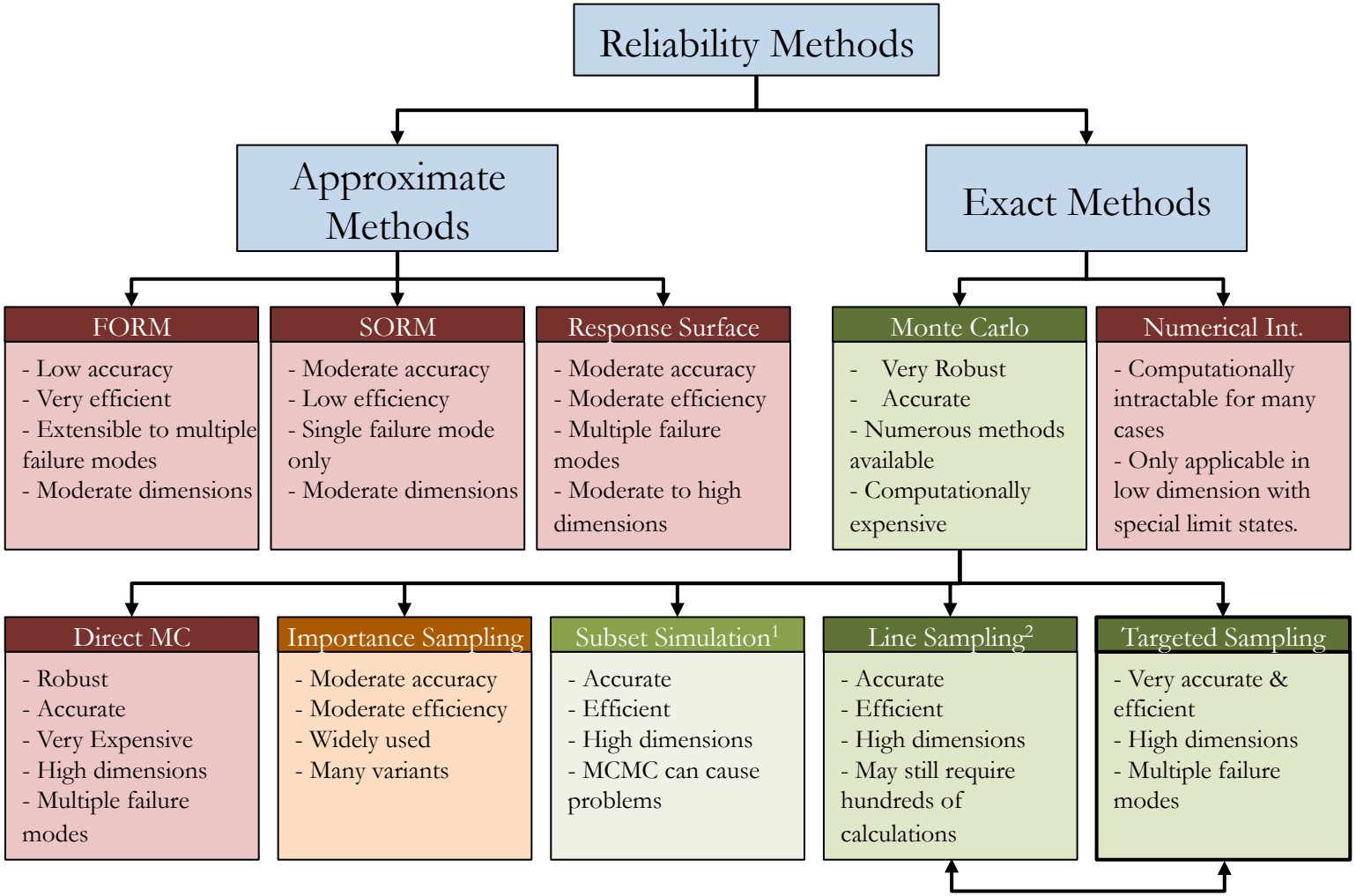
Motivation of Present Work

- Monte Carlo Simulation (MCS) is the most robust and accurate (and sometimes the *only*) means of reliability analysis
- MCS is often (rightfully so) criticized for its high computational cost
- The cost of MCS has been greatly reduced through the advent of sophisticated sampling and simulation methods
- We can still do better!

Outline of Presentation

- Overview of Reliability Methods
  - Emphasis on Monte Carlo Simulation based methods
- Refined Stratified Sampling
  - Concept / Theory
  - Algorithms / Applications
- Targeted Random Sampling
  - Concept / Theory
  - Algorithm
  - Examples
- Challenges & Future Directions

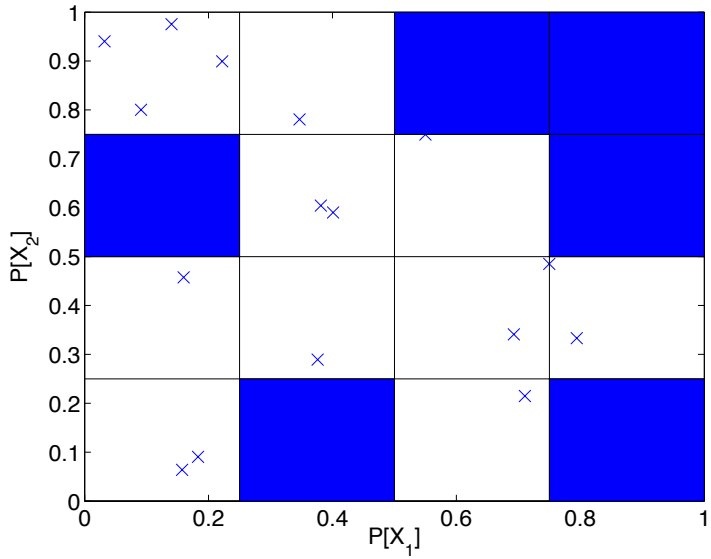
Reliability Analysis



1. Au, S.-K. and Beck, J.L. (2001). "Estimation of small failure probabilities in high dimension by subset simulation" *Probabilistic Engineering Mechanics*
2. Koutsourelakis, P.S., Pradlwarter, H.J., Schueller, G.I. (2004). "Reliability of structures in high dimension, part I: algorithms and applications" *Probabilistic Engineering Mechanics*

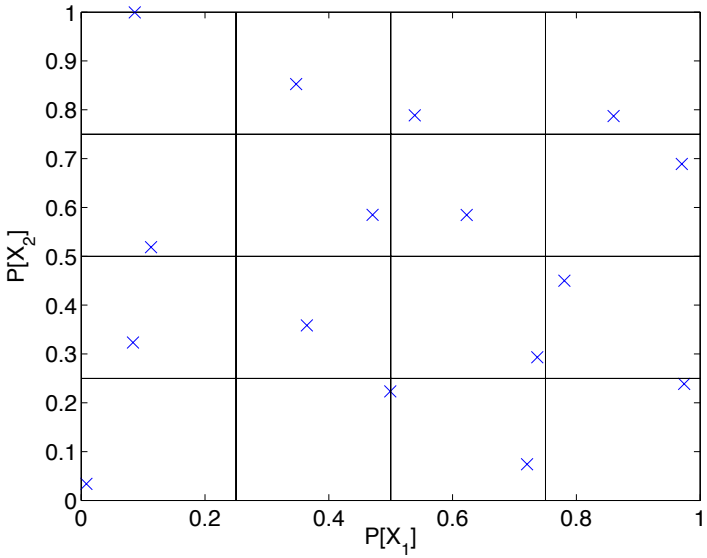
Stratified Sampling

Random Sampling



- Sample size extension is straightforward
- No control over where new samples are placed
- No need to compute sample weights

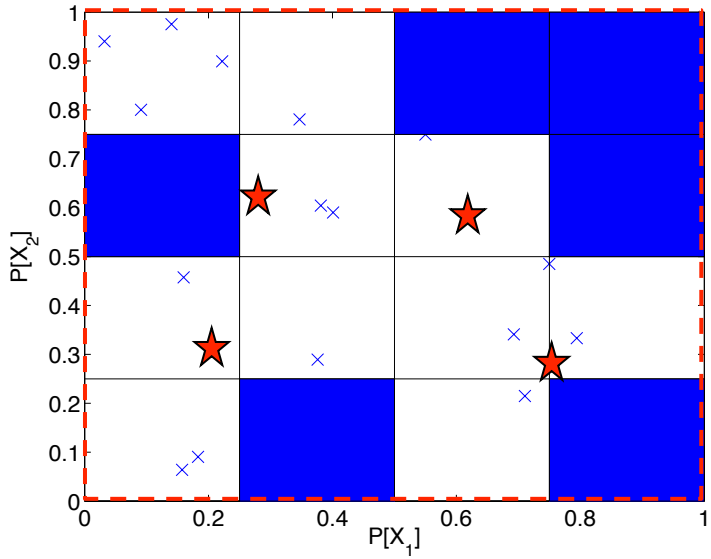
Stratified Sampling



- Almost limitless control over how samples are added
- Carefully select, based on the available information, which strata to add samples to
- Is there a better way?

