10th World Congress on Structural and Multidisciplinary Optimization
May 19 -24, 2013, Orlando, Florida, USA

Parallel Genetic Algorithm with Population-Based Sampling Approach to Discrete Optimization under Uncertainty

Nithya Subramanian[1], William A. Crossley[2]

[1] Purdue University, West Lafayette, Indiana, USA, nsubram@purdue.edu
[2] Purdue University, West Lafayette, Indiana, USA, crossley@purdue.edu

1. Abstract

This paper presents the Genetic Algorithm (GA) with a Population-Based Sampling (PBS) technique that enables optimization under uncertainty with discrete variables at a lower computational expense than using Monte Carlo Sampling (MCS) for every fitness evaluation. The optimization of composite laminates using ply orientation angles as discrete variables provides an example to demonstrate further developments of the GA with Population-Based Sampling for discrete optimization under uncertainty. The focus problem aims to reduce the expected weight of the composite laminate while treating the laminate's fiber volume fraction and externally applied loads as uncertain quantities following normal distributions. Results indicate a reduction in computational cost by three orders of magnitude through the implementation of GA-PBS in comparison to the GA-MCS. The study also investigates two different implementations of parallel computation in the GA-PBS and compares the results from these two schemes.

2. Keywords: discrete optimization, uncertainty, composite laminate design, population-based sampling

3. Introduction

Researchers from several engineering and non-engineering disciplines have demonstrated the use of probabilistic design approaches to avoid excessively conservative designs. Design with a safety margin typically yields expensive products without an accurate risk assessment. The design approach using safety margins only gives a point estimate of the structural performance whereas probabilistic methods predict an interval estimate for the probability of successful structural performance. However, design optimization under uncertainty can result in high computational costs.

Making use of traditional sampling techniques such as the Monte Carlo Sampling (MCS) in an optimization framework proves to be computationally expensive because a large number of samples are needed to predict the expected objective and / or constraint values of each design generated in the optimization run to provide high confidence levels for the predicted values. From the analysis of previous research efforts, the MCS technique is a popular sampling method accompanying zero-order methods such as the GA in reliability based optimization.

Cantoni et al. [1] demonstrated the use of GA with MCS for an optimal industrial plant design under conflicting safety and economic constraints. A profit function is employed as the fitness function to be maximized by altering the value of five design variables. The study also deemed the full run of MCS for each potential design impractical and suggested a modified selective sampling approach. The combination of GA with MCS was also used to assess the reliability of a water quality system by obtaining optimum waste load allocation solutions [2]. The total treatment cost is minimized while achieving a specified level of reliability of meeting the water quality standard at a critical location. The authors introduced a First Order Reliability Method (FORM) to replace the computationally expensive MCS scheme without significant loss of accuracy.

Population-Based Sampling (PBS) aims to overcome the challenge of high computational costs associated with probabilistic optimization methods. Crossley [3] originally proposed this sampling concept for the GA using a three-bar truss as an example problem. Subsequently, Hassan [4] further refined the approach and used it for the discrete optimization problem presented by the conceptual design of spacecraft. This study adds some improvisations to the existing GA with PBS technique for optimization under uncertainty. The work proposes and demonstrates the use of two new sample accumulation techniques for PBS – design count-based increase in sample size and stepwise increase in sample size. The design count-based sample accumulation technique exploits the convergence characteristics of the GA to assign samples to designs in a generation. The stepwise sample accumulation technique uses the current generation number of the GA run to assign an appropriate sample size. In addition, parallel computation exploits the independence of individual fitness function evaluations and conducts these calculations in parallel. In general, the new sample accumulation schemes coupled with parallel computing efforts has shown significant reduction in computational time and cost of the GA.

Design of composite materials is complicated due to the number of variables involved both at the material and structural level. In laminated composite structures, each ply has its greatest stiffness and strength properties along the direction of the fibers. By orienting each layer at different angles relative to each other, the structure can be designed for a specific loading environment. In this study, the discrete optimization problem with PBS aims to reduce the expected weight of a composite laminate while treating the externally applied loads and the variation in fiber volume fraction as uncertain quantities following a normal distribution. A sample consists of a single instance from each of the uncertain parameter distributions. The constraints enforced include the probability of satisfying the Tsai-Hill failure criterion and the maximum strain limit. The calculations to establish the expected values of constraints and fitness values use the Classical Laminate Theory (CLT).

