

Comparison of Alternative Strategies for Multilevel Optimization of Hierarchical Systems

Saber DorMohammadi, Masoud Rais-Rohani

Center for Advanced Vehicular Systems, Mississippi State University, Starkville, USA
Email: Masoud@ae.msstate.edu

Received July 10, 2012; revised August 10, 2012; accepted August 17, 2012

ABSTRACT

The augmented Lagrangian penalty formulation and four different coordination strategies are used to examine the numerical behavior of Analytical Target Cascading (ATC) for multilevel optimization of hierarchical systems. The coordination strategies considered include augmented Lagrangian using the method of multipliers and alternating direction method of multipliers, diagonal quadratic approximation, and truncated diagonal quadratic approximation. Properties examined include computational cost and solution accuracy based on the selected values for the different parameters that appear in each formulation. The different strategies are implemented using two- and three-level decomposed example problems. While the results show the interaction between the selected ATC formulation and the values of associated parameters, they clearly highlight the impact they could have on both the solution accuracy and computational cost.

Keywords: Analytical Target Cascading; Multilevel Design Optimization; Augmented Lagrangian; Method of Multipliers

1. Introduction

A complex optimization problem may be decomposed into two or more subsystems with partitioned design variables and separate objective functions and design constraints. The layered architecture of hierarchically decomposed multilevel systems is illustrated by the example in **Figure 1**; the hierarchy can be expanded to include several levels with each containing multiple elements. Because the number of design variables in each element represents a fraction of the total set, dimensionality of each element optimization problem is reduced. Decomposed optimization problems require a coordination strategy that in the ensuing iterative solution process ensures satisfaction of system-level design criteria and proper convergence to an optimum design.

Analytical target cascading (ATC) [1,2] was developed for systems such as that shown in **Figure 1**. In the initial formulation of ATC [2-5], deviation tolerances

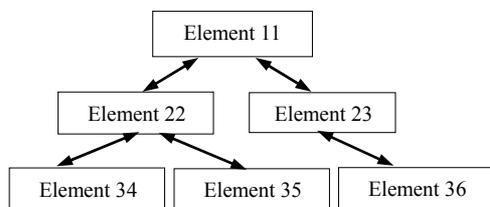


Figure 1. An illustrative model of a hierarchically decomposed multilevel system.

were defined for the responses and targets as well as the linking (or shared design) variables. The multilevel optimization problem was solved while minimizing the deviation tolerances and satisfying the design constraints.

ATC solution has been shown to converge to a point that satisfies the necessary optimality conditions of the original design optimization problem [6]. Using a formulation of ATC with similarities to that in [3], the inequality constraints on deviation tolerances were brought into the objective function to form an augmented objective function; this formulation included the addition of weight factors to the deviation tolerances.

The scaled tolerance formulation [3] was used in [7] to investigate the numerical behavior of the ATC methodology and the local convergence properties of different coordination strategies. They examined the effects of linking variables, subproblem solution accuracy, and the number of significant digits on numerical stability.

The commonly used ATC formulations are based on quadratic penalty (QP) functions [7-10]. Numerical experiments with these formulations show significant computational effort to obtain accurate solutions. The QP functions minimize the consistency constraints (equality or inequality) to force targets and responses to match. Ideally, these consistency constraints have to be relaxed, allowing inconsistencies between targets and responses that are gradually eliminated in the iterative solution process. For the QP function, in general, large weight

factors are required to find accurate solutions [11]. Due to lack of a mathematical relationship between weight factors and solution accuracy, the weight factors are given arbitrarily large values that may cause computational difficulties [8,10].

An iterative method was presented in [8] for finding the minimal penalty weight factors that provide converged solutions within user-specified inconsistency tolerances, and its effectiveness was demonstrated with several examples. This method contains an inner and an outer loop. The inner loop solves the decomposed ATC problem with a coordination scheme. The outer loop updates the penalty weight factors based on information obtained from the inner loop. The iterative method calculates the Lagrange multipliers and derivatives of the response function to update the weight factors.

In the separable ordinary Lagrangian (OL) approach, a large-scale convex nonlinear programming problem is formulated and decomposed using the ATC [12]. By combining the classical Lagrangian duality and the augmented Lagrangian duality, a simple method was proposed in [13] for decomposition without imposing restrictive conditions to alleviate the difficulty of convexity requirement. The modified Lagrangian dual formulation and coordination enhances the ATC performance [14] over those proposed earlier in the literature.

ATC problem relaxation with an augmented Lagrangian penalty (ALP) function using the method of multipliers (AL) and the alternating direction method of multipliers (AL-AD) was proposed and investigated in [15]. By means of the ALP relaxation, ill-conditioning is reduced for the inner loop because accurate solutions can be obtained for smaller weight factors. This formulation was later adopted in [16] that used Diagonal Quadratic Approximation (DQA) and Truncated DQA (TDQA) for parallelization of ATC. Similarly, the ALP formulation was also applied in [17], but three different updating methods were used in the outer loop.

In this paper, the (ALP) function using the method of multipliers with four different coordination strategies (*i.e.*, AL, AL-AD, DQA, and TDQA) is used to study the numerical behavior of ATC. Moreover, the role of two penalty parameters that can have large influence on solution accuracy and computational cost is investigated. The effects of the penalty parameter updating coefficient in the outer loop and the initial guessed values for the decision variables to start the multilevel optimization process are examined by solving three example problems.

2. Overview of ATC

For a decomposed system with N levels and M elements, as shown in **Figure 2**, the subscripts ij denote the j th element in the i th level [15]. The vector of local variables is denoted by \mathbf{x}_{ij} with \mathbf{t}_{ij} is the vector of target variables

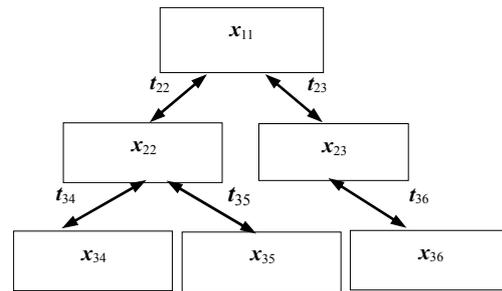


Figure 2. Variable allocation in a hierarchical system.

shared by element ij and its parent at level $i - 1$; E_i is the set of elements at level i (e.g., $E_3 = \{4, 5, 6\}$ in **Figure 2**); $D_{ij} = \{k_1, \dots, k_{D_{ij}}\}$ is the set of children of element ij (e.g., $D_{22} = \{4, 5\}$); f_{ij} is the local objective; \mathbf{g}_{ij} is the vector of local inequality constraints; and \mathbf{h}_{ij} is the vector of local equality constraints. Hence, an all-in-one (AIO) problem of such a system is defined as

$$\begin{aligned} \min_{\mathbf{x}_{ij}, \mathbf{t}_{ij}} \sum_{i=1}^N \sum_{j \in E_i} f_{ij} \left(\mathbf{x}_{ij}, \mathbf{t}_{ij}, \mathbf{t}_{(i+1)k_1}, \dots, \mathbf{t}_{(i+1)k_{D_{ij}}} \right) \\ \text{s.t. } \mathbf{g}_{ij} \left(\mathbf{x}_{ij}, \mathbf{t}_{ij}, \mathbf{t}_{(i+1)k_1}, \dots, \mathbf{t}_{(i+1)k_{D_{ij}}} \right) \leq 0 \quad (1) \\ \mathbf{h}_{ij} \left(\mathbf{x}_{ij}, \mathbf{t}_{ij}, \mathbf{t}_{(i+1)k_1}, \dots, \mathbf{t}_{(i+1)k_{D_{ij}}} \right) = 0 \\ \forall j \in E_i, i = 1, \dots, N \end{aligned}$$