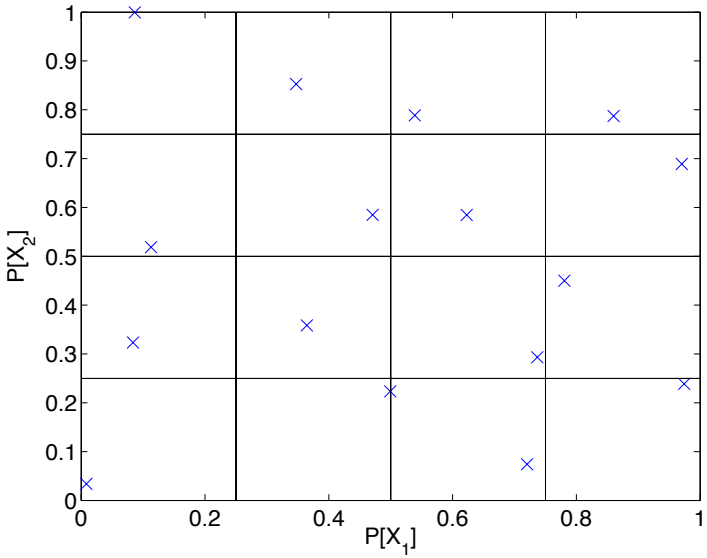
Stratified Sampling

Random Sampling



- Sample size extension is straightforward
- No control over where new samples are placed
- No need to compute sample weights

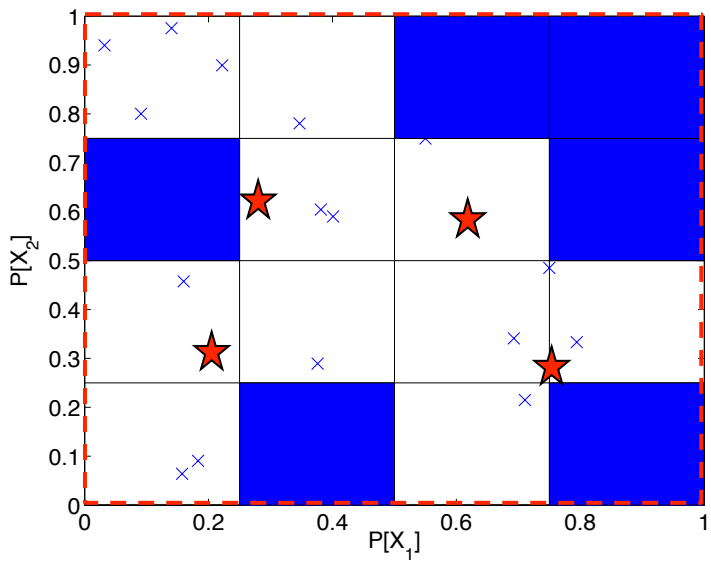
Stratified Sampling



- Almost limitless control over how samples are added
- Carefully select, based on the available information, which strata to add samples to
- Is there a better way?

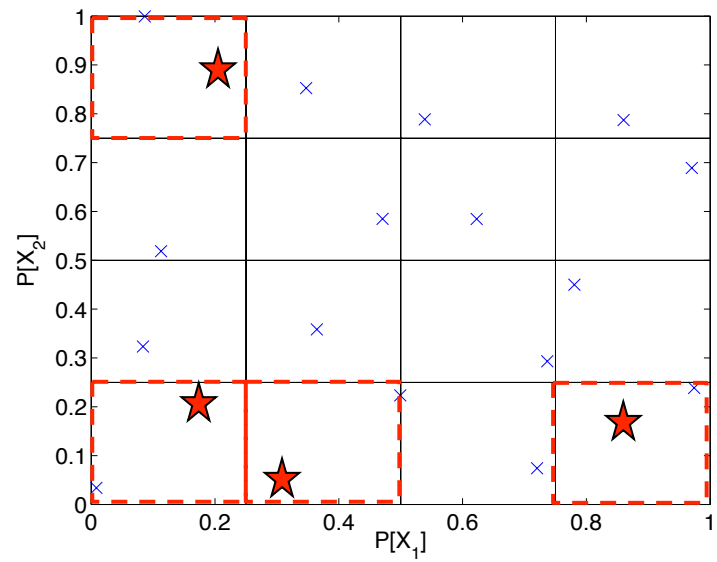
Stratified Sampling

Random Sampling



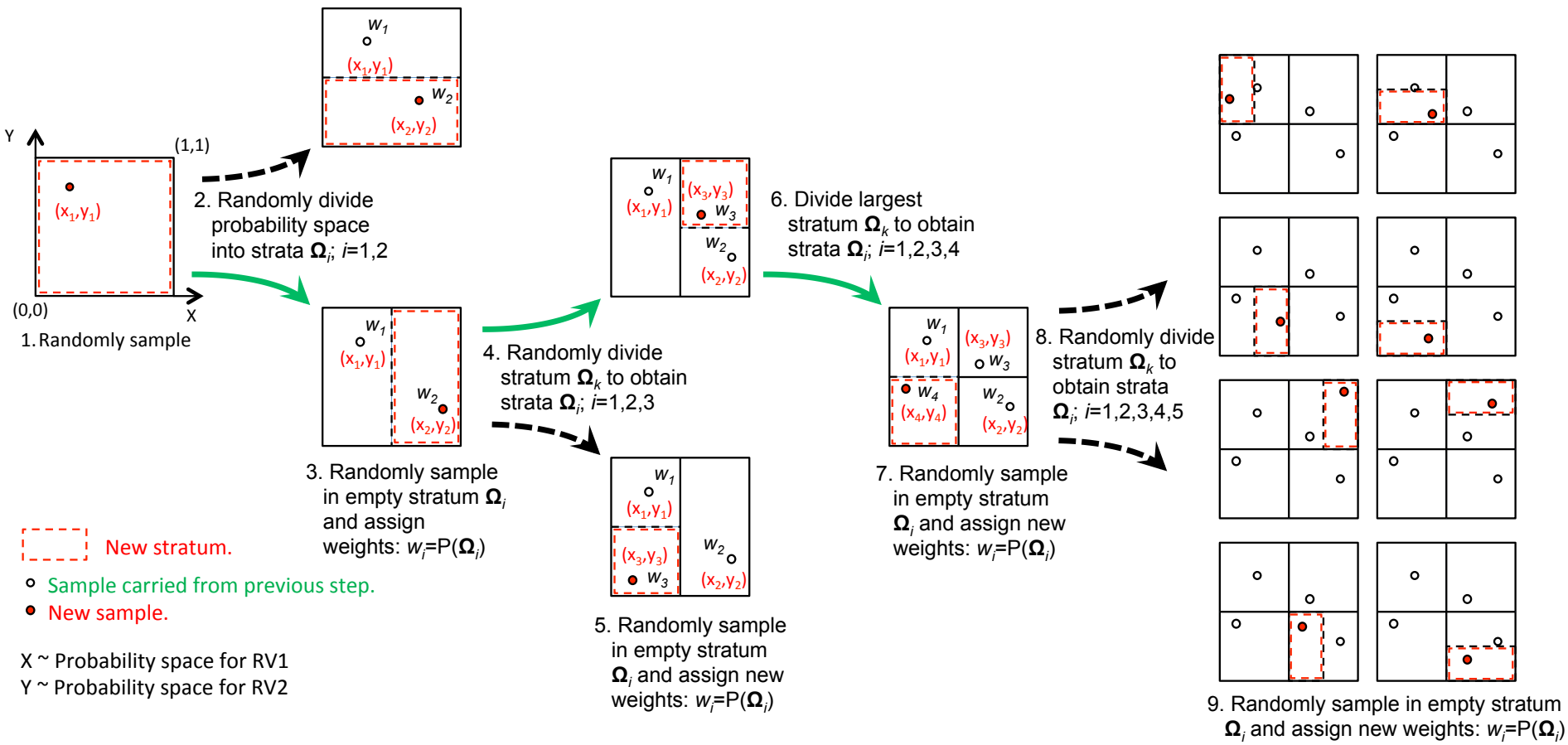
- Sample size extension is straightforward
- No control over where new samples are placed
- No need to compute sample weights

Stratified Sampling



- Almost limitless control over how samples are added
- Carefully select, based on the available information, which strata to add samples to
- Is there a better way?

