Coarse Grained Variable-Complexity Modeling for Multidisciplinary Optimization in Aircraft Design

S. Burgee, L. T. Watson
Departments of Computer Science and Mathematics

A. A. Giunta, V. Balabanov, B. Grossman
R. T. Haftka, W. H. Mason
Department of Aerospace & Ocean Engineering

Virginia Polytechnic Institute and State University
Blacksburg, VA 24061-0106

Abstract

A parallel variable-complexity modeling approach, permitting the efficient use of emerging parallel computing in multidisciplinary optimization (MDO) technology, is presented for the particular instance of High Speed Civil Transport (HSCT) design. In this method simple analyses are used to limit the approximation domain. Points are selected within the approximation domain based on the D-optimality criterion through the use of a genetic algorithm. At the D-optimal points, a refined analysis is performed. The optimization code is composed of a sequence of analysis cycles between aerodynamic and structural calculations. Results for coarse grained parallelization of the aerodynamic and structural codes on an Intel Paragon are presented for an example HSCT design problem involving only two variables. The full HSCT design problem employs twenty-eight design variables.

Key words: *multidisciplinary optimization, variable complexity modeling, coarse grained parallelization, high speed civil transport design.*

1: Introduction

The design of modern aerospace vehicles involves multidisciplinary interactions and presents a formidable challenge to the designer in a competitive marketplace. Multidisciplinary optimization (MDO) technology is a major tool for addressing this need, but its current use in the design process is limited by its enormous computational burden. To address this challenge, Unger et al. recently developed a variable-complexity modeling approach, involving the use of refined and computationally expensive models together with simple and computationally inexpensive models [12]. This approach was applied to the combined aerodynamic-structural optimization of transport wings and focuses on the aerodynamic-structural optimization of the High Speed Civil Transport (HSCT) [7].

The goal of this paper is to present a parallel multipoint approximation technique as applied to the aerodynamic and structural design of the HSCT. In this method, the design space is limited through the use of simple, less computationally expensive analyses for the evaluation of both the objective function and constraints for numerous HSCT configurations. Here the simple analyses are used to screen out HSCT designs which are not close to a possible optimum design. A genetic algorithm then is used to determine a set of points within the limited approximation domain which satisfy the D-optimality criterion. For these points, the refined, more computationally expensive analyses are performed to provide a more accurate assessment of the merit of the selected HSCT configurations. Prior to the use of the refined analyses, response surfaces are constructed to provide a smooth approximation to the aerodynamic drag calculations and to the structural weight estimations.

A coarse grained parallelization of both the simple and refined analysis codes used for HSCT design was implemented to reduce the time required to evaluate the numerous HSCT configurations. Results on an Intel Paragon have shown reasonable speedup and efficiency, although further improvements will require a combination of coarse and fine grained parallelism. This paper presents an HSCT design problem involving two variables and the multipoint approximation technique previously used by Giunta et al. [5], initial speedup and efficiency results obtained for the parallelization of both the simple and refined aerodynamic analysis codes, and parallelization results for two structural analysis codes which are under consideration for use in the HSCT design procedure.

2: HSCT configuration and mission

Successful aircraft configuration optimization requires a simple yet meaningful mathematical characterization of the geometry. The model used here completely defines the wing-body-tail configuration, using twenty-eight design variables. The wing planform is described by eight design variables, and the airfoil thickness distribution by an additional five. The engine nacelles move axially with the trailing-edge of the wing, and two parameters define their spanwise locations. The axisymmetric fuselage requires eight parameters to specify both the axial positions and radii of the four fuselage restraint locations. Two additional variables describe the size of the horizontal and vertical stabilizers. Details of the geometry specification appear in [8] and [9]. While the configuration is defined using this set of parameters, the aircraft geometry is actually stored as a discrete numerical description in the Craidon format [3].