In the ATC formulation adopted from [15], response copies \mathbf{r}_{ij} are introduced to make the objective function and constraints separable, which leads to the addition of consistency constraints expressed as $\mathbf{c}_{ij} = \mathbf{t}_{ij} - \mathbf{r}_{ij} = 0$, where \mathbf{c}_{ij} is a measure of inconsistency between the targets and the corresponding responses in element ij . Moreover, the objective function is augmented by the addition of a penalty term π that leads to the relaxed form of the AIO problem formulated as

$$\begin{aligned} \min_{\bar{\mathbf{x}}_1, \dots, \bar{\mathbf{x}}_{NM}} \sum_{i=1}^N \sum_{j \in E_i} f_{ij} \left(\bar{\mathbf{x}}_{ij} \right) + \pi \left(\mathbf{c}_{ij} \right) \\ \text{s.t. } \mathbf{g}_{ij} \bar{\mathbf{x}} \left(\bar{\mathbf{x}}_{ij} \right) \leq 0 \\ \mathbf{h}_{ij} \left(\bar{\mathbf{x}}_{ij} \right) = 0 \quad (2) \\ \mathbf{c}_{ij} = \mathbf{t}_{ij} - \mathbf{r}_{ij} = 0 \end{aligned}$$

with

$$\bar{\mathbf{x}}_{ij} = \left[\mathbf{x}_{ij}, \mathbf{r}_{ij}, \mathbf{t}_{ij}, \mathbf{t}_{(i+1)k_1}, \dots, \mathbf{t}_{(i+1)k_{D_{ij}}} \right]$$

$$\forall j \in E_i, i = 1, \dots, N$$

where $\mathbf{c} = [\mathbf{c}_{22}, \dots, \mathbf{c}_{NM}]$ in the hierarchy.

Now, the relaxed AIO problem in Equation (2) can be decomposed into separate subproblems (e.g., P_{ij} for element ij) involving only a subset of decision variables $\bar{\mathbf{x}}_{ij}$ given by

$$\begin{aligned} \min_{\bar{x}_{ij}} & f_{ij}(\bar{x}_{ij}) + \pi(\mathbf{c}_{ij}) \\ \text{s.t.} & \mathbf{g}_{ij}(\bar{x}_{ij}) \leq 0 \\ & \mathbf{h}_{ij}(\bar{x}_{ij}) = 0 \end{aligned} \quad (3)$$

with

$$\bar{x}_{ij} = \left[\mathbf{x}_{ij}, \mathbf{r}_{ij}, \mathbf{t}_{(i+1)k_1}, \dots, \mathbf{t}_{(i+1)k_{D_{ij}}} \right]$$

In QP, OL, and ALP, the penalty term takes the form

$$\pi_Q(\mathbf{c}_{ij}) = \|\mathbf{w}_{ij} \circ \mathbf{c}_{ij}\|_2^2 \quad (4)$$

$$\pi_L(\mathbf{c}_{ij}) = \lambda_{ij}^T \mathbf{c}_{ij} \quad (5)$$

$$\pi_{AL}(\mathbf{c}_{ij}) = \lambda_{ij}^T \mathbf{c}_{ij} + \|\mathbf{w}_{ij} \circ \mathbf{c}_{ij}\|_2^2 \quad (6)$$

The ALP method contains two loops. In the inner loop, the decomposed ATC problem is solved for fixed penalty parameters (λ and w) whereas in the outer loop, an algorithm is applied to update both λ and w as

$$\lambda^{(k+1)} = \lambda^{(k)} + 2w^{(k)} \circ w^{(k)} \circ c^{(k)} \quad (7)$$

$$w^{(k+1)} = \beta w^{(k)} \quad (8)$$

where the penalty parameter updating coefficient β is required to be ≥ 1 for convex objective functions [15].

The double-loop approach in AL avoids setting arbitrarily large weight factors that can often cause ill-conditioning in the solution. The weight factors are updated using the information obtained from the inner loop. Whereas the inner loop is very computationally expensive, the outer loop is very inexpensive. It has been shown in the literature that the AL method can significantly reduce the computational cost of solving a problem with ATC without loss of accuracy.

2.1. Alternative Coordination Strategies

For the ALP formulation, the four alternative coordination strategies are described by the algorithms outlined in Figures 3 and 4.