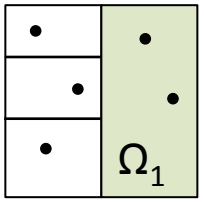
Refined Stratified Sampling



1. Shields, M.D., Teferra, K., Hapij, A. and Daddazio, R. (2014) "Refined stratified sampling for efficient Monte Carlo based uncertainty quantification" *Reliability Engineering & System Safety* (forthcoming)

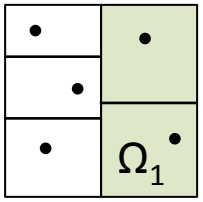
Why Refine Strata?

- Consider a N samples from stratum Ω_1:



$$Var\left[T_S^1\right] = \underbrace{\frac{p_1^2}{N}\sigma_1^2}_{\Omega_1} + \underbrace{\sum_{j=1}^{M}\frac{p_j^2}{n_j}\sigma_j^2}_{\text{Other strata}}$$

- Consider stratum refinement with a single sample in each stratum:



$$Var\left[T_S^2\right] = \underbrace{\frac{p_1^2}{N^2}\sum_{i=1}^{N}\sigma_i^2}_{\Omega_1} + \underbrace{\sum_{j=1}^{M}\frac{p_j^2}{n_j}\sigma_j^2}_{\text{Other strata}}$$

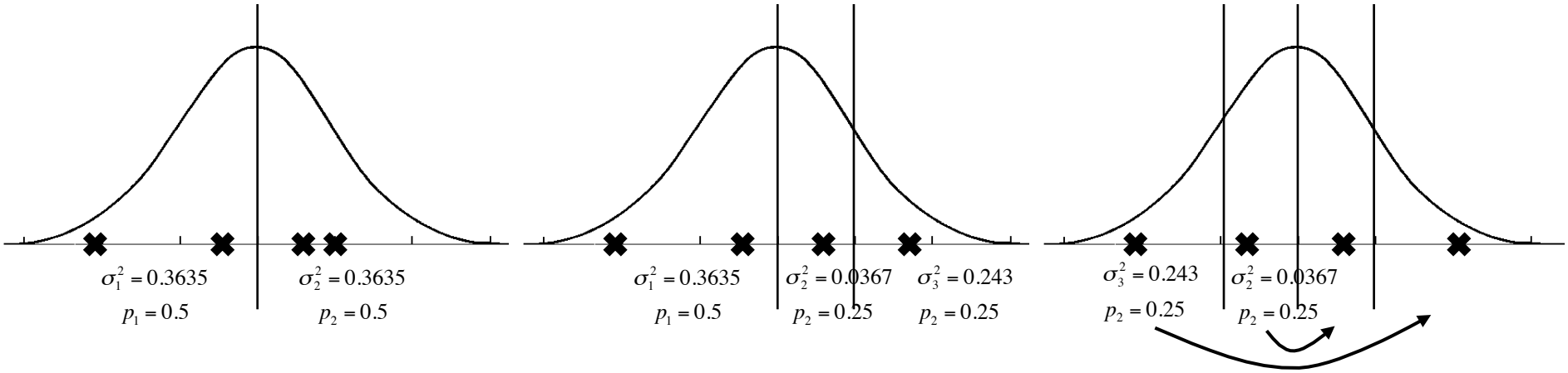
- Evaluate the difference in the variances:

$$Var\left[T_S^1\right] - Var\left[T_S^2\right] = \frac{p_1^2}{N}\left(\sigma_1^2 - \frac{1}{N}\sum_{i=1}^{N}\sigma_i^2\right) > 0$$

(Assumes a "balanced" sub-stratification)

Variance Savings by RSS: Example

- Consider samples drawn from x~N(0,1) with $T_S = E[x]$:



$$Var\left[T_S^b\right] = \frac{\left(\frac{1}{2}\right)^2}{2}\sigma_1^2 + \frac{\left(\frac{1}{2}\right)^2}{2}\sigma_2^2$$

$$Var\left[T_S^b\right] = 0.090875$$

$$Var\left[T_S^u\right] = \frac{\left(\frac{1}{2}\right)^2}{2}\sigma_1^2 + \left(\frac{1}{4}\right)^2\sigma_2^2 + \left(\frac{1}{4}\right)^2\sigma_3^2$$

$$Var\left[T_S^u\right] = 0.0629$$

$$Var\left[T_S^u\right] = \left(\frac{1}{4}\right)^2\sigma_1^2 + \left(\frac{1}{4}\right)^2\sigma_2^2 + \left(\frac{1}{4}\right)^2\sigma_3^2 + \left(\frac{1}{4}\right)^2\sigma_4^2$$

$$Var\left[T_S^u\right] = 0.0349$$

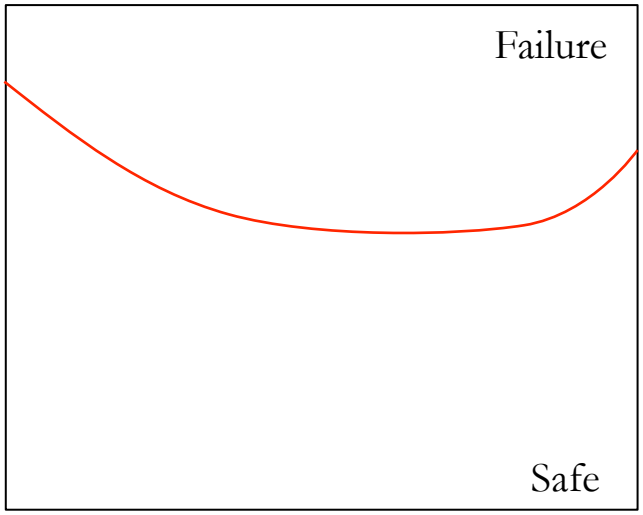
>30% reduction in variance by adding one stratum

>60% reduction in variance by fully stratifying

RSS Algorithms

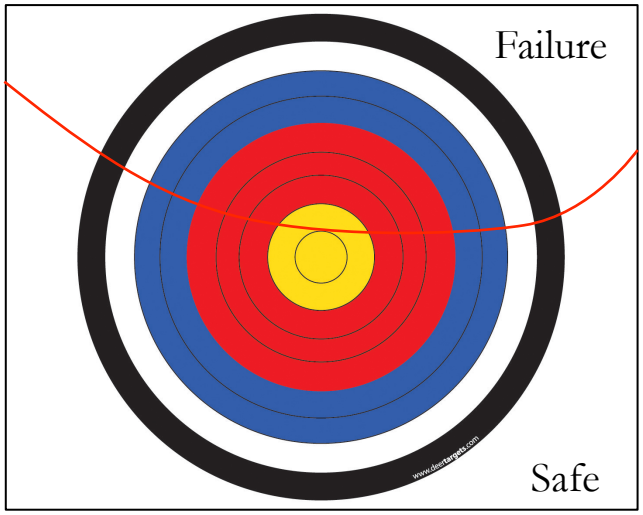
1. Random Stratum Division
2. Variance Minimization
3. Sensitivity-Based Stratum Division
4. Enhanced Space-Filling
5. Targeted Random Sampling
6. Others (Stochastic Search, Stochastic Field Simulation, etc.)