The optimization problem is to minimize the takeoff gross weight of an HSCT configuration with a range of 5500 nautical miles and a cruise speed of Mach 2.4 while transporting 251 passengers. For this mission, in addition to the geometric parameters mentioned above, three variables define the idealized cruise mission. One variable is the mission fuel and the other two are the initial cruise altitude and the constant climb rate used in the range calculation. Fifty-five constraints which include performance/aerodynamic constraints and geometric constraints, such as fuselage volume and tail scrape angle, prevent the optimizer from creating physically improbable designs [8].

3: Variable-complexity modeling

The detailed aerodynamic analysis code utilized the Harris program [6] for the supersonic volumetric wave drag, a Mach-box [2] type method for supersonic drag due to lift, and a vortex-lattice program for landing performance. When compared to the computational costs of current computational fluid dynamics analysis techniques, the aerodynamic analysis methods used in this study are relatively inexpensive. However, when implemented in design optimization, where the same calculation may be repeated thousands of times, the cost associated with these techniques quickly becomes substantial. Therefore, a simple, less computationally intensive analysis in a variable-complexity modeling approach [7] is employed. In earlier work, this methodology was used within a sequential approximate optimization technique whereby the overall design process was composed of a sequence of optimization cycles. At the beginning of each cycle, approximations to the wave drag and drag due to lift were constructed using either linear, scaled, or global-local approximations [7]. The scaled approximation method was utilized in this study and employed a constant scaling function, σ, given as $\sigma(x_0) = f_d(x_0)/f_s(x_0)$, where f_d represents a detailed model analysis result, and f_s represents a simple model analysis result, both evaluated at a specified design point, x_0, at the beginning of an optimization cycle. During an optimization cycle, the scaled approximate analysis results, $f(x)$, were calculated as $f(x) = \sigma(x_0) f_s(x)$. Move limits were then imposed on the design variables to avoid large errors.

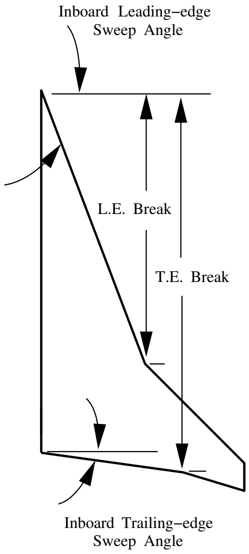

Figure 1. Design variable definition for the example problem.

4: Two design variable problem

A simplified, two variable HSCT design problem previously was considered by Giunta et al. [5]. In this problem the two design variables were the the axial locations for the leading-edge and trailing-edge break points on the HSCT wing planform (Fig. 1). The design variables were varied by $\pm$ ten percent from a baseline HSCT configuration. This corresponded to a variation of the inboard leading-edge sweep angle from $77°$ to $79°$ and a variation of the inboard trailing-edge sweep angle from $-55°$ to $50°$. Minimizing supersonic drag due to lift is the objective function for this design problem.

The design space is discretized using a 21×41 uniform grid, and supersonic drag due to lift is calculated for the HSCT configuration at each point on the mesh. This produces a detailed map of the objective function, where drag due to lift is plotted versus the range of leading and trailing-edge sweep angles (Fig. 2). This demonstrates the jagged surface features which created the multiple spurious local minima found in an earlier study [5]. Similar results are obtained for the volumetric wave drag.

The irregular surface features can be attributed to the techniques used in the supersonic drag due to lift calculations. The methods of Carlson et al. [2] utilize a paneling scheme that is sensitive to planform changes. Thus, slight modifications to the leading and trailing-edge sweep angles along with changes in the location where the Mach angle intersected the leading-edge produced discontinuous variations in the predicted drag. The variations are small enough so that at all points the accuracy of the drag is acceptable. However, the oscillatory behavior creates difficulties for gradient based optimization techniques [5].

5: Multipoint approximation

This work focuses on the use of multipoint approximations to fit response surfaces to the noisy objective function of the example problem. The term *multipoint* refers to the practice of concurrently analyzing many different HSCT configurations. The multipoint analyses yield a database containing both the independent design variables and objective function data. *Response surfaces* are smooth functions which define the relationship between the independent and dependent variables [11].