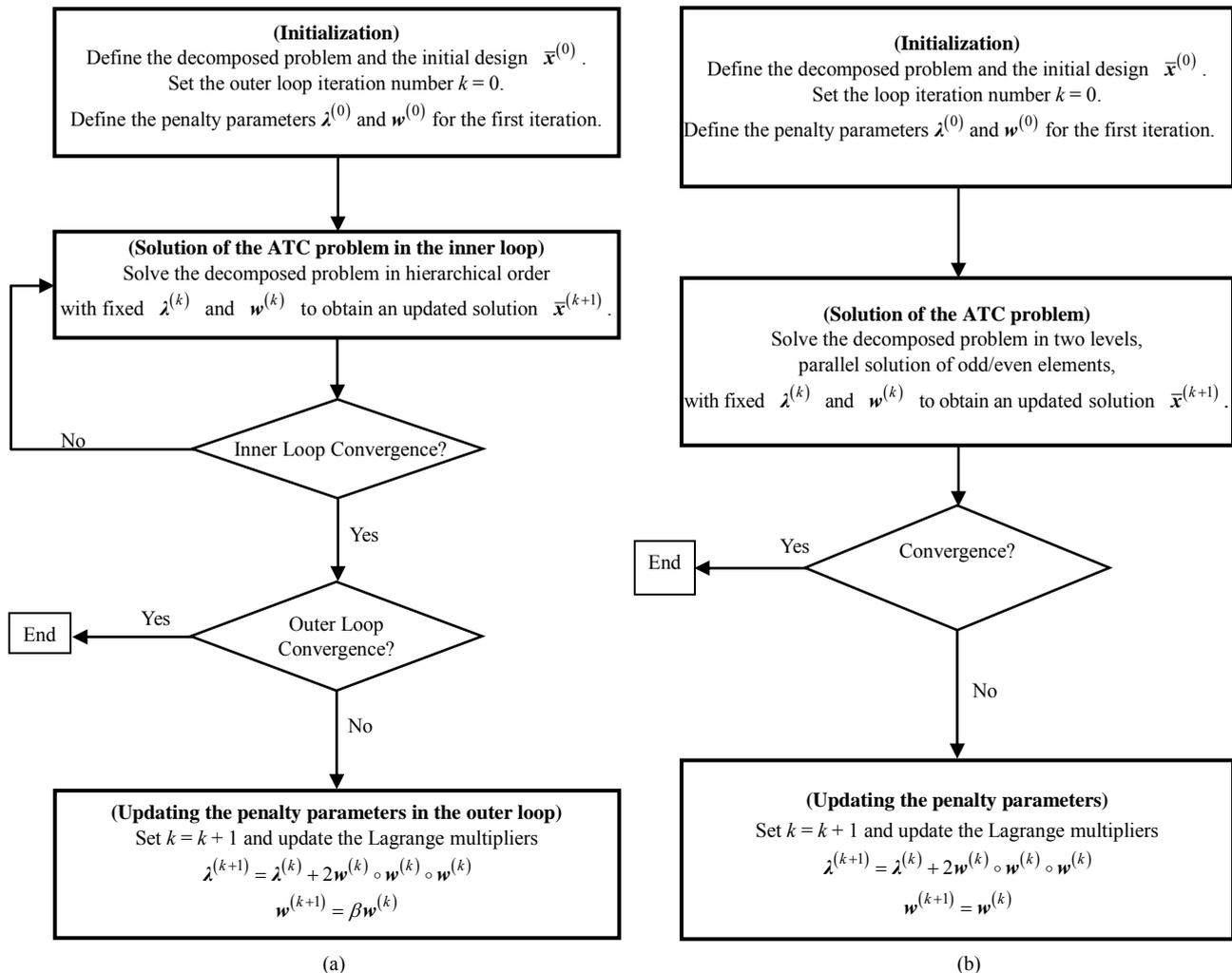


Figure 3. Flowcharts of (a) AL and (b) AL-AD algorithms.

For AL and AL-AD in **Figure 3**, the outer-loop convergence criterion is satisfied when the reduction of inconsistencies in two successive solutions is sufficiently small (i.e., $\|c^{(k)} - c^{(k-1)}\| < \tau$, where k denotes the outer loop counter and τ is a user-defined termination tolerance). The inner loop convergence criterion is reached when the difference in the objective function values in two consecutive inner loop iterations is less than $\tau_{ATC} = \tau/10$.

In the DQA and TDQA algorithms in **Figure 4**, the convergence criteria are defined as

$$\max \left(\left\| \mathbf{t}^{(k+1,s+1)} - \mathbf{t}^{(k+1,s)} \right\|, \left\| \mathbf{r}^{(k+1,s+1)} - \mathbf{r}^{(k+1,s)} \right\| \right) \leq \sigma_{in}$$

$$\max \left(\left\| \mathbf{t}^{(k+1)} - \mathbf{t}^{(k)} \right\|, \left\| \mathbf{r}^{(k+1)} - \mathbf{r}^{(k)} \right\| \right) \leq \sigma_{out}$$

where σ_{in} and σ_{out} are the inner and outer loop termination tolerances with $\sigma_{in} = \sigma_{out}/10$ and $\sigma_{out} = \tau$.

2.2. Illustrative Example Problems

The effect of β on the accuracy and computational cost has not been addressed in the literature. Although it has

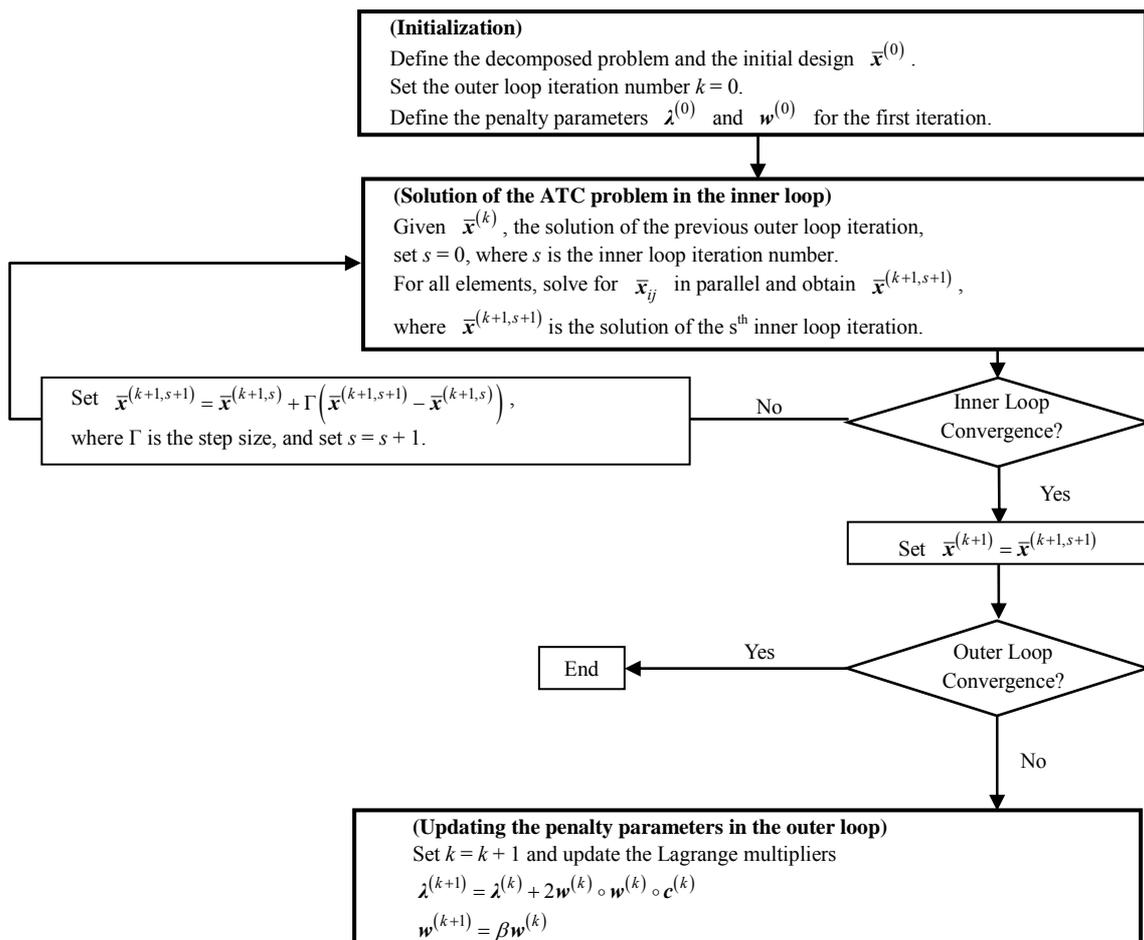
been mentioned that β can take a wide range of values, it is unclear what value must be chosen with respect to the desired levels of accuracy and computational cost as well as the selected ATC solution methodology and coordination strategy. Furthermore, since in ATC the initial values for response/target and linking variables are selected at random, it is unclear what effects these values would have on the ATC results.

To examine these effects, three different example problems are solved using the four different methods of ATC described in the previous section. For each method, the solution starts from different initial guessed values (IGV) that correspond to different randomly selected design points relative to the optimum point. The solution is repeated for 20 different values of β and every IGV.