Targeted Random Sampling



Targeted Random Sampling

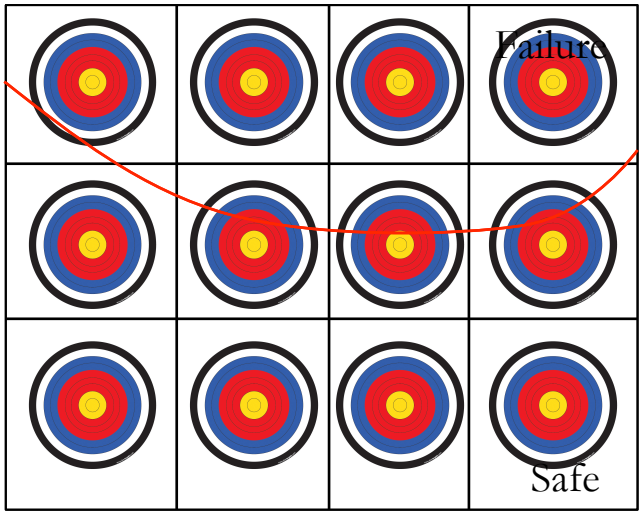
Random (iid) sampling



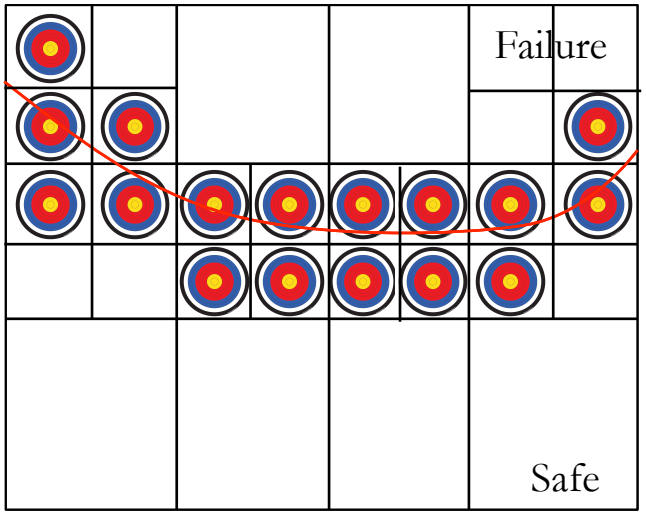
Targeted Random Sampling

Stratified Sampling/

Refined Stratified Sampling



Targeted Random Sampling



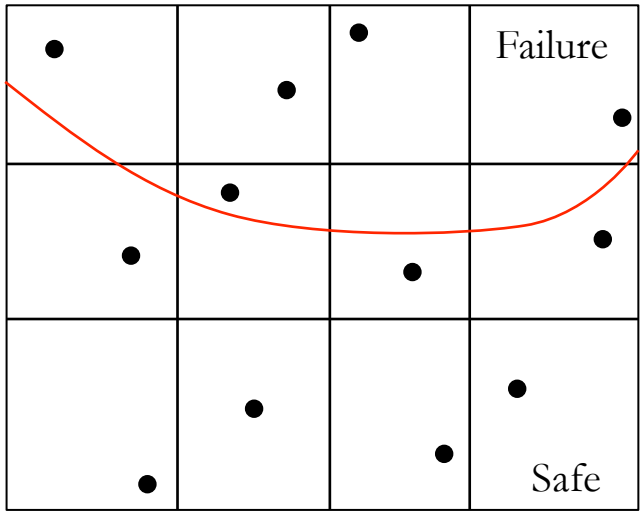
Targeted Random Sampling

The biggest problem with Monte Carlo methods is that they are <u>wasteful</u>

- Many simulations in regions that are, statistically, of little or no use
- This is especially problematic for reliability analysis where probability of failure is <u>very low</u>

Targeted Random Sampling

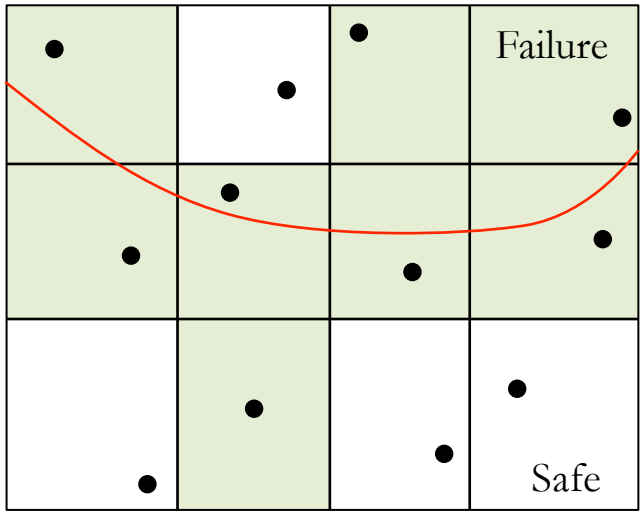
1. Define an initial stratification of the domain & sample
2. Identify point pairs on opposite sides of the failure surface
3. Identify the pair with the largest separation and divide the stratum at a point in between (alternatively, pair with largest associated probability)
4. Sample in the new stratum
5. Refine the target space
6. Repeat



Targeted Random Sampling

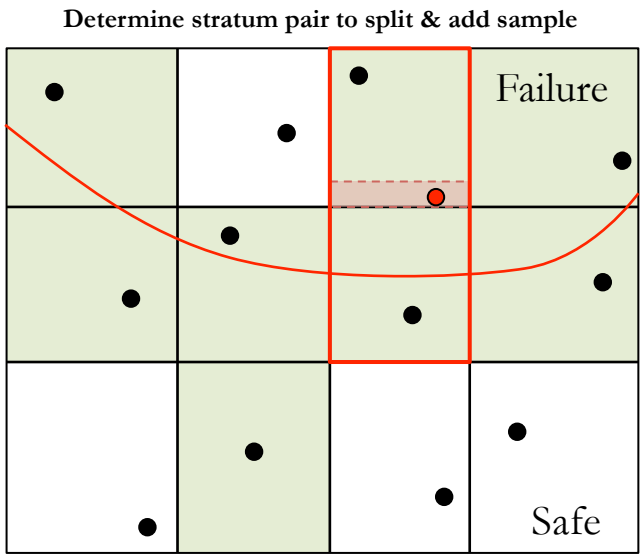
1. Define an initial stratification of the domain & sample
2. Identify point pairs on opposite sides of the failure surface
3. Identify the pair with the largest separation and divide the stratum at a point in between (alternatively, pair with largest associated probability)
4. Sample in the new stratum
5. Refine the target space
6. Repeat

Identify Strata Adjacent to Failure Surface



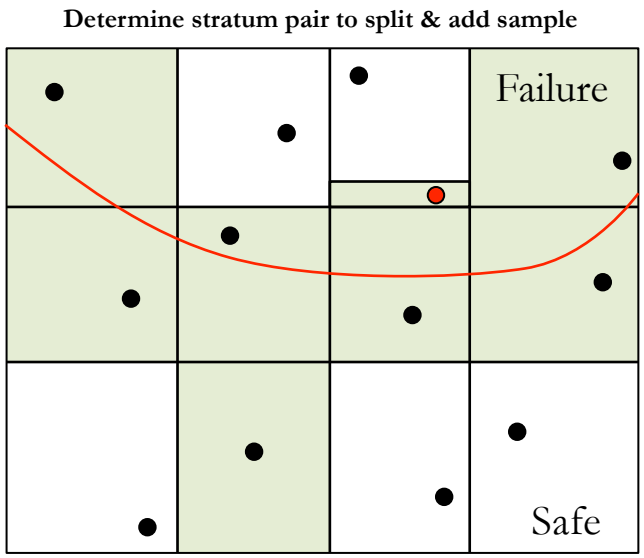
Targeted Random Sampling

1. Define an initial stratification of the domain & sample
2. Identify point pairs on opposite sides of the failure surface
3. Identify the pair with the largest separation and divide the stratum at a point in between (alternatively, pair with largest associated probability)
4. Sample in the new stratum
5. Refine the target space
6. Repeat



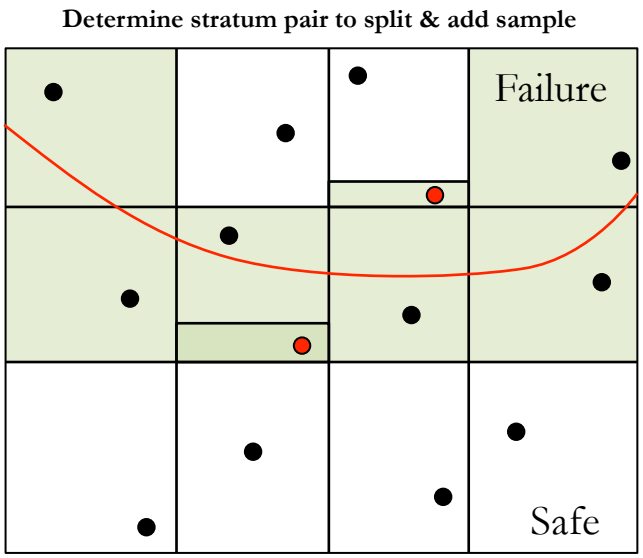
Targeted Random Sampling

1. Define an initial stratification of the domain & sample
2. Identify point pairs on opposite sides of the failure surface
3. Identify the pair with the largest separation and divide the stratum at a point in between (alternatively, pair with largest associated probability)
4. Sample in the new stratum
5. Refine the target space
6. Repeat