4. Relevant Methods and Analysis Tools

4.1. Analysis of Composite Laminates

The Classical Laminate Theory is the analysis tool employed in the study for discrete optimization under uncertainty to determine the performance characteristics of the laminate. CLT combines material properties and fiber orientation to account for extensional, flexural and torsional deformations and the coupling effects between these deformations. The reduced stiffness matrix for each lamina is constituted based on its mechanical properties [5].

The matrix *A* is called the extensional stiffness matrix, *B*, the coupling stiffness matrix and *D*, the bending stiffness matrix. The in-plane deformations in the laminate are related to the external loads by Eq. (1).

$$\begin{bmatrix} N \\ M \end{bmatrix} = \begin{bmatrix} A & B \\ B & D \end{bmatrix} \begin{Bmatrix} \varepsilon_o \\ \kappa \end{Bmatrix} \tag{1}$$

Symmetric laminates have no extension-bending coupling – a pure in-plane loading on these laminates will not cause an out-of-plane deformation. Hence, the coupling matrix, *B*, for such laminates is a zero matrix. Symmetric laminates are highly preferred in manufacturing for this same reason and this study restricts all laminate designs to be symmetric.

Fiber composites are a heterogeneous media, but a macro homogeneous solid with certain effective moduli that describe the average material properties of the composite may effectively represent them. The square fiber model is used as the micromechanics model to evaluate the average material properties. This model assumes the fiber to be a square with fiber volume fraction v_f. Based on this assumption, the effective moduli along the *x*, *y*, and *z* directions are obtained from the reduced stiffness coefficients using Eqs. (2), (3), and (4).

$$\overline{E}_x = \frac{1}{A_{11}t} \tag{2}$$

$$\overline{E}_y = \frac{1}{A_{22}t} \tag{3}$$

$$\overline{G}_{xy} = \frac{1}{A_{66}t} \tag{4}$$

For isotropic materials, the von Mises yield criterion states that the yielding of a material occurs when the distortion energy density equals the distortion energy density at yield in pure tension. The Tsai-Hill theory is an extension of the von Mises failure criterion for orthotropic materials [6]. The Tsai-Hill criterion assumes known values of the failure strengths in the principal directions. The yield criterion obtained from this appears in Eq. (5), where values above 1.0 suggest failure; this criterion becomes a constraint in the discrete optimization problem.

$$\left(\frac{\sigma_{11}}{X}\right)^2 + \left(\frac{\sigma_{22}}{Y}\right)^2 - \left(\frac{\sigma_{11}\sigma_{22}}{X^2}\right) + \left(\frac{\sigma_{12}}{S}\right)^2 = 1 \tag{5}$$

4.2. Genetic Algorithm

The Genetic Algorithm (GA) is a heuristic search process that mimics the process of natural selection. The probabilistic elements of the search technique make the GA likely to search across the entire design space and not be trapped in local minima. This adaptive search process, which includes some randomized features, has been widely used in the design optimization of complex systems; but the GA requires a fairly high computational cost compared to calculus-based methods. The GA, however, can also address problems that calculus-based methods cannot, including designs with discrete variables. Before implementing the GA, the potential range of the design variables are coded into strings that the computer can process. It is common to encode the design variables as binary strings of 1's and 0's.

This study uses the MATLAB GA and employs Gray Coding, tournament selection, scattered crossover and uniform mutation. The tournament selection process picks two individuals at random without replacement, from the population of each generation. The fitness values of these two individuals are compared and the individual with a better (lower for a minimization problem) fitness value is copied over to the mating pool while the other individual is discarded. The scattered crossover process uses a crossover fraction P_c of 0.8, indicating 80% of a given population undergo crossover to form children for the next generation (the other 20% pass to the next generation directly). Scattered crossover generates a random binary vector to describe the crossover; when the vector element is 1, the gene from the first parent is selected and when the vector element is 0, the corresponding gene from the second parent is selected. The GA used here only generates one child from a set of two parents. Figure 1 illustrates the crossover procedure with an example.

$$\begin{aligned} \text{parent 1}: & \begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 0 \end{bmatrix} \\ \text{parent 2}: & \begin{bmatrix} 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix} \\ \text{binary vector}: & \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 1 \end{bmatrix} \\ \text{child 1}: & \begin{bmatrix} 1 & 1 & 0 & 1 & 1 & 0 \end{bmatrix} \end{aligned}$$

Figure 1: Illustration of scattered crossover procedure

The mutation process introduces diversity in the population by randomly switching a gene and encourages better exploration of the design space. Mutation is often a secondary operator performed with a low probability. The mutation operator used for this work is the uniform mutation operator with a probability P_m, of 1.98×10^{-3}; the value of P_m is obtained based on Eq. (6) [7].

$$P_m = \frac{Chromosome\ length + 1}{2 \times Population\ size \times Chromosome\ length} \tag{6}$$

A good mutation rate encourages exploration but does not turn the GA into a random search. If a mutation rate is good, a design with the newly introduced trait will have a good fitness value and survive to pass that trait to its offspring. The poor designs generated as a result of mutation will have poor fitness and will not survive the tournament selection.