Two performance metrics are considered: the computational cost that is captured by the number of function evaluations, and the error, which is defined as

$$e = \left\| \mathbf{x}^* - \mathbf{x}^{ATC} \right\|_{\infty} \tag{9}$$

where \mathbf{x}^* is the exact optimum design point and \mathbf{x}^{ATC} is the solution found by ATC. All of the ATC formulations



(a)

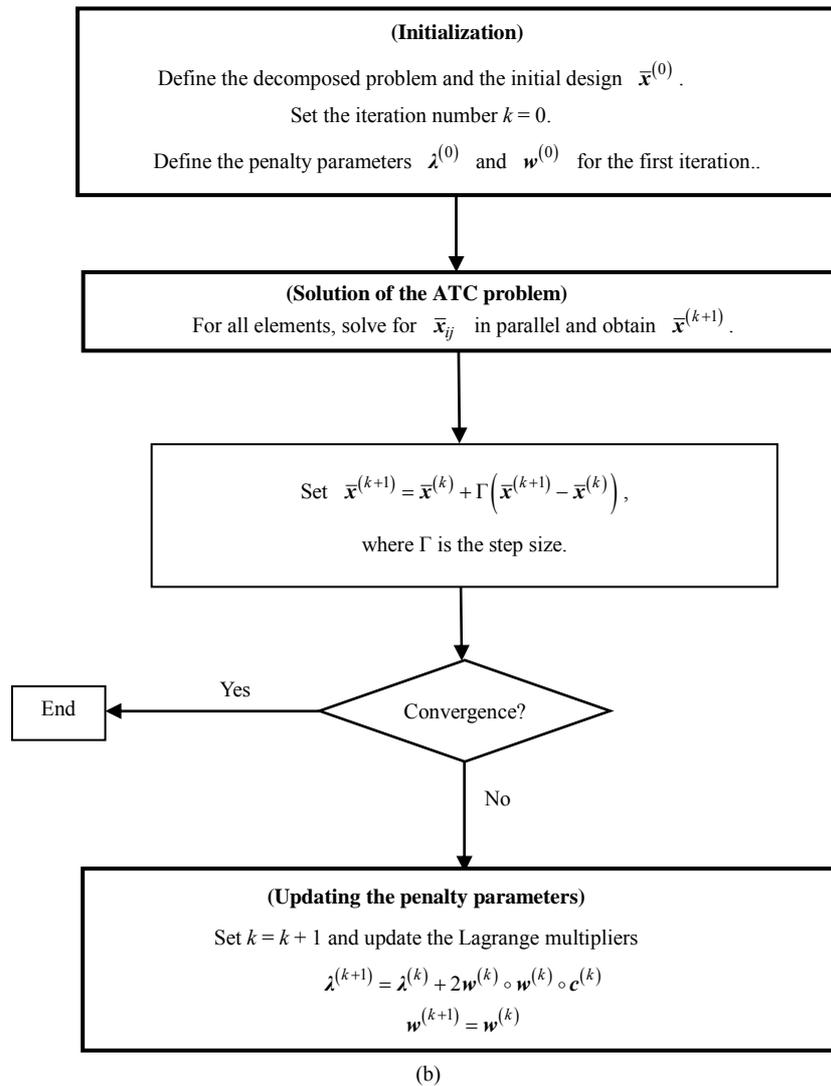


Figure 4. Flowcharts of (a) DQA and (b) TDQA algorithms.

cited were developed into separate MATLAB codes and used to solve the following example problems.

Problem 1: This is a 7-variable geometric programming problem with the AIO formulation expressed as

$$\begin{aligned} \min_{x_1, \dots, x_7} f &= x_1^2 + x_2^2 \\ \text{s.t. } g_1 &= \frac{x_3^{-2} + x_4^2}{x_5^2} - 1 \leq 0 \\ g_2 &= \frac{x_5^2 + x_6^{-2}}{x_7^2} - 1 \leq 0 \\ h_1 &= x_1^2 - x_3^2 - x_4^{-2} - x_5^2 = 0 \\ h_2 &= x_2^2 - x_5^2 - x_6^2 - x_7^2 = 0 \\ x_1, x_2, \dots, x_7 &\geq 0 \end{aligned}$$

where the point of optimum is at $\mathbf{x}^* = [2.15, 2.06, 1.32,$

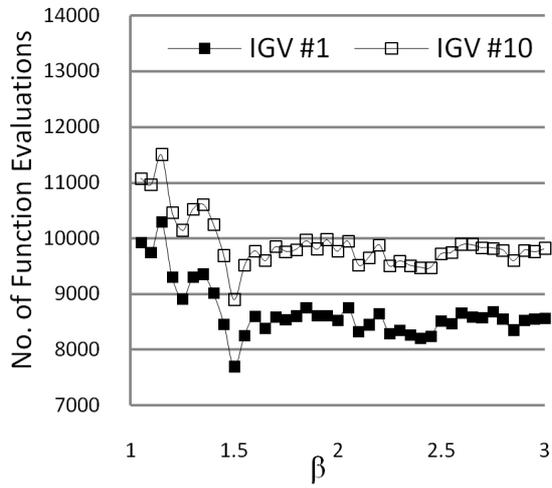
$0.76, 1.07, 1.0, 1.47]$ with all four constraints active.

This problem is decomposed into a two-level hierarchy [10] with a single element at the top level and another element at the bottom level. Local variables in the top element are x_1, x_3 and x_4 along with $f_1 = x_1^2$ as the objective function subject to the inequality constraint g_1 and equality constraint h_1 . Variables x_2, x_6 and x_7 are the local design variables for the bottom element with the objective function $f_2 = x_2^2$ and constraints g_2 and h_2 . The response/target variable for the two elements is x_5 .

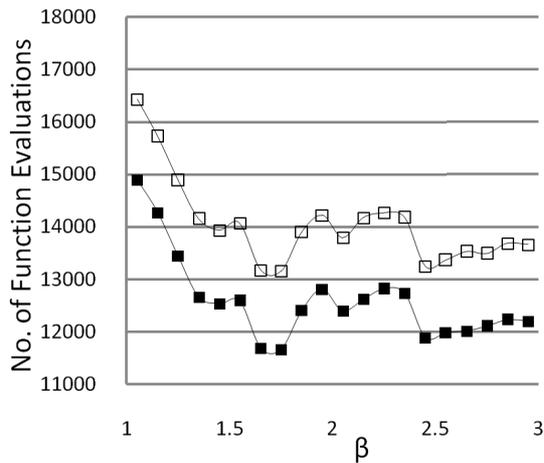
The initial values for the penalty parameters are defined as $\lambda^{(0)} = 0$ and $w^{(0)} = 1$. The starting design point is $\mathbf{x}^{(0)} = [3, 3, 3, 3, 3, 3, 3]$ for all the formulations. The ten initial guessed values (IGV), i.e., IGV #1, \dots #10 for x_5 are chosen as $\{0, 2, 4, 6, 8, 10, 20, 40, 70, 100\}$. For AL and DQA, β is given different values in the range of $\{1.1, 1.2, 1.3, \dots, 3.0\}$, whereas for AL-AD and TDQA, $\beta = 1$. The IGV for x_5 and β are chosen arbitrarily to

simply diversify the iterative solution process. The termination tolerance is chosen as $\tau = \{10^{-2}, 10^{-3}, 10^{-4}\}$.