The goal of multipoint approximation is to model the objective function either on the entire design space or portions of it using smooth response surfaces. Since the topography of a multidimensional objective function generally is unknown and may have many local minima, the smooth functions are selected so that the prominent features of the objective function are retained. Thus, in the optimization process, the region where the global minimum exists may be readily found while spurious local minima may be avoided. Although the drag due to lift calculations are noisy, the variation in the drag is only 2.5 percent. This allows low order response surfaces based on linear and quadratic functions to be used to model the topography of the objective function.

Because of the computational costs associated with evaluating the large number of point designs, the present work focuses on methods to determine the minimum number of calculations required to construct various response surfaces. This work with two design variables will be extended later to design problems involving many variables where computational costs are significant.

Response surface approximations are constructed using both polynomial functions and rational functions. The bilinear tensor product model has the form

$$f(x,y) = (a_1 x + a_2)(a_3 y + a_4).$$

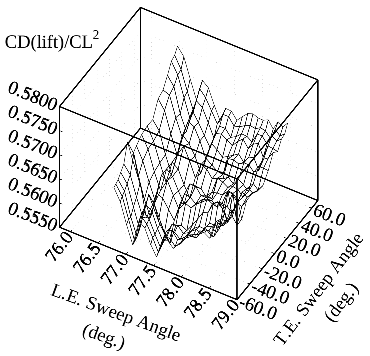

Figure 2. Noisy drag due to lift objective function data in the design space.

Also employed were the quadratic polynomial

$$f(x,y) = a_1 x^2 + a_2 xy + a_3 x + a_4 y^2 + a_5 y + a_6$$

and the biquadratic tensor product

$$f(x,y) = (a_1 x^2 + a_2 x + a_3)(a_4 y^2 + a_5 y + a_6).$$

In addition to these models, the (1,1) rational function

$$f(x,y) = (a_1 x + a_2 y + a_3)/(a_4 x + a_5 y + a_6)$$

was utilized.

The construction and evaluation of a response surface requires several steps. In the first step a number of points were selected using the D-optimal selection technique (discussed below). Supersonic drag due to lift data are then calculated at each of the design points and a response surface is fit to the data. For the bilinear tensor product, quadratic polynomial, and biquadratic tensor product the response surfaces were found by performing a linear least squares fit to the drag due to lift data. For the rational function approximation, a Levenberg-Marquardt nonlinear least squares fitting routine was used. For this problem, the residual was defined as the difference between the objective function values on the computed surface and corresponding model values on the response surface. From the residual data the average error, RMS error, and maximum error of the residual were calculated on a grid of 861 points to evaluate the quality of the response surface fits (Tables 1, 2, 3). Here, m is the number of data points on which the response surface calculation is based. The data in the tables represents the average of three trials for each value of m.

m	Avg. Error	RMS Error	Max. Error
D-optimal Points			
4	0.00413	0.00502	0.01490
5	0.00403	0.00492	0.01468
10	0.00381	0.00470	0.01416
15	0.00409	0.00496	0.01502
20	0.00378	0.00464	0.01414

Table 1. Calculated error for the bilinear tensor product response surface fit.

m	Avg. Error	RMS Error	Max. Error
D-optimal Points			
6	0.00189	0.00247	0.00933
10	0.00215	0.00284	0.01102
20	0.00192	0.00252	0.01034
30	0.00192	0.00234	0.00855
40	0.00193	0.00244	0.00892
50	0.00211	0.00258	0.00779

Table 2. Calculated error for the quadratic polynomial response surface fit.

m	Avg. Error	RMS Error	Max. Error
D-optimal Points			
9	0.00199	0.00259	0.01044
10	0.00201	0.00263	0.00996
20	0.00202	0.00251	0.00774
30	0.00199	0.00240	0.00810
40	0.00183	0.00232	0.00823
50	0.00184	0.00233	0.00893
60	0.00189	0.00234	0.00782
70	0.00188	0.00240	0.00875
80	0.00182	0.00231	0.00907
90	0.00181	0.00231	0.00900
100	0.00180	0.00226	0.00855

Table 3. Calculated error for the biquadratic tensor product response surface fit.