The GA cannot handle constraints explicitly; the fitness function, therefore, reflects the 'goodness' of a design and it incorporates the objective function and information about any violated constraints. Commonly, penalty methods are used to handle violated constraints in the GA. Penalty methods convert a constrained optimization problem into an unconstrained one by adding a penalty term to the objective function when the constraints are violated. This work implements a quadratic exterior penalty method to handle violated constraints and the general construct of this penalty method appears in Eqs. (7) and (8).

$$\varphi\left(\bar{x}\right) = f\left(\bar{x}\right) + r_p \times P\left(\bar{x}\right) \tag{7}$$

$$P\left(\bar{x}\right) = \left\{\max\left[0, g\left(\bar{x}\right)\right]\right\}^2 \tag{8}$$

Handling constraints using the exterior penalty method function approach discourages the propagation of infeasible designs but allows designs with very small constraint violations to survive. The appropriate choice of value for the penalty multiplier r_p, is generally problem dependent, and the r_p value used here is 50.

5. Problem Description

In probabilistic approaches, for a predetermined accuracy of prediction, the confidence level associated with the predicted values of the uncertain aggregate function is proportional to the number of samples ($N_{samples}$) used to predict the expected values of those uncertain functions. This study treats the objective function and constraints as functions of input from uncertain parameters. The mathematical formulation of the expected value of the objective function and constraints appear in Eqs. (9) through (13).

Minimize:
$$E\left(f(\bar{x}, \xi)\right) = E\left(mass\ of\ laminate\right) \tag{9}$$

Subject to: $$G_j(\bar{x},\xi) = \frac{\sum_{i=1}^{N_{samples}} g_j(\bar{x},\xi)}{N_{samples}} \leq 0; \; j = 1,2,3 \tag{10}$$

The constraints g_j are as follows.

$$g_1(\bar{x}) = Tsai\text{-}Hill\ factor\ (\bar{x}) - 1 \leq 0 \tag{11}$$

$$g_2(\bar{x}) = 1 - \frac{Laminate\ Effective\ Modulus\ (\bar{x})}{Minimum\ Modulus} \leq 0 \tag{12}$$

$$g_3(\bar{x}) = \frac{Mid\text{-}plane\ Strain\ (\bar{x})}{\varepsilon_{limit}} - 1 \leq 0 \tag{13}$$

Each sample (ξ) in this study contains a random instance from the normal PDF of external loads and the variation in fiber volume fraction.

$$\xi = \begin{Bmatrix} \Delta v_f \sim N\left(0, \sigma_{v_f}^{\ 2}\right) \\ \left[F_x, F_y\right] \sim N\left(\mu_F, \sigma_F^{\ 2}\right) \end{Bmatrix} \tag{14}$$

Fitness function: $$\varphi(\bar{x},\xi) = E\left(f(\bar{x},\xi)\right) + r_p \times \sum_{j=1}^{3} \max\left[0, G_j(\bar{x},\xi)\right]^2 \tag{15}$$

Sixteen design variables describe the composite laminate. These variables correspond to one half of the stack of ply orientation angles for the plate. This formulation limits the maximum number of plies in the plate to 32. Eight discrete options are available for the ply orientation angles, 0º, 90º, 45º, 30º, 60º, -45º, -30º, -60º. A 'no-ply' option is coded into the algorithm as a discrete option for a design variable, making it possible to have any even-numbered value between 0 and 32 for the total number of plies. The chromosome length for this case study is 64 bits (16 variables, each represented by four bits). Based on the chromosome length, a constant population size of 256 (4 × chromosome length) is employed for the GA run.