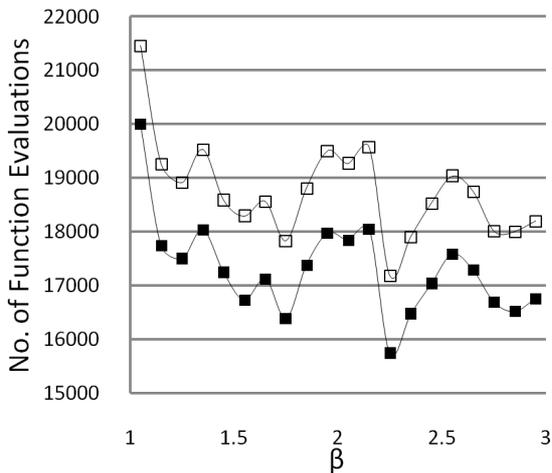
Figures 5 and 6 show the plots of function



(a)



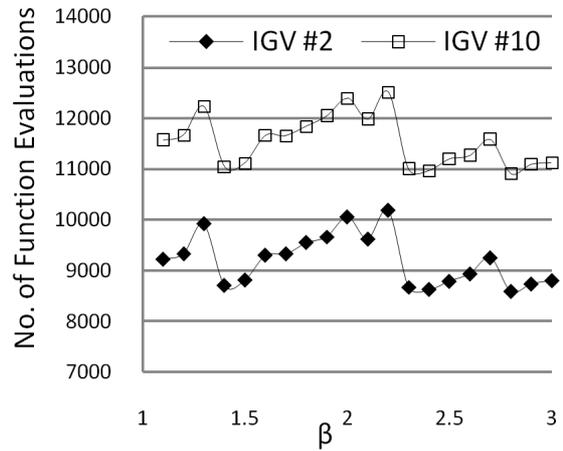
(b)



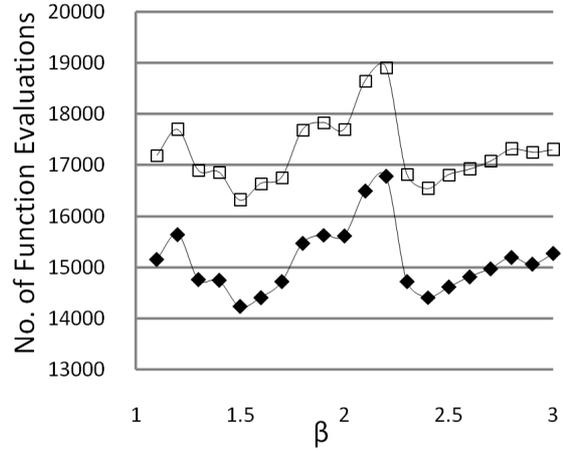
(c)

Figure 5. Cost trends for AL-based solution of Problem 1 using (a) $\tau = 10^{-2}$, (b) $\tau = 10^{-3}$, and (c) $\tau = 10^{-4}$.

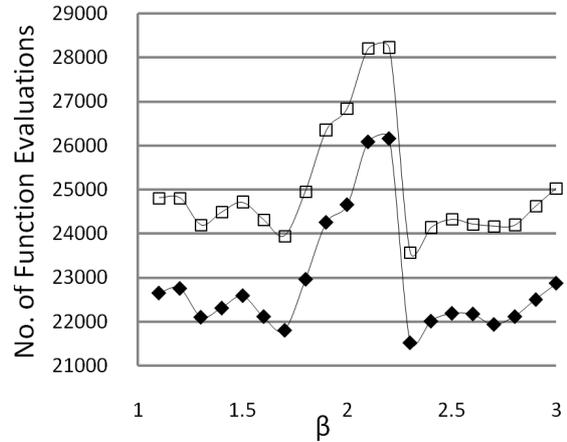
evaluations number (cost) versus β for AL and DQA, respectively, using different IGV for x_5 . These figures show that the cost is affected by the choice of β . The optimum β value to minimize cost depends on the termination tolerance used, but it appears to be near 1.5



(a)



(b)



(c)

Figure 6. Cost trends for DQA-based solution of Problem 1 using (a) $\tau = 10^{-2}$, (b) $\tau = 10^{-3}$, and (c) $\tau = 10^{-4}$.

or 2.3 for most cases. For different IGTV, the relationship between cost and β is similar, but it is not necessarily monotonic. Due to this similarity, only the upper and lower bounds are shown for each case using the corresponding IGTV numbers. It appears that the value of β also has an influence on the error, especially for larger tolerances as shown in **Figure 7**. The solution error trends for different IGTV are identical; hence, the plot of error from Equation (9) versus β is shown for only one case. **Figure 8** is used to further highlight the effect of IGTV on both function evaluations and error under different solution strategies and convergence tolerances. The plots are shown only for $\beta = 2$ case with three convergence tolerance values. The dependency of error on IGTV for AL-AD and TDQA is very high at $\tau = 10^{-2}$, but minimal or nonexistent at $\tau = 10^{-3}$ and $\tau = 10^{-4}$. Thus, TDQA and AL-AD are much more dependent on IGTV than DQA and AL. The efficiency of AL and DQA

methods changes drastically with tighter termination tolerance, while solution error for AL and DQA does not change very much. Hence, for larger τ , AL and DQA are less costly, whereas for smaller τ , AL-AD and TDQA are more efficient.

Problem 2: This is a 14-variable geometric programming problem with the AIO formulation expressed as [5]

$$\begin{aligned} \min_{x_1, x_2, \dots, x_{14}} \quad & f = x_1^2 + x_2^2 \\ \text{s.t.} \quad & g_1 = \frac{x_3^{-2} + x_4^2}{x_5^2} \leq 1 \\ & g_2 = \frac{x_5^2 + x_6^{-2}}{x_7^2} \leq 1 \\ & g_3 = \frac{x_8^2 + x_9^2}{x_{11}^2} \leq 1 \end{aligned}$$