6: Approximation domain definition

The approximation domain is centered around a nominal design point, usually obtained from the previous approximate optimization cycle. The approximation domain is constrained by requiring that the approximation to each constraint (fifty-five in this case) not be violated by more than 10 percent. The domain is bound by additional limits on the values of the constraints and the objective function obtained from an inexpensive approximation. When the simple analysis is several orders of magnitude cheaper than the refined analysis, then the simple analysis defines directly the approximation domain. Otherwise, a linear approximation to the simple analysis is used. Constraints can bound the approximation domain by requiring that *the approximation to the constraint not be violated by more than a certain margin*, $\pm 10\%$ in this case. Similarly, the objective function can bound the approximation domain by *limiting the change in the objective function.*

7: D-optimal criterion

One method for selecting points to be used in response surface construction is to choose m points from a set of $l > m$ candidate points existing in the design space that will maximize the determinant $|\mathbf{X}^T\mathbf{X}|$, known as the D-optimality criterion [1]. The model here is $Y = \mathbf{X}c$, where Y is the $m \times 1$ vector of objective function values, c is the $k \times 1$ vector of coefficients to be estimated, and $\mathbf{X}$ is an $m \times k$ matrix of constants having rank k.

Create a set of $l > m$ points by forming a grid over the design space. Points of intersecting grid lines become candidate points. Conceivably, one could consider each of the $\binom{l}{m}$ combinations of m points from the set of l candidate points, evaluate $|\mathbf{X}^T\mathbf{X}|$, and identify the set with the largest determinant. This nonconvex combinatorial optimization problem is difficult because the discrete feasible set is huge, and $|\mathbf{X}^T\mathbf{X}|$ may have local maxima. For lack of a better alternative a genetic algorithm (GA) is used to maximize $|\mathbf{X}^T\mathbf{X}|$.

8: Genetic algorithm for D-optimality

The genetic algorithm (GA) employed here is roughly modeled after the algorithm described in the paper by Furuya and Haftka [4]. The intent of the GA is to select the group of m points that satisfy D-optimality for the purpose of constructing a least squares approximation to the objective function. Any group of m distinct points among the l points is called a candidate. To start the algorithm, an initial population of b candidates is created. Each candidate is formed by randomly selecting m distinct points from the set of l points spanning the design space. A genetic string, made from the coordinates of the m points, is used to describe the candidate.

The candidates are selected as parents for breeding based on the fitness $|\mathbf{X}^T\mathbf{X}|$. The rank fitness is defined as $b + 1 - r$, where r is the rank of the candidate in the population in terms of the value of $|\mathbf{X}^T\mathbf{X}|$. After two parents have been selected, the breeding process begins. A random integer j between one and $G - 1$ (the number of genes-1) is generated. The child is constructed from the first j genes of the first parent and the last $G-j$ genes from the second parent. Each child has a 15% chance of mutating, whereby a randomly selected gene is changed. Finally the child is checked for uniqueness and duplicate points. If it is a duplicate or has duplicate points, it is destroyed.

After $b-1$ children are created, the parent generation is replaced, retaining only the best parent for the next generation. After GA termination the candidate with the highest $|\mathbf{X}^T\mathbf{X}|$ is used for forming the response surface. There is no guarantee that the output of the genetic algorithm is truly D-optimal. However, known D-optimal sets of points for test cases have been recovered using the GA.

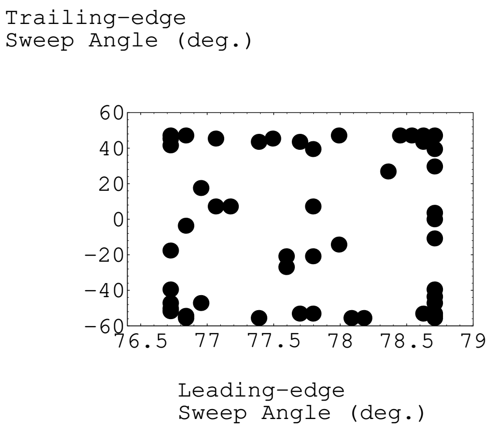

Figure 3. Fifty D-optimal design points chosen by GA for a quadratic response surface fit.