The stopping criterion for the GA with Monte Carlo Sampling (MCS) runs two checks; the stall generation limit is set to be 40 and the average change in the best fitness value beyond the stall generation limit must be less than 1×10^{-3}. Furthermore, the best design should have satisfied all constraints for at least 99% of the samples evaluated. At a confidence level of 99%, each individual in a generation uses 10,000 samples to evaluate the expected values of the objective function and constraints. Thus, the number of fitness evaluations for a single generation is 2,560,000. The study, therefore, implements the Population-Based Sampling (PBS) scheme to perform design optimization under uncertainty using a GA with significantly lower computational cost than the traditional Monte-Carlo Sampling method.

6. Sample Size Assignment in Population-Based Sampling

The PBS approach makes use of the large number of individual designs evaluated by the GA in each generation. Good designs tend to appear in multiple generations and eventually, the population has multiple copies of such good designs in the same generation. If each fitness evaluation for the good designs uses a few samples, the opportunity exists to accumulate a large number of samples for the same design. This large number of samples gives good estimates for the expected objective function values and constraints for design solutions with desirable design characteristics. Hassan and Crossley [8, 9] investigated various reliability-based methods to optimize the configuration of a spacecraft system. The PBS method was mathematically formulated and successfully applied to the optimization of spacecraft configuration with less than 0.2% variation in accuracy in comparison to MCS. The use of PBS method showed a reduction in the number of function evaluations by two orders of magnitude.

This work aims to further improve the computational efficiency of the PBS technique with two different sample accumulation methods – design count-based increase in sample size and stepwise increase in sample size. The mean of the objective function and constraints are calculated with the available sample size in both methods. Every individual evaluated by the GA is stored and fitness function values are updated as the individual accumulates more samples over the course of the GA run. The mathematical formulation of these two sample accumulation methods appears in Table 1. The wall clock time of the GA-PBS is reduced by incorporating parallel computation.

Table 1: Sample accumulation methods

Design Count-Based Increase in Sample Size	Stepwise Increase in Sample Size
$p(\bar{x}_{best})_{gen-1} = P(\bar{X} = \bar{x}_{best})_{gen-1}$ $N_{samples,gen} = k_1 \times p(\bar{x}_{best})_{gen-1}$	$N_{samples}(\bar{x})_{n_{gen}} = \begin{cases} c_1 & if \quad n_{gen} < 20 \\ ceil\left(\dfrac{c_2}{\left(\left\lvert f(\bar{x}) - f(\bar{x}_{best}, \xi)_{n_{gen}-1}\right\rvert + \varepsilon\right)}\right) & if \quad 20 \le n_{gen} \le 40 \\ ceil\left(\dfrac{c_3}{\left(\left\lvert f(\bar{x}) - f(\bar{x}_{best}, \xi)_{n_{gen}-1}\right\rvert + \varepsilon\right)}\right) & if \quad n_{gen} > 40 \end{cases}$

The first sample accumulation method counts the number of occurrences of the best design in the previous generation and assigns a proportional sample size to each individual in the current generation. The constant of proportionality (k_1) used for the GA-PBS runs is 18. This constant value was chosen based on trial runs to strike a balance between the risk of early convergence of the GA (resulting in a far from optimal solution) and the risk of prolonged run time (resulting in high computational expense). Based on this sample size assignment technique, if the best design appeared three times in the previous generation, each individual in the current generation would receive 54 (18×3) different samples of ξ.

The design count-based sample assignment process makes no distinction in sample sizes assigned to good designs and poor designs in the same generation. Because the sample size for every design in a generation depends only on the frequency of occurrence of the best design, the individuals with good fitness values and those with poor fitness values in the same generation get the same number of samples. The PBS scheme with stepwise increase in samples first determines an individual's fitness function value with one sample. This sample is just a collection of the deterministic design variables and the mean values of the uncertain parameters. The magnitude of difference between the individual's fitness function value and the fitness value of the best design (up to that point in the GA run) is determined, and a sample size that is inversely proportional to this difference is assigned to the individual. Thus, the designs with poor fitness function values receive fewer samples.

The sample size also depends on how far the GA run has progressed. The first 20 generations have been largely exploratory for this specific focus problem (this needs prior assessment of convergence trends) and fewer samples are expended at this stage of the GA run. A large value of c_1, c_2 and c_3 would cause the GA to converge faster but introduces more samples per design evaluation. Since the stopping criterion requires the best design to have accumulated a total of 10,000 samples, large values for these constants will cause the GA to converge in fewer generations. However, the caveat here is the possibility of premature convergence. In this case study, c_1, c_2 and c_3 are assigned values 5, 1 and 2, respectively. The constant value ε is positive and prevents division by zero if the fitness value of the current design equals that of the best design. The constant ε is assigned a value of 0.1 in this study. The ceiling function ensures that every design gets at least one sample during the PBS process.