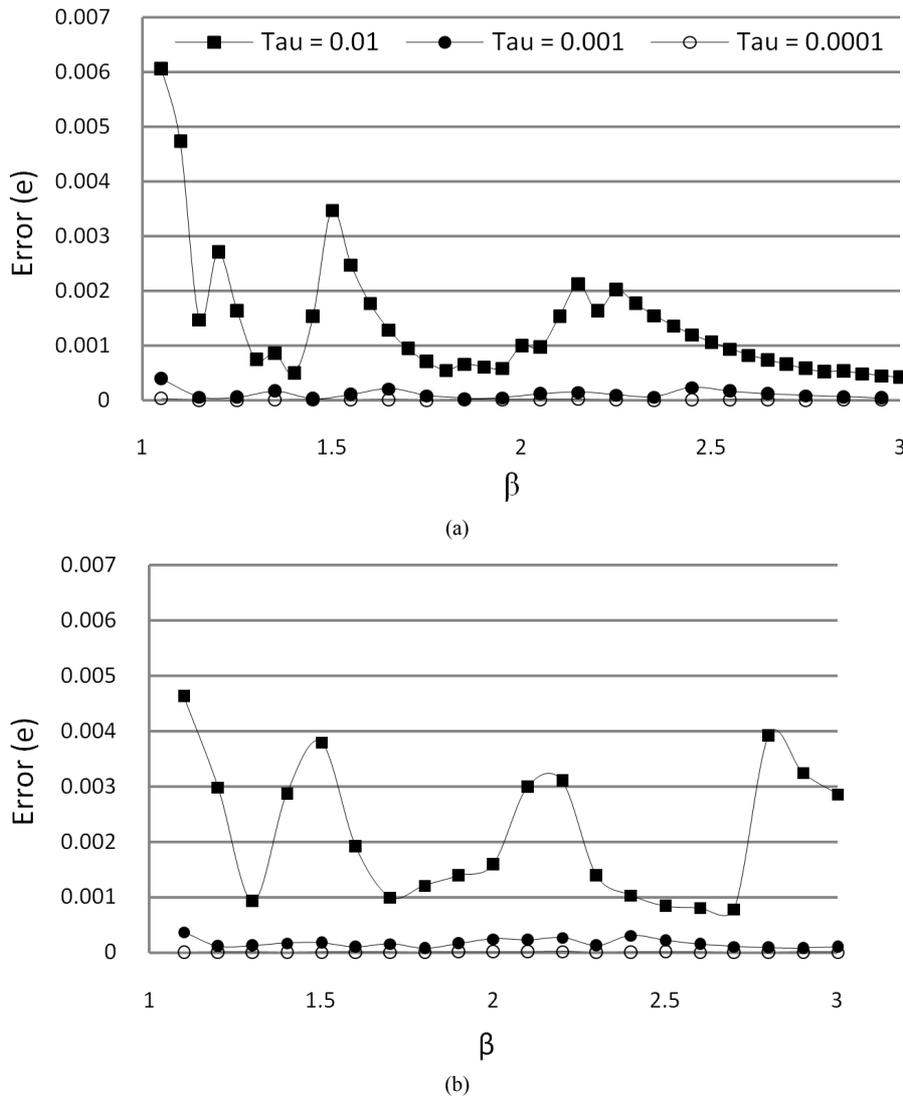


Figure 7. Error trends for (a) AL and (b) DQA solutions of Problem 1 using $x_5 = 6$ with $\tau = \{10^{-2}, 10^{-3}, 10^{-4}\}$.

$$g_4 = \frac{x_8^{-2} + x_{10}^2}{x_{11}^2} \leq 1$$

$$g_5 = \frac{x_{11}^2 + x_{12}^{-2}}{x_{13}^2} \leq 1$$

$$g_6 = \frac{x_{11}^2 + x_{12}^2}{x_{14}^2} \leq 1$$

$$h_1 = x_1^2 - (x_3^2 + x_4^{-2} + x_5^2) = 0$$

$$h_2 = x_2^2 - (x_5^2 + x_6^2 + x_7^2) = 0$$

$$h_3 = x_3^2 - (x_8^2 + x_9^{-2} + x_{10}^{-2} + x_{11}^2) = 0$$

$$h_4 = x_6^2 - (x_{11}^2 + x_{12}^2 + x_{13}^2 + x_{14}^2) = 0$$

$$x_1, x_2, \dots, x_{14} \geq 0$$

The global optimum is located at $x^* = [2.84, 3.09, 2.36, 0.76, 0.87, 2.81, 0.94, 0.97, 0.87, 0.80, 1.30, 0.84, 1.76, 1.55]$ with $f^* = 17.6$ and all constraints active.

The decomposition model selected for this problem [15] has five elements in three levels: A top-level element (1) with two children (2 and 3) at level 2, each with one child (4 and 5, respectively) at the bottom level. Local variables in elements 2, 3, 4, and 5 are $\{x_4\}$, $\{x_7\}$, $\{x_8, x_9, x_{10}\}$ and $\{x_{12}, x_{13}, x_{14}\}$ with design constraints being $\{g_4\}$, $\{g_1, h_1\}$, $\{g_2, h_2\}$, $\{g_3, g_4, h_3\}$ and $\{g_5, g_6, h_4\}$, respectively. The parameters x_1, x_2, x_3 and x_6 are the responses/targets between elements 1 - 2, 1 - 3, 2 - 4, and 3 - 5,

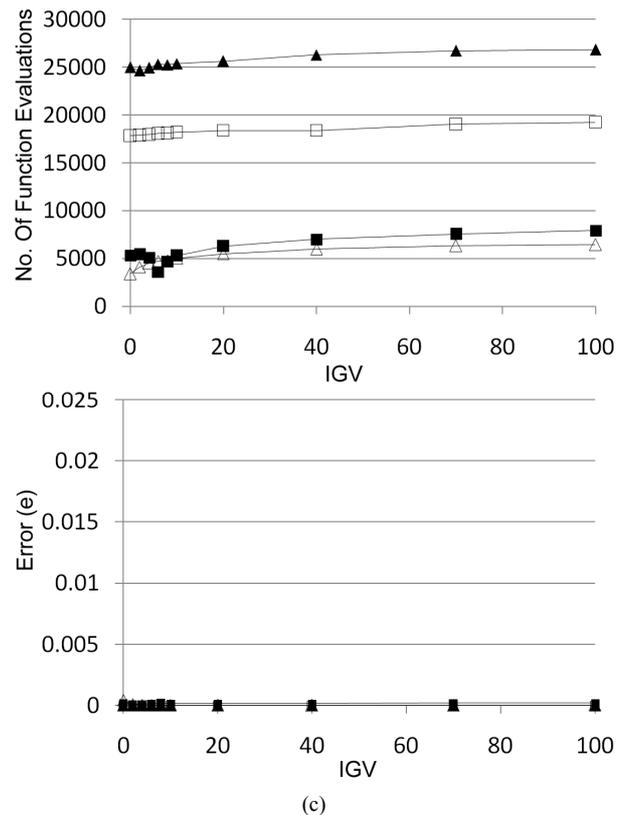
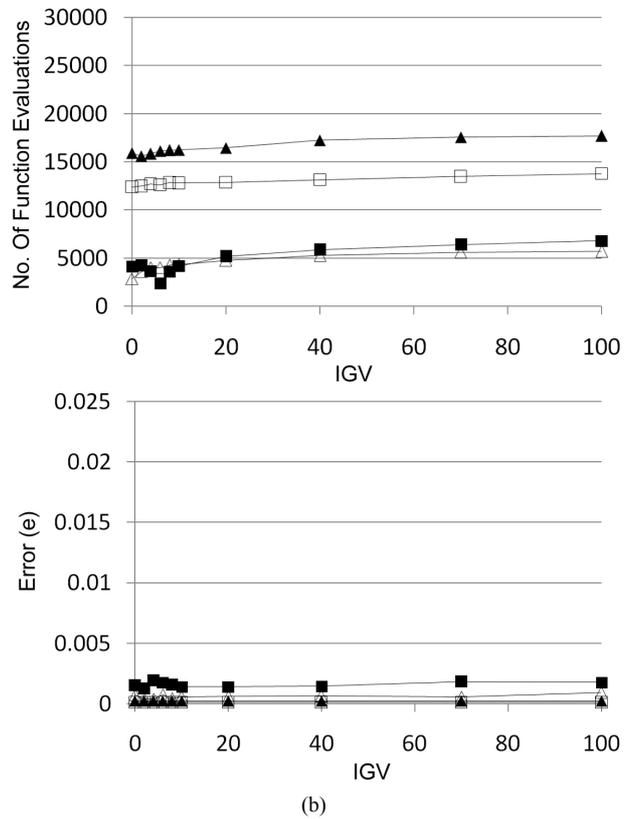
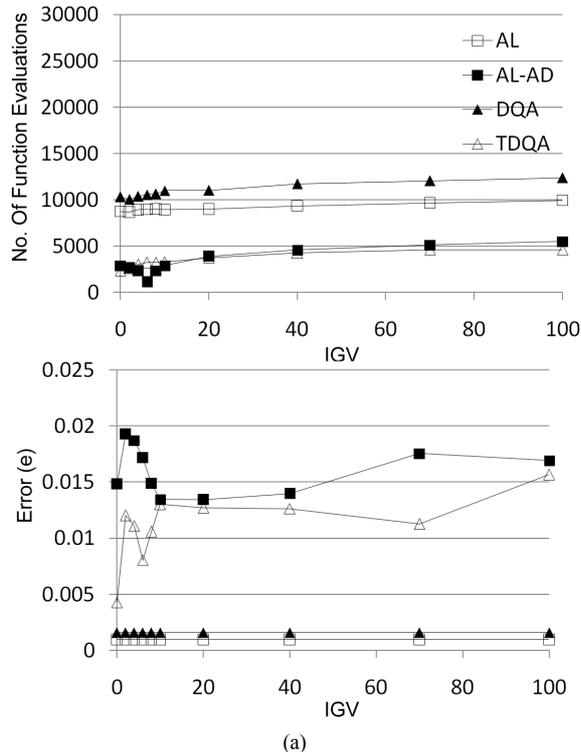