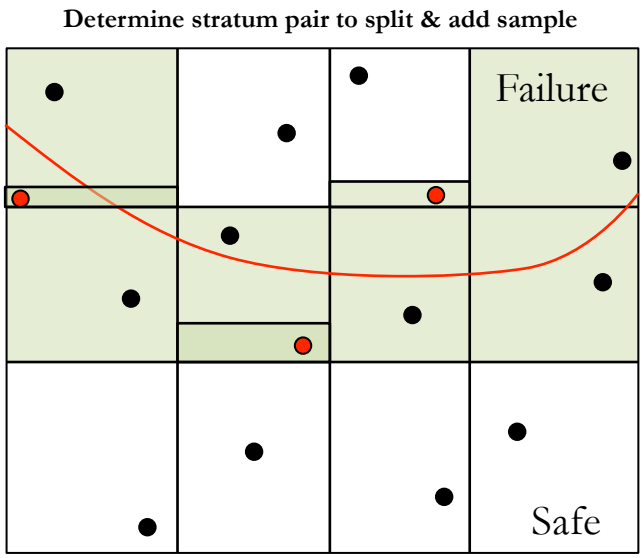
Targeted Random Sampling

1. Define an initial stratification of the domain & sample
2. Identify point pairs on opposite sides of the failure surface
3. Identify the pair with the largest separation and divide the stratum at a point in between (alternatively, pair with largest associated probability)
4. Sample in the new stratum
5. Refine the target space
6. Repeat



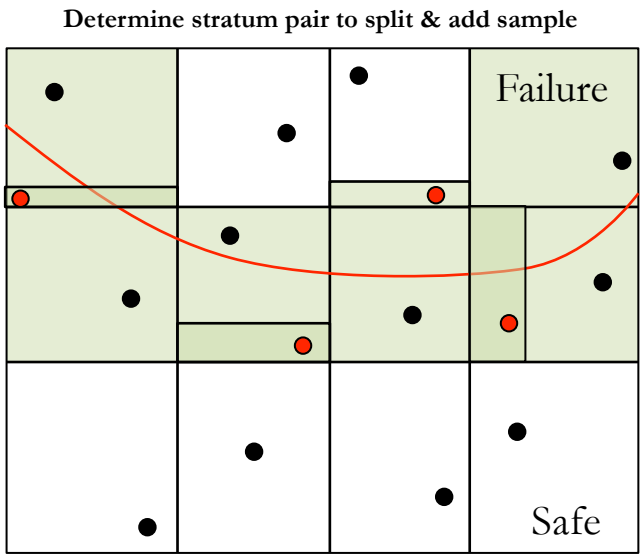
Targeted Random Sampling

1. Define an initial stratification of the domain & sample
2. Identify point pairs on opposite sides of the failure surface
3. Identify the pair with the largest separation and divide the stratum at a point in between (alternatively, pair with largest associated probability)
4. Sample in the new stratum
5. Refine the target space
6. Repeat



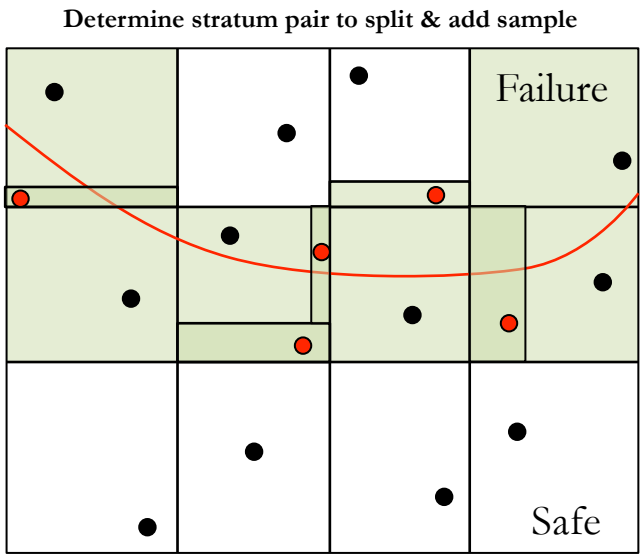
Targeted Random Sampling

1. Define an initial stratification of the domain & sample
2. Identify point pairs on opposite sides of the failure surface
3. Identify the pair with the largest separation and divide the stratum at a point in between (alternatively, pair with largest associated probability)
4. Sample in the new stratum
5. Refine the target space
6. Repeat



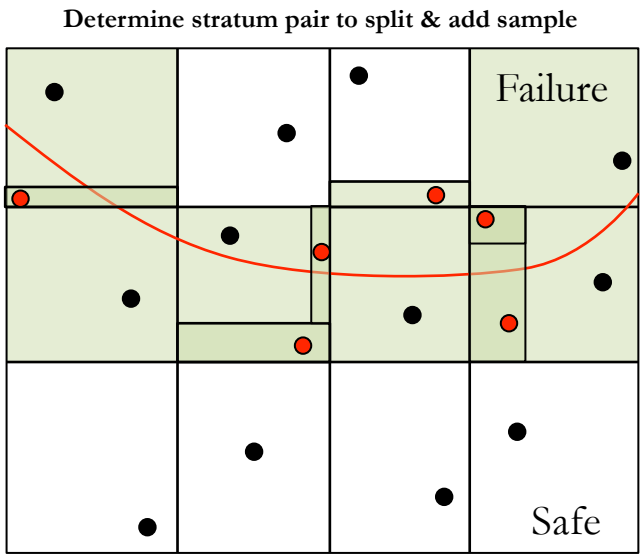
Targeted Random Sampling

1. Define an initial stratification of the domain & sample
2. Identify point pairs on opposite sides of the failure surface
3. Identify the pair with the largest separation and divide the stratum at a point in between (alternatively, pair with largest associated probability)
4. Sample in the new stratum
5. Refine the target space
6. Repeat



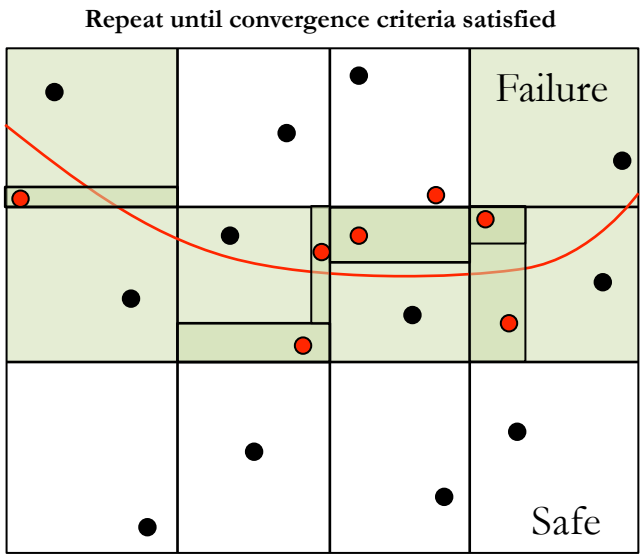
Targeted Random Sampling

1. Define an initial stratification of the domain & sample
2. Identify point pairs on opposite sides of the failure surface
3. Identify the pair with the largest separation and divide the stratum at a point in between (alternatively, pair with largest associated probability)
4. Sample in the new stratum
5. Refine the target space
6. Repeat