The stopping criterion to halt the GA run with the PBS approach has three parts. The first part checks the total number of samples accumulated by the best design. The run is not stopped until the design with the best fitness value has accumulated at least a specified number of samples, depending on the confidence level required. The value of the total sample limit based on the required confidence interval appears in Table 2. The second part checks for the average change in the best fitness value in the last ten generations to be less than 1×10^{-3}. The last part of the stopping criterion ensures that the best design has not changed in the last five generations.

Table 2: Total sample limit based on required confidence level

Confidence Level (%)	Total number of samples
90	100
95	400
97.5	1600
99	10,000

7. Results

7.1. Sample Accumulation Results

Figures 2 and 3 depict the accumulation of samples by the best design from the beginning to the end of a GA run from the design count-based and stepwise increase sample assignment techniques, respectively. The stopping criterion is checked only at the end of each generation. Therefore, the total number of samples accumulated by the best design may cross 10,000 at the end of the last generation.

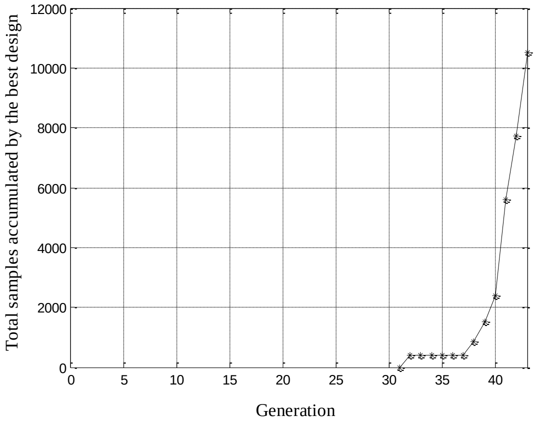

Figure 2: Sample accumulation from design count-based increase in samples

For this particular GA run that lasted 44 generations, the optimal design is identified in the 32^{nd} generation and it accumulated 10,512 samples at the end of the 44^{th} generation. This activates the stopping criterion of the GA –PBS at 99% confidence level.

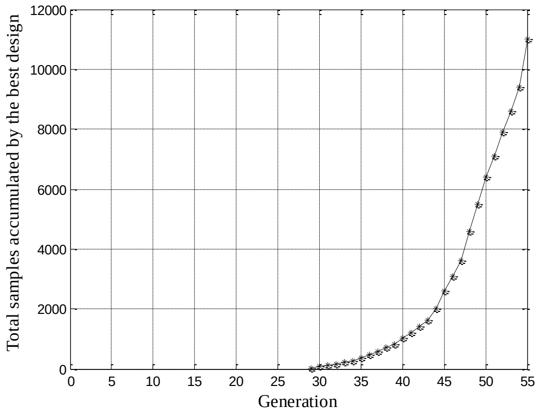

Figure 3: Sample accumulation from stepwise increase in samples

In this GA run with stepwise increase in samples, the best design first appeared in the 30^{th} generation. This design received 20 samples in this generation based on the comparison of its fitness value to the best value previously encountered by the GA. Eventually, there were 93 copies of the best design in the 55^{th} generation; the individual received a total of 11,000 samples at the end of the GA run.

7.2. Comparison of Results from GA-MCS and GA-PBS

Figures 4 through 6 summarize the results of the two non-deterministic approaches – MCS and PBS. Three GA runs were conducted at four confidence levels – 90%, 95%, 97.5%, and 99%. The mean value of the laminate mass from the three GA runs is plotted at four confidence levels in Fig 4. Figures 5 and 6 illustrate the mean computational time and cost associated with the GA runs with the two non-deterministic techniques at the four confidence levels. Note that Figs. 5 and 6 use a $\log_{10}$ scale for clarity of variation across large orders of magnitude.