9: D-optimal points in the example

Our example problem required the selection of m D-optimal points from the set of 861 drag due to lift data points. Once the best estimate of a set of D-optimal points was obtained, the appropriate response surface functions were fit to the data. For the GA used in this problem, the population size for each generation was fixed at ten and the number of generations varied from 2000 to 25000. When the best $|\mathbf{X}^T\mathbf{X}|$ remained unchanged for approximately 1000 generations, the breeding process was terminated. Figure 3 shows an example set of fifty points obtained by the GA for a quadratic polynomial response surface model. Note that the points lie predominately along the perimeter of the design space. This is expected since points near the perimeter space have the strongest influence on $|\mathbf{X}^T\mathbf{X}|$.

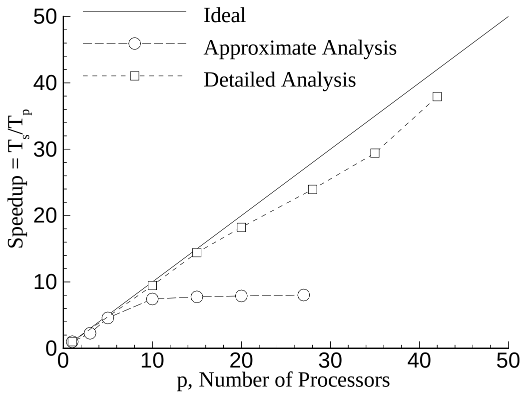

Figure 4. Speedup obtained from parallelization of aerodynamic code.

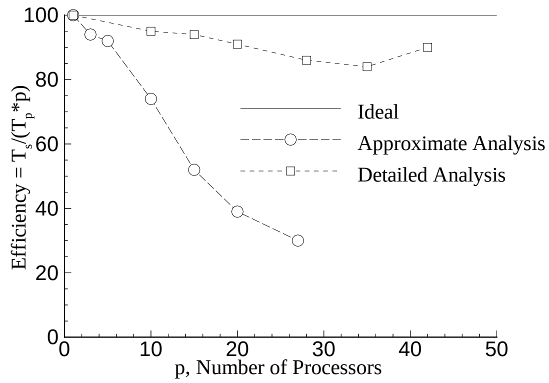

Figure 5. Efficiency obtained from parallelization of aerodynamic code.

10: Parallelization of aerodynamic code

The aerodynamic code computes either detailed or approximate analyses at the selected points in the approximation domain. These analyses are then used to compute a least squares approximation (the response surface) to some criterion function (e.g., drag), which is then optimized over the approximation domain. A similar process is done for the structural design component of the HSCT design problem.

The coarse grained parallelization of the aerodynamic analysis modules within the full HSCT analysis code makes use of a master-slave paradigm on the Paragon whereby one designated master node controls the data transfer and file input/output (I/O) of the remaining slave nodes. This coarse grained approach is used for the numerous independent analyses required for response surface construction.

To initiate the parallel multipoint analyses, a group of predetermined analysis points is input to the master node. The master node then computes the subset of the points which each slave node will analyze and sends that information to the appropriate slave. Both the master and slave nodes then analyze their respective subsets of the selected points and store the results in an array local to each node. When each slave has finished its portion of the analyses, it sends the array of analysis values to the master node for output.

Parallel efficiencies approach ninety percent for the detailed aerodynamic analysis methods, but the efficiencies for the approximate aerodynamic analysis methods have been unreasonable. When the number of nodes is small (less than 10), reasonable efficiency was obtained (above eighty percent), but when the number of nodes is increased, efficiency becomes drastically reduced (near thirty percent). The low efficiency for the approximate aerodynamic analysis methods is a result of the large amount of intermediate I/O occurring during the analysis of each point. A serial portion of code at the beginning and end of the aerodynamic code also contributed to the unreasonably low speedups and efficiencies for the approximate aerodynamic analysis code.

11: Structural optimization

Structural optimization acutely requires parallelization because it is repeated many times within the overall HSCT optimization. The first step in the application of parallel computing to the HSCT structural analysis was to choose a finite element program that could be efficiently run on the Paragon. Two software packages were considered: Genesis and Maestro.