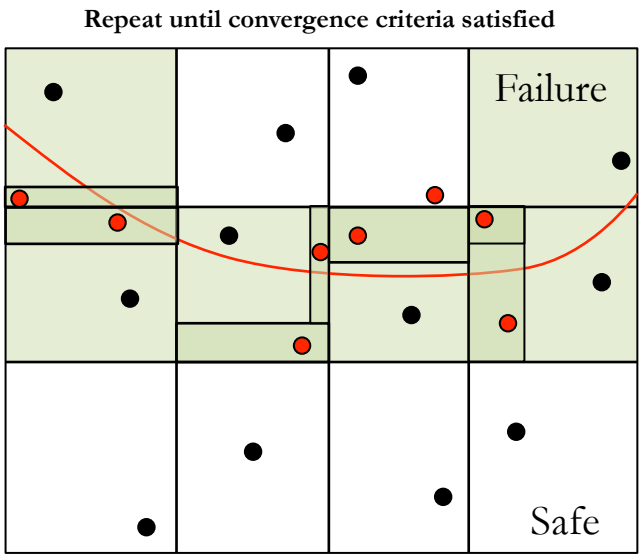
Targeted Random Sampling

1. Define an initial stratification of the domain & sample
2. Identify point pairs on opposite sides of the failure surface
3. Identify the pair with the largest separation and divide the stratum at a point in between (alternatively, pair with largest associated probability)
4. Sample in the new stratum
5. Refine the target space
6. Repeat



Targeted Random Sampling

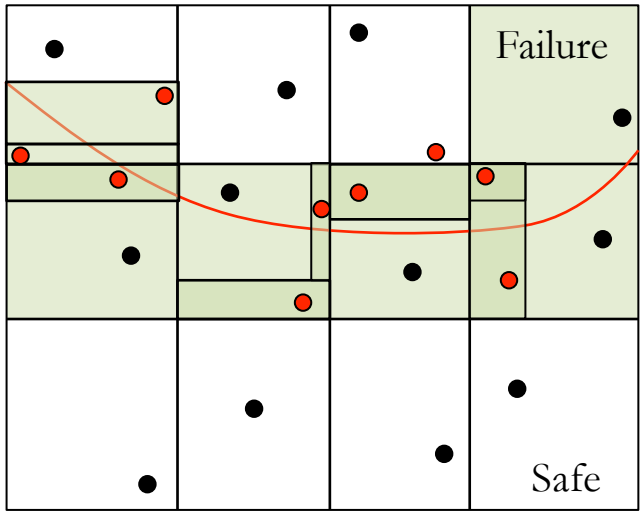
1. Define an initial stratification of the domain & sample
2. Identify point pairs on opposite sides of the failure surface
3. Identify the pair with the largest separation and divide the stratum at a point in between (alternatively, pair with largest associated probability)
4. Sample in the new stratum
5. Refine the target space
6. Repeat



Targeted Random Sampling

1. Define an initial stratification of the domain & sample
2. Identify point pairs on opposite sides of the failure surface
3. Identify the pair with the largest separation and divide the stratum at a point in between (alternatively, pair with largest associated probability)
4. Sample in the new stratum
5. Refine the target space
6. Repeat

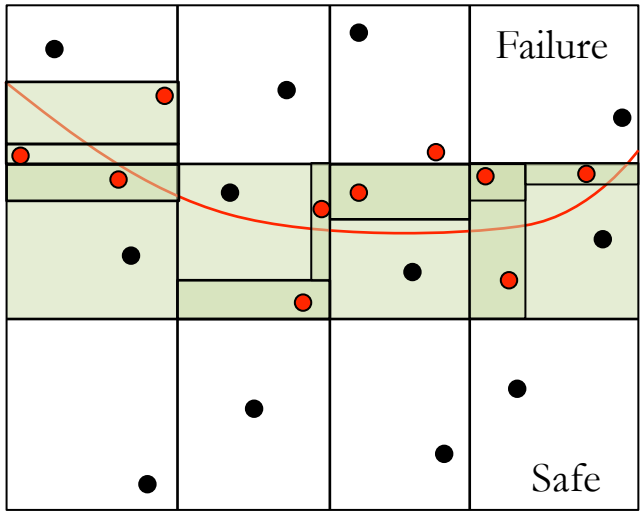
Repeat until convergence criteria satisfied



Targeted Random Sampling

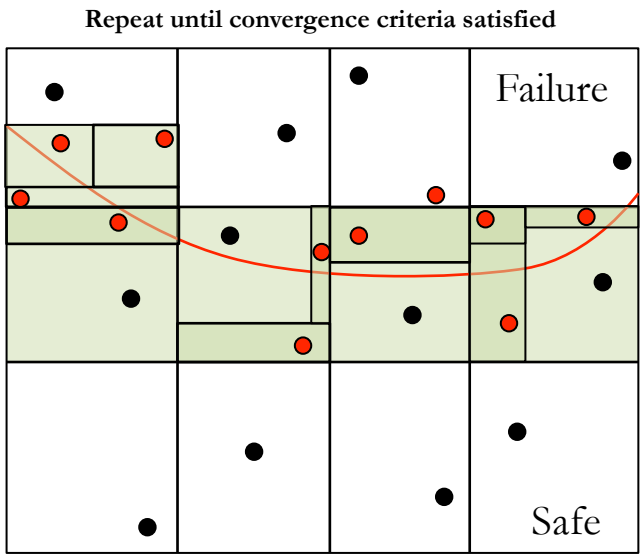
1. Define an initial stratification of the domain & sample
2. Identify point pairs on opposite sides of the failure surface
3. Identify the pair with the largest separation and divide the stratum at a point in between (alternatively, pair with largest associated probability)
4. Sample in the new stratum
5. Refine the target space
6. Repeat

Repeat until convergence criteria satisfied



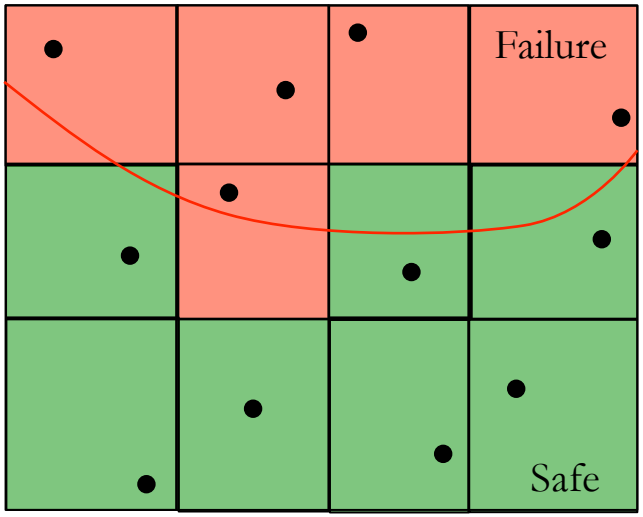
Targeted Random Sampling

1. Define an initial stratification of the domain & sample
2. Identify point pairs on opposite sides of the failure surface
3. Identify the pair with the largest separation and divide the stratum at a point in between (alternatively, pair with largest associated probability)
4. Sample in the new stratum
5. Refine the target space
6. Repeat



Targeted Random Sampling

1. Define an initial stratification of the domain & sample
2. Identify point pairs on opposite sides of the failure surface
3. Identify the pair with the largest separation and divide the stratum at a point in between (alternatively, pair with largest associated probability)
4. Sample in the new stratum
5. Refine the target space
6. Repeat



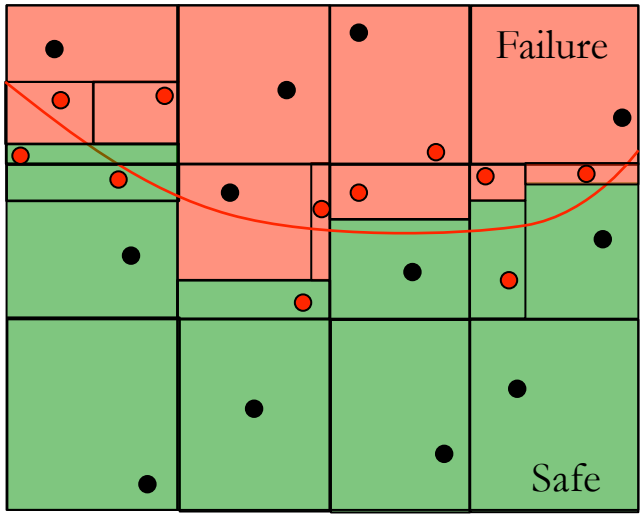
Targeted Random Sampling

1. Define an initial stratification of the domain & sample
2. Identify point pairs on opposite sides of the failure surface
3. Identify the pair with the largest separation and divide the stratum at a point in between (alternatively, pair with largest associated probability)
4. Sample in the new stratum
5. Refine the target space
6. Repeat