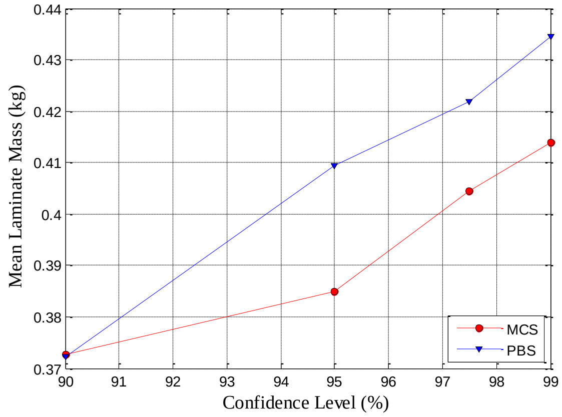

Figure 4: Mean laminate mass found using MCS and PBS at various confidence levels

The total mass of the optimal laminate design at 90% confidence from both approaches is same up to three decimal places. However, at higher confidence levels, the PBS approach results in a heavier composite laminate. Every estimation in the GA-MCS technique has the same level of accuracy because they have the same number of samples. However, in GA-PBS, the designs eliminated in the initial phase are associated with poor estimations based upon small sample sizes. 'Useful' designs may get eliminated in the initial phase of the GA-PBS owing to poor fitness estimation and not re-emerge in the GA run, thus resulting in a heavier optimal laminate configuration. The average variation in laminate mass using the two approaches is calculated to be 0.0157 kg, which corresponds to a difference of one ply in the laminate.

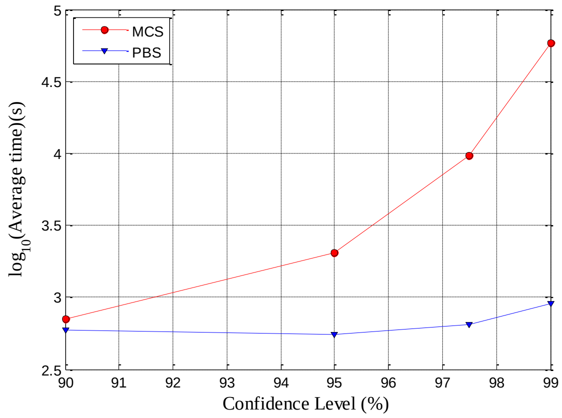

Figure 5: Average time taken by GA-PBS and GA-MCS at various confidence levels

The fitness analyses for both approaches were run in parallel on an eight core server for meaningful comparison. At a confidence level of 90%, the average GA-PBS run time is observed to be 596.48 seconds. However, the average run time for the same algorithm at 95% confidence level is 7.4% less despite the four-fold increase in total sample size. The MCS fitness analyses and sampling processes, on the other hand, are independent of each other – no data transfer is required among the processors. Hence, the average computational time is linearly related to the total number of samples expended in the MCS technique. At the highest confidence level of 99%, the GA-PBS run is found to be 98.44% faster than its MCS counterpart.

Figure 6 compares the total number of samples expended in a single run of both algorithms at different confidence levels. Note that this could also be interpreted as the total number of function evaluations by the GA in a single run.

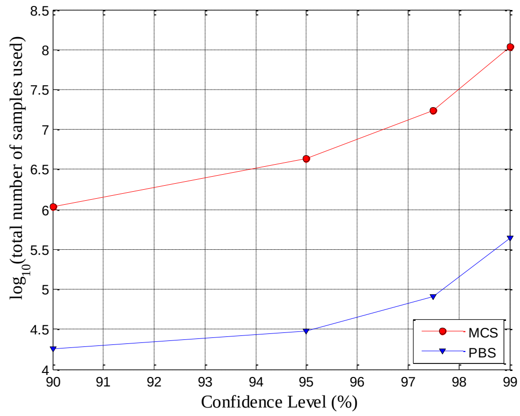

Figure 6: Total number of samples used in GA-MCS and GA-PBS at various confidence levels

At 90% confidence level, the MCS used 1,075,200 samples whereas the PBS used 17,819. At the largest confidence level, the GA-MCS evaluated 99.59% more samples relative to the PBS approach. The MCS approach has a large computational cost because each fitness evaluation, whether for poor or good designs, requires a fixed number of samples to be used. The improved sample accumulation techniques reduced the number of function evaluations in GA-PBS by three orders of magnitude relative to the MCS approach. This reduction in computational cost due to the improved sample accumulation methods is an order of magnitude better than previously achieved results [4].

Tables 3 and 4 show the laminate configurations resulting from the GA-MCS and GA-PBS runs. While this study conducted six GA-PBS runs to help demonstrate repeatability, the high computational expense restricted the number of GA-MCS runs to three.