Genesis is a finite element structural optimization code. An attempt was made to develop a coarse grained parallel version of this code. However, Genesis relies heavily on disk I/O which limits its parallel performance on the Paragon (Figs. 6, 7). Thus, it was not possible to obtain speedup beyond a factor of 2.3 even using a large number of nodes.

Maestro is a computer program for optimum design of large complex thin-walled structures, developed by Owen Hughes of Virginia Tech, and is extensively used for ship design. It has the following features:

i. a rapid design-oriented finite element analysis for the complete structure,
ii. an explicit evaluation of all limit states both at member and multimember levels,
iii. formulation of the linearized state equations, which utilize partial safety factors to achieve a consistent degree of structural reliability,
iv. multiobjective optimization based on any combination of individual measures of merit.

Like Genesis, Maestro suffers from excessive disk I/O. However, due to the availability of the source code, a coarse grained parallel version of Maestro was developed by replacing disk I/O with memory references. As shown in Figures 6 and 7, the initial parallel performance of Maestro on the Paragon is promising.

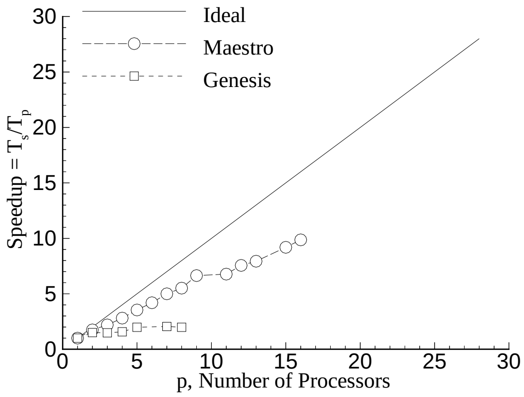

Figure 6. Speedup obtained from parallelization of FEM codes.

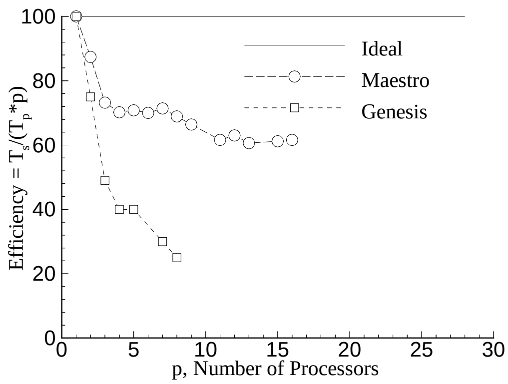

Figure 7. Efficiency obtained from parallelization of FEM codes.

The results from the coarse grained parallel versions of Genesis and Maestro were compared to results obtained from a serial version of the NASTRAN code, the de facto standard in structural analysis. To facilitate this comparison, several geometry translators were developed to allow the HSCT geometry in the Craidon format to be easily transferred to Genesis, Maestro, or NASTRAN.

12: Future research directions

We have completed a design problem involving two design variables in which the response surface approximation methods were demonstrated. Currently we are investigating design problems with four to eight design variables to further validate this methodology. After we have used the inexpensive analysis methods in our variable-complexity modeling approach to identify the feasible regions of the design space, we then apply principal factor analysis [10] to identify the design variables which have the most impact on the overall HSCT design. Those variables having the most effect will be modeled using higher order response surface functions while those having less effect will be modeled using lower order functions. Eventually we will apply this variable-complexity response surface design methodology to the full HSCT design problem which involves twenty-eight design variables. We have begun initial evaluation of an Euler/Navier-Stokes equations solver for use with our HSCT design methodology.

We plan to apply the coarse grained parallel version of Maestro to variable-complexity structural optimization, and to integrate Maestro into the HSCT design process. In addition, we plan to develop a fine grained parallel version of the finite element program, used for two-level structural optimization and for the same tasks as the coarse grained parallel version. Future work on parallel computation includes attaining more reasonable speedups and efficiency for the approximate aerodynamic analysis code for more than ten processors. Our work will focus on using both coarse and fine grained parallelization of the aerodynamic codes to produce large speedups and high efficiency.