Notes:

- TRS constructs a piecewise approximation of the failure surface
- Need to identify at least one failure point per disjoint failure region
- Uses available information to inform stratification

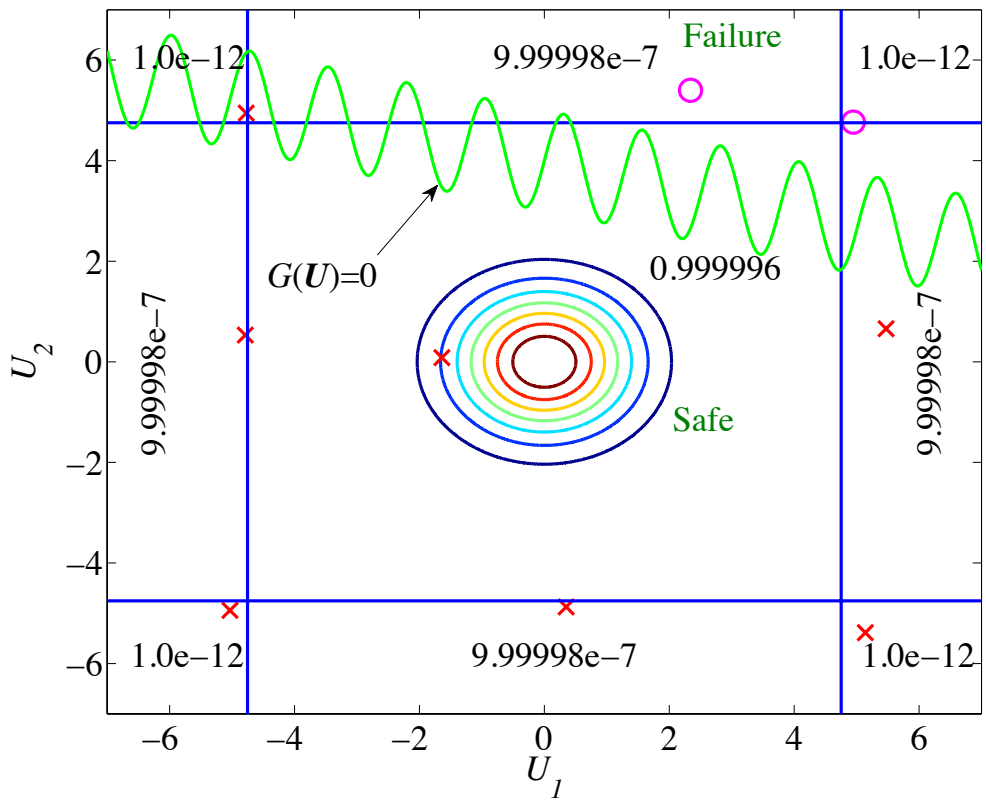
Repeat until convergence criteria satisfied



Example 1: Initial Sample

Limit State: $G(\mathbf{U}) = -\frac{U_1}{4} + \sin\left(5U_1\right) + 4 - U_2; \quad \mathbf{U} \sim N(0,1)$

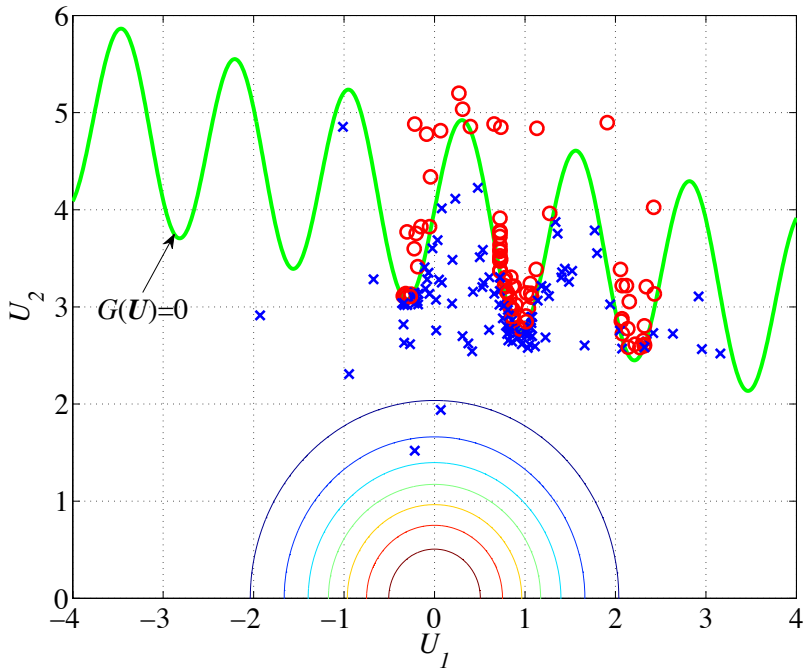
Probability of Failure: $p_f \approx 4.154 \times 10^{-4}$

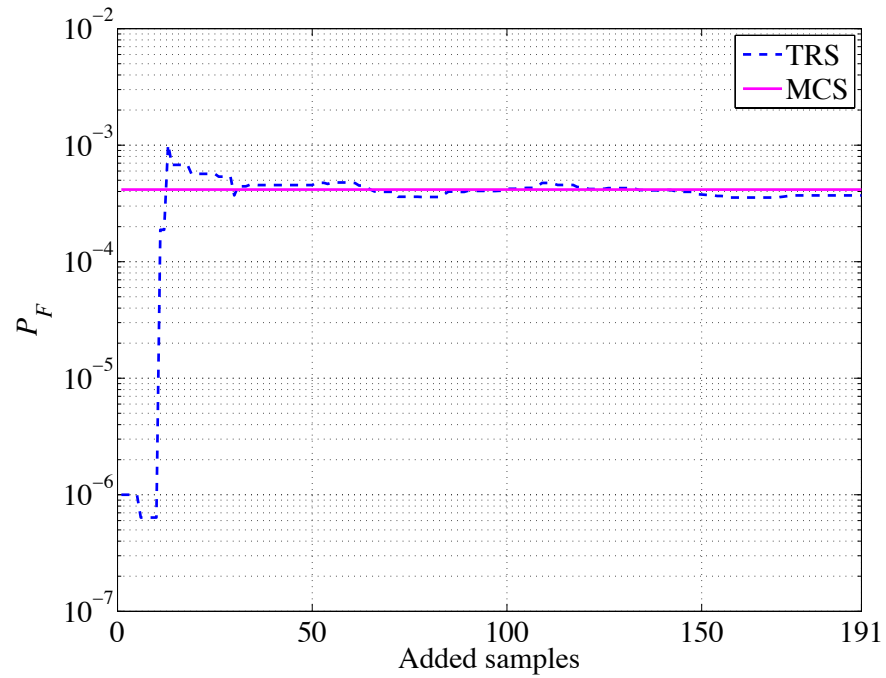


Example 1: 200 Samples

Limit State: $G(\mathbf{U}) = -\frac{U_1}{4} + \sin\left(5U_1\right) + 4 - U_2; \quad \mathbf{U} \sim N(0,1)$