Table 3: Summary of results from study with GA-MCS

Run	Number of iterations	Ply orientation sequence (deg)	Fitness Value	Deterministic mass (kg)
1	42	[60 0 -60 0 90 0]s	0.37249	0.3725
2	42	[60 0 -60 0 90 0]s	0.37255	0.3725
3	42	[45 90 0 90 0 0]s	0.37252	0.3725

All runs result in a 12-ply laminate with a deterministic mass of 0.3725 kg; there are two different ply orientations identified. The deterministic mass is calculated using the mean value of the uncertain parameters and other fixed values. Note that the expected fitness value differs from the deterministic fitness value listed because the expected fitness value depends on the sample of uncertain parameters used in the calculation.

Table 4: Summary of results from study with GA-PBS

Run	Number of iterations	Ply orientation sequence (deg)	Fitness Value	Deterministic mass (kg)	Reliability
1	53	[90 0 30 0 0 90 90]s	0.43462	0.43463	1
2	49	[90 0 30 0 0 90 90]s	0.4345	0.43463	1
3	54	[90 0 30 0 0 90 90]s	0.43458	0.43463	0.9994
4	64	[45 90 0 90 0 0]s	0.37251	0.3725	0.9901
5	62	[-60 0 90 30 0 0 90]s	0.43456	0.43463	0.9998
6	44	[90 0 30 0 0 90 90]s	0.43447	0.43463	0.9972

Four out of six runs indicate that the best laminate configuration with 10,000 accumulated samples is [90 0 30 0 0 90 0]$_s$. The average fitness value for this design from the six GA runs is 0.43454 kg. The deterministic mass is calculated by using the mean value of the uncertain parameters and the value for this configuration is 0.43463 kg. This laminate design has high values of reliability indicating that all probabilistic constraints are satisfied for more than 9,900 samples. The laminate design with the lowest mass appeared as the optimal solution only once in the six runs and the reliability of this design just satisfied the constraint. This laminate configuration with a stacking sequence of [45 90 0 90 0 0]$_s$ was also one of the results obtained via GA with MCS approach. Perhaps, this laminate could be the global minimum here; however, its lack of repeatability indicates that the reliability of the laminate design could be in fact less than the cut-off value (0.99) with a different set of samples.
The optimal laminate designs produced by both approaches include a combination of 0º and 90º plies to improve the effective stiffness in the *x* and *y* directions. The addition of 30º and 60º plies reduces mid-plain strains and curvature in the laminate.

7.3. Parallel Computing Implementations in GA-PBS

A reduction in the computational cost associated with the GA-PBS was successfully achieved with the improved sample accumulation methods. However, the computational time is dependent not only on the complexity of the algorithm, but also on how advantageous the implementation is, and the capacity of the processor. Two relatively obvious different modes exist to introduce parallel processing to this algorithm: performing the fitness analyses of multiple designs in parallel or performing the sampling and constraint evaluations for a given design in parallel.
In the parallel fitness analysis implementation, the population of new individuals is passed to multiple processors. Each processor decodes a single individual and generates samples of uncertain parameters depending on the 'goodness' of the deterministic fitness value of that individual. Furthermore, the expected values of the constraints and fitness function are calculated, and this data is transferred back to the controller. In the parallel sampling implementation, the GA generates a population of individuals and decodes each of them in series. Once the sample size for an individual is determined, the samples are generated in parallel by multiple processors. The probabilistic constraints are also evaluated in parallel mode and this data is transferred back to the controller to calculate the penalty function and the expected fitness value of the design.
An additional confidence level of 99.5% was investigated to compare the two parallel processing implementations. Figure 7 illustrates the computational time of the GA-PBS at five different confidence levels when parallel computation techniques are implemented. Note that the implementation of parallel processing does not have an effect on the quality of solutions or the total number of fitness evaluations. The maximum and minimum run times at each confidence interval for both parallel implementations also appear in Fig. 7.