Figure 8. Costand error trends from different solutions of Problem 1 using $\beta = 2$ for AL and DQA with (a) $\tau = 10^{-2}$, (b) $\tau = 10^{-3}$, and (c) $\tau = 10^{-4}$.

respectively, whereas x_5 and x_{11} are the linking variables between elements 2 - 3 and 4 - 5, respectively, both of which are coordinated in element 1.

The initial values for the penalty parameters in all the formulations are taken as $\lambda^{(0)} = 0$ and $w^{(0)} = 1$. The initial design point is $\mathbf{x}^{(0)} = [5, 5, 2.76, 0.25, 1.26, 4.64, 1.39, 0.67, 0.76, 1.7, 2.26, 1.41, 2.71, 2.66]$ for all the formulations, which is the same as that used in the previous studies cited. The IGV for $x_1, x_2, x_3, x_4, x_5, x_6$ and x_{11} are randomly selected in the design domain with a relative distance of $\{0, 2, 4, 6, 8, 10, 20, 40, 70, 100\}$ from the optimum point with the corresponding values shown in **Table 1**. These variables need to have predefined values to start the ATC solution sequence. For example in AL, it is necessary to guess values for response/linking variables x_1, x_2, x_5 and x_{11} from the lower level elements to solve element 1, response value for x_3 from element 4 to solve element 2 and response value for x_6 from element 5 to solve element 3.

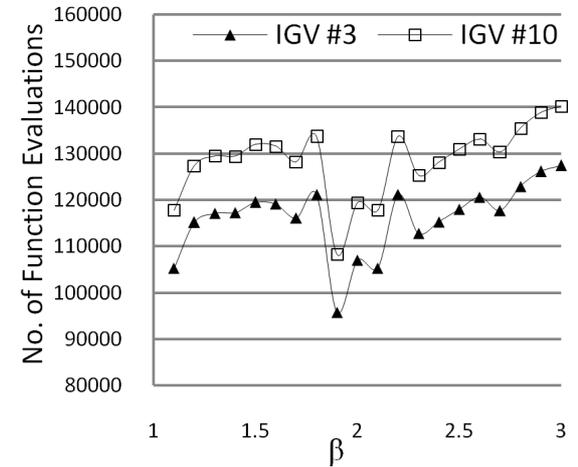
For AL and DQA, $\beta = \{1.1, 1.2, 1.3, \dots, 2.9, 3\}$, whereas for AL-AD and TDQA, $\beta = 1$. The termination tolerances were set to $\tau = \{10^{-2}, 10^{-3}, 10^{-4}\}$. Computational cost in AL and DQA is affected by β values but it follows a non-monotonic manner. It has the minimum computational cost near $\beta = 2$ and it generally increases with higher β values as shown in **Figures 9** and **10**.

The plots in **Figure 11** show that error in both AL and DQA depends on the β value, especially with $\tau = 10^{-2}$, and this is very critical for the DQA method. The error in AL is nearly uniform for $\beta > 1.5$ while in DQA it has an ascending mode.

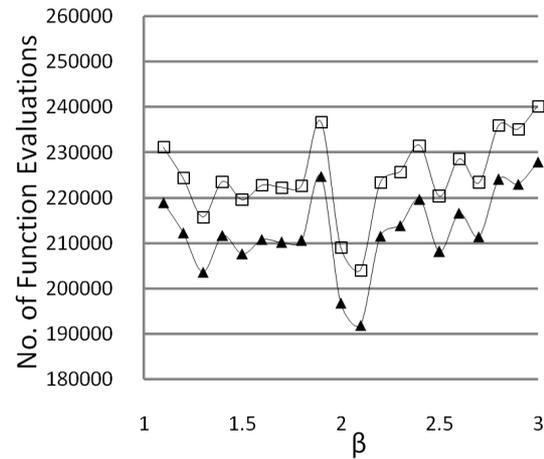
Figure 12 indicates that the dependency of error on IGV for AL-AD and DQA is observable at $\tau = 10^{-2}$, diminishes slightly for TDQA at $\tau = 10^{-3}$, and vanishes at $\tau = 10^{-4}$. It can be concluded that TDQA and, to some extent, AL-AD are much more dependent on the IGV than DQA and AL. The computational costs of AL and DQA

Table 1. List of IGV for response/target and linking variables in Problem 2.

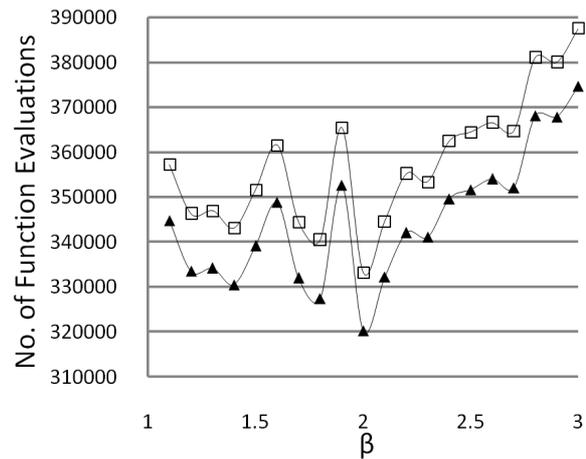
IGV	$[x_1, x_2, x_3, x_5, x_6, x_{11}]$
1	[2.835, 3.090, 2.355, 0.870, 2.812, 1.301]
2	[2.979, 4.094, 1.417, 2.231, 2.886, 0.895]
3	[5.764, 2.848, 0.412, 1.748, 1.222, 0.777]
4	[0.125, 0.835, 3.382, 5.370, 1.457, 1.080]
5	[6.731, 3.675, 3.192, 7.602, 3.964, 2.437]
6	[7.444, 10.626, 2.843, 3.127, 6.366, 3.160]
7	[2.740, 4.545, 7.056, 18.179, 10.027, 5.410]
8	[1.582, 23.522, 19.805, 7.305, 8.027, 29.399]
9	[15.582, 53.774, 12.037, 6.821, 37.460, 29.94]
10	[0.141, 46.214, 8.356, 81.508, 19.002, 36.692]