Probability of Failure: $p_f \approx 4.154 \times 10^{-4}$

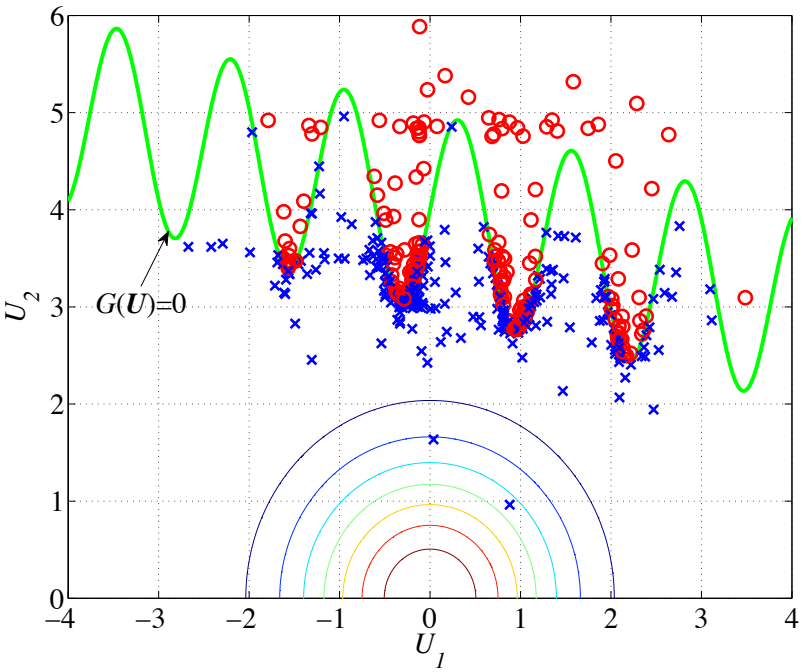


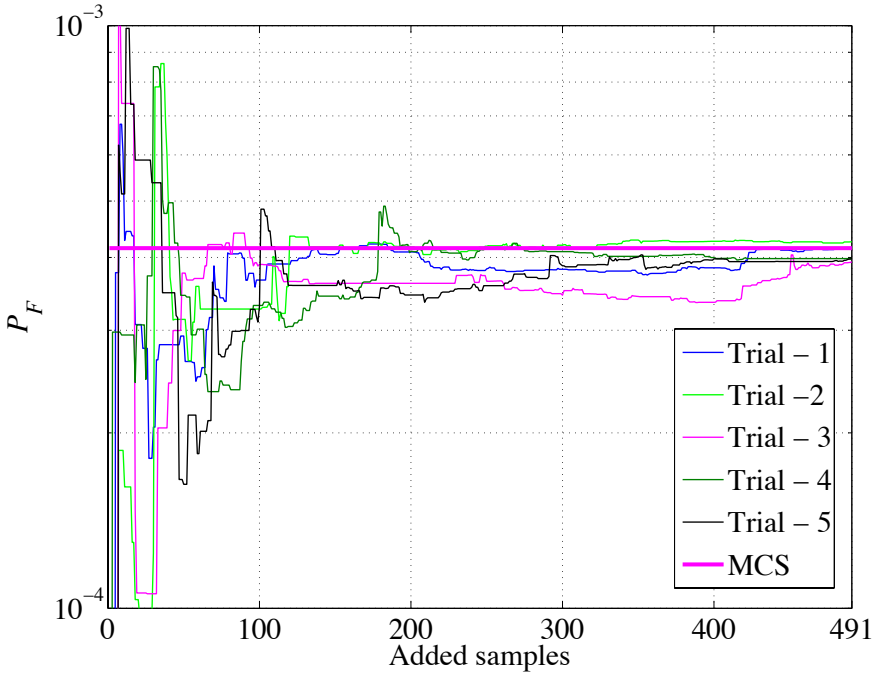


Example 1: 500 Samples

Limit State: $G(\mathbf{U}) = -\frac{U_1}{4} + \sin\left(5U_1\right) + 4 - U_2; \quad \mathbf{U} \sim N(0,1)$

Probability of Failure: $p_f \approx 4.154 \times 10^{-4}$

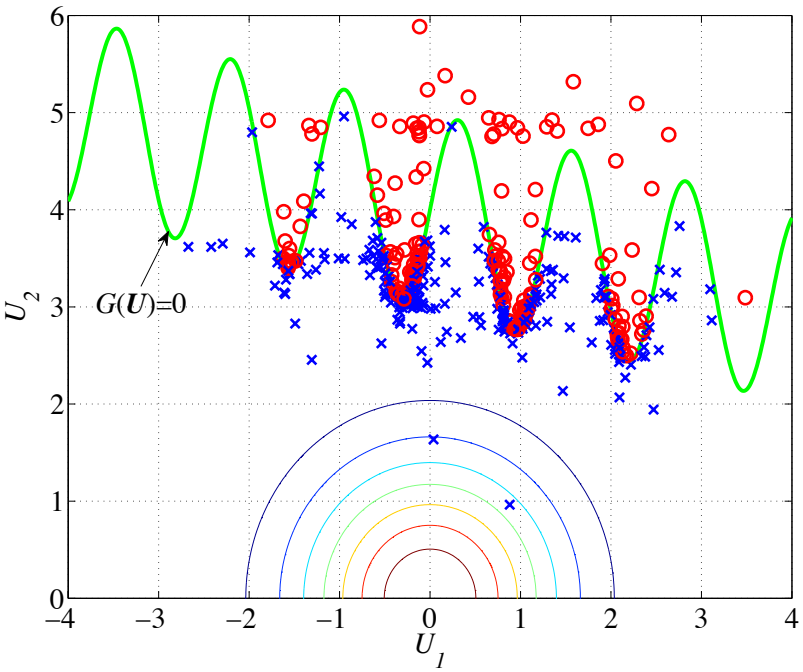




Example 1: Comparison

Limit State: $G(\mathbf{U}) = -\frac{U_1}{4} + \sin\left(5U_1\right) + 4 - U_2; \quad \mathbf{U} \sim N(0,1)$

Probability of Failure: $p_f \approx 4.154 \times 10^{-4}$

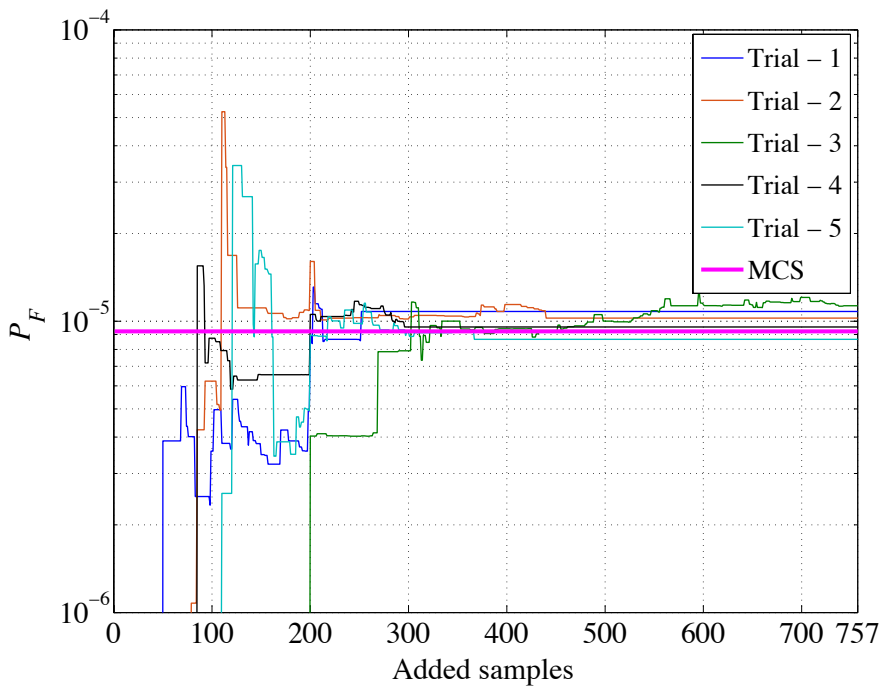


Method	Mean P_F	COV
MCS (10^8 samples)	4.154e-4	
FORM	N/A	
SORM	N/A	
Importance Sampling (IS)	N/A	
SS(500; 50; 10) (~1000 samples)	4.2641e-4 (100)	66%
TRS (9+491) (500 samples)	3.9998e-4 (100)	4%

Example 2: Comparison

Limit State*: $G(\mathbf{X}) = \sum_{i=1}^{5} X_i - C; \quad X_i \sim Exp(\lambda_i); \lambda_i = 1 \, \forall i; C = 0.26715$

Probability of Failure: $p_f \approx 9.2300 \times 10^{-6}$



Method	Mean P_F	COV
MCS (10^8 samples)	9.2300e-6	3%
FORM	1.3913e-4	
SORM	1.4847e-4	
IS (1000 samples)	8.3715e-6	
SS(500; 50; 10) (~1000 samples)	9.8664e-6 (100)	50%
TRS (243+757) (1000 samples)	9.4830e-6 (100)	20%

1. Engelund, S. and Rackwitz, R. (1993). A benchmark study on importance sampling techniques in structural reliability. *Structural Safety*. 12: 255-276.

Challenges & Future Directions

- Defining an appropriate initial stratification in high dimension
  - $3^{20} \sim 3.5e+9$
  - Use MCMC to explore the space and post-stratify
- Dynamic Problems & Stochastic Processes/Fields

- Other RSS designs
  - Variance Minimization for UQ
  - Sensitivity-based Sampling
  - Space-filling sample designs
  - Etc.