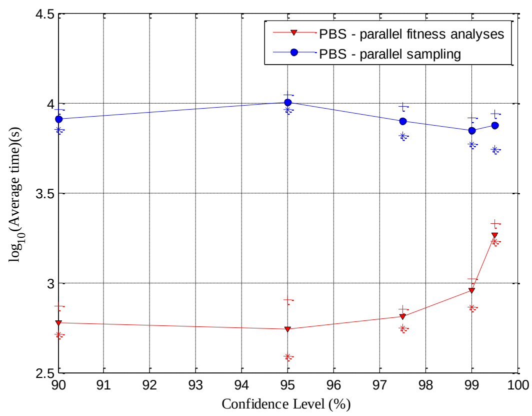

Figure 7: Comparison of average run time using parallel computation techniques

The average run time for the GA-PBS with fitness analyses in parallel follows an increasing trend when increasing the total number of samples. This trend makes sense, because the evaluation of a larger number of samples takes more time. However, this same trend is not valid for the GA-PBS with parallel sampling and constraint evaluation. At a confidence level of 90%, the stopping criterion requires only a total of 100 samples for the best design. In the first few generations, each individual requires only one or two samples. The constraint function evaluations with such a low number of samples do not warrant the use of parallel computation. However, as the number of samples gets larger, the parallel evaluations are more effective in reducing run time. The average run time beyond the confidence level of 95% decreases with increase in sample size. An additional confidence level of 99.5% was investigated to better interpret the trend. There is a slight increase in the average run time at 99.5% confidence level using the parallel sampling implementation; however, the variation in the average run time is seen to be more drastic for the parallel fitness analyses GA-PBS at 99.5% confidence level. Comparing the two slopes, it can be argued that, beyond a certain confidence level requiring a very large number of samples, the GA-PBS with parallel sampling might be more effective than its parallel fitness analyses counterpart i.e., the upward trend for parallel fitness evaluation might eventually cross the curve for parallel sampling.

8. Conclusions

The objectives of this work can be categorized into three parts. The first objective established a probabilistic approach to characterize uncertainty in design factors. Employing an effective and computationally inexpensive sampling process without compromising the solution quality served as the second objective. The third objective attained a significant reduction in run time by implementing parallel computing processes in the optimization scheme.

Results from the GA-MCS runs indicate a high probability of the optimal solution satisfying all constraints for a fixed sample size. However, at a confidence level of 99%, a single GA-MCS run performed 107,520,000 function evaluations, lasting more than 16 hours on average. As an alternative to high computational costs, PBS with improved sample accumulation methods was used to address uncertainties associated with the material properties and the applied loads on the laminate. The results from GA-PBS yielded a laminate constituting 14 plies. The total number of function evaluations performed to generate an optimum solution was reduced by three orders of magnitude through the implementation of GA-PBS, without deterioration of solution quality relative to the designs generated by GA-MCS.

The variation in computational time via the parallel fitness analyses algorithm and the parallel sample evaluation algorithm was compared. Performing the fitness analyses in parallel was more effective for the confidence levels investigated in this study; however, the data could be extrapolated to suggest that, at a much higher confidence level (closer to 100%) requiring a very large sample size, the parallel sample evaluation scheme could be equally effective.

However, the lack of published data in this field prompted a Gaussian distribution assumption for the uncertain parameters. Experimental studies and manufacturers may provide a more reasonable estimate of the distributions. Future research could also include the modeling of interactions and correlations between uncertain parameters to generate the probability distributions. Nonetheless, the results indicate that the GA-PBS with improved sample accumulation methods can be a computationally efficient and reliable baseline approach to discrete optimization under uncertainty.

9. References

[1] M. Cantoni, M. Marseguerra, and E. Zio, Genetic algorithms and Monte Carlo simulation for optimal plant design, *Reliability Engineering & System Safety*, 68.1, 29-38, 2000.

[2] J.A. Vasquez et al, Achieving water quality system reliability using genetic algorithms, *Journal of environmental engineering,* 126.10 , 954-962, 2000.

[3] W.A. Crossley, A Genetic Algorithm with Population-Based Sampling for Optimization under Uncertainty, AIAA 1999 – 1427, *40th Structures, Structural Dynamics and Materials Conference and Exhibit*, April 1999.

[4] R. Hassan, Genetic Algorithm approaches for conceptual design of spacecraft systems including multi-objective optimization and design under uncertainty, School of Aeronautics & Astronautics, West Lafayette, Purdue University, Doctor of Philosophy, 2004.

[5] R.M. Jones, Mechanics of composite materials, Vol. 2. London: Taylor & Francis, 1975.

[6] R. Hill, Mathematical Theory of Plasticity, Oxford University Press, New York, l950.

[7] E.A. Williams and W.A. Crossley, Empirically-derived population size and mutation rate guidelines for a genetic algorithm with uniform crossover, In *Soft computing in engineering design and manufacturing* , Springer London, 163-172, 1998.

[8] R. Hassan, et al, Spacecraft reliability-based design optimization under uncertainty including discrete variables, *Journal of Spacecraft and Rockets*, 45.2, 394-405, 2008.

[9] R. Hassan and W. Crossley, Approach to discrete optimization under uncertainty: The population-based sampling genetic algorithm, *AIAA journal*, 45.11, 2799-2809, 2007.