(a)



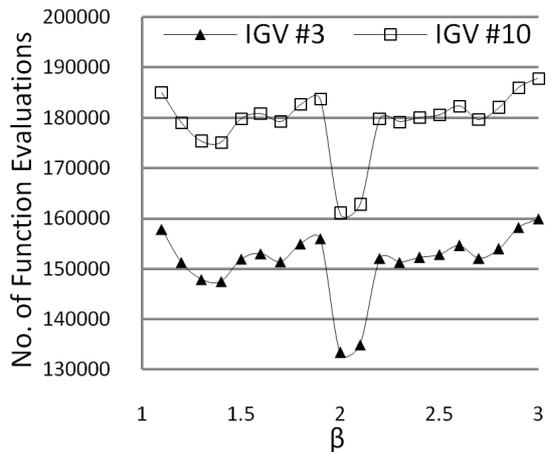
(b)



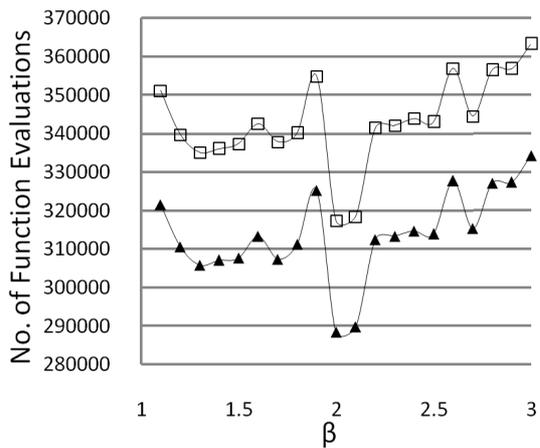
(c)

Figure 9. Cost trends for AL-based solution of Problem 2 using (a) $\tau = 10^{-2}$, (b) $\tau = 10^{-3}$, and (c) $\tau = 10^{-4}$.

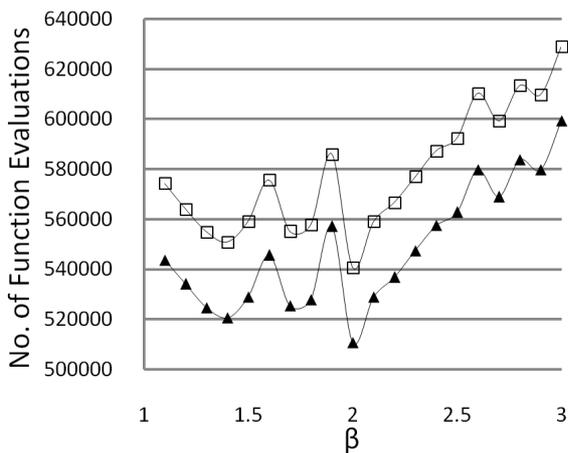
drastically change with tighter termination tolerance, while solution errors in AL and DQA do not change very much. In contrast to AL and DQA, the error in AL-AD and TDQA changes with different τ values while the



(a)



(b)

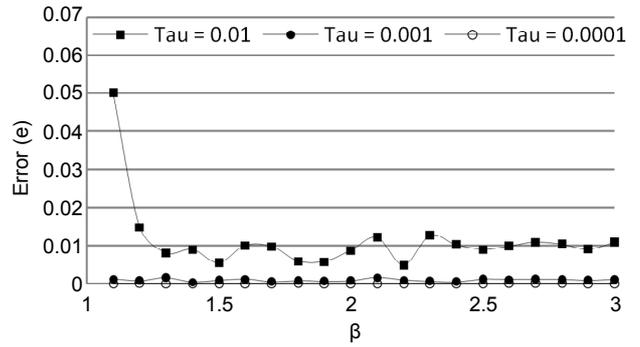


(c)

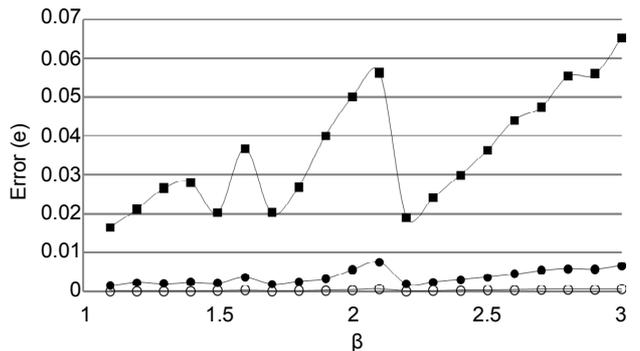
Figure 10. Cost trends for DQA-based solution of Problem 2 using (a) $\tau = 10^{-2}$, (b) $\tau = 10^{-3}$, and (c) $\tau = 10^{-4}$.

computational costs are nearly similar. Hence, for larger τ , AL and DQA are better choices, whereas for tighter τ , AL-AD and TDQA are more efficient.

Problem 3: This is a seven-variable geometric programming problem with only inequality constraints. The



(a)



(b)

Figure 11. Error trends for (a) AL and (b) DQA solutions of Problem 2 using IG V #4 with $\tau = \{10^{-2}, 10^{-3}, 10^{-4}\}$.

corresponding AIO problem is expressed as

$$\begin{aligned} \min f &= (x_1 - 10)^2 + 5(x_2 - 12)^2 + x_3^4 + 3(x_4 - 11)^2 \\ &\quad + 10x_5^6 + 7x_6^2 + x_7^4 - 4x_6x_7 - 10x_6 - 8x_7 \\ \text{s.t. } g_1 &= -127 + 2x_1^2 + 3x_2^4 + x_3 + 4x_4^2 + 5x_5 \leq 0 \\ g_2 &= -282 + 7x_1 + 3x_2 + 10x_3^2 + x_4 - x_5 \leq 0 \\ g_3 &= -196 + 23x_1 + x_2^2 + 6x_6^2 - 8x_7 \leq 0 \\ g_4 &= 4x_1^2 + x_2^2 - 3x_1x_2 + 2x_3^2 + 5x_6 - 11x_7 \leq 0 \\ &\quad -10 \leq x_i \leq 10, i = 1, \dots, 7 \end{aligned}$$

where $\mathbf{x}^* = [2.3305, 1.9513, -0.4775, 4.3657, -0.6245, 1.0371, 1.5942]$ is the unique optimal solution.

The problem is decomposed into three elements in two levels: A top-level element with elements 2 and 3 at level 2. There is no local variable or constraint at the top level. Local variables of element 2 are x_4 and x_5 along with constraints g_1 and g_2 . Local variables of element 3 are x_6 and x_7 with inequality constraints g_3 and g_4 . There is no target variable in this decomposed structure. The linking variables x_1, x_2 and x_3 are shared between elements 2 and 3 and coordinated in element 1.

The starting design point is $\mathbf{x}^{(0)} = [0, 0, 0, 0, 0, 0, 0]$ for all the formulations. The IG V for x_1, x_2 and x_3 are randomly selected in design domain at a distance nearly

equal to {0, 2, 4, 6, 8, 10} from the optimum point with the corresponding values shown in **Table 2**.

For AL and DQA, $\beta = \{1.1, 1.2, 1.3, \dots, 2.