13: Conclusion

A parallel variable-complexity modeling approach for the particular instance of High Speed Civil Transport design on an Intel Paragon has been presented in this paper. Beyond the conceptual question of how best to accomplish MDO, the practical usage of MDO requires the integration of heterogeneous, mismatched disciplinary codes. It is interesting that many of the features of existing software artifacts that make MDO difficult are precisely the same features that make parallelization difficult. These features include

- nonstandard programming constructs,
- use of disk as extended memory,
- dynamic creation of noncontiguous data structures,
- interactive code ill-suited for inclusion in a larger code,
- mismatched data structures for subtasks.

Results from the coarse grained parallelization of the detailed aerodynamic and structural codes have shown reasonable speedups and efficiencies. Improvements need to be made on the speedup and efficiency results for the approximate aerodynamic code. There are two main factors limiting the performance of the coarse grained parallelization of the approximate aerodynamic code: inherently serial portions of the code, and necessary intermediate I/O. Converting the necessary intermediate I/O to memory references still may not provide reasonable speedup and efficiency results for the approximate aerodynamic code, because of the inherently serial portions of the coarse grained code. Therefore, the best that can be done is to maximize the speedup of the parallelizable portions of the coarse grained code. A complete reorganization of the computational algorithms with fine grained parallelism would be required to reduce the serial portions of the approximate aerodynamic code.

Acknowledgement:

This research was supported by NASA Grant NAG-1-1562 and DOE grant DE-FG05-88ER25068/A004.

References:

[1] Box, M. J. and Draper, N. R., *Factorial Designs, the $|X^T X|$ Criterion, and Some Related Matters*, Technometrics, vol. 13, No. 4, 1971, pp. 731–742.

[2] Carlson, H. W., and Mack, R. J., *Estimation of Leading-Edge Thrust for Supersonic Wings of Arbitrary Planforms*, NASA TP-1270, 1978.

[3] Craidon, C. B., *Description of a Digital Computer Program for Airplane Configuration Plots*, NASA TM X-2074, 1970.

[4] Furuya, H. and Haftka, R. T., *Locating Actuators for Vibration Suppression on Space Trusses by Genetic Algorithms*, ASME Winter Annual Meeting, 1993.

[5] Giunta, A. A., Dudley, J. M., Narducci, R., Grossman, B., Haftka, R. T., Mason, W. H., and Watson, L. T., *Noisy Aerodynamic Response and Smooth Approximations in HSCT Design*, AIAA Paper 94-4376, Sept. 1994.

[6] Harris, R. V., Jr., *An Analysis and Correlation of Aircraft Wave Drag*, NASA TM X-947, 1964.

[7] Hutchison, M. G., Huang, X.-M., Mason, W. H., Haftka, R. T., and Grossman, B., *Variable-Complexity Aerodynamic-Structural Design of a High-Speed Civil Transport Wing*, AIAA Paper 92-4695, Sept. 1992.

[8] Hutchinson, M. G., *Multidisciplinary Optimization of High Speed Civil Transport Configurations Using Variable-Complexity Modeling*, Ph.D. Dissertation, VPI&SU, March 1993.

[9] Hutchinson, M. G., Unger, E. R., Mason, W. H., Grossman, B., and Haftka, R. T., *Aerodynamic Optimization of an HSCT Wing Using Variable-Complexity Modeling*, AIAA Paper 93-0101, Jan. 1993.

[10] Khuri, A. I., and Cornell, J. A., *Response Surfaces: Designs and Analyses*, Marcel Dekker, Inc., New York, N. Y., 1987.

[11] Lawley, D. N., and Maxwell, A. E., *Factor Analysis as a Statistical Method*, American Elsevier Publishing Co., New York, N. Y., 1971, pp. 15-18.

[12] Unger, E. R., Hutchison, M. G., Rais-Rohani, M., Haftka, R. T., and Grossman, B., *Variable-Complexity Design of a Transport Wing*, Int. J. Sys. Automation: Res. and Appl., to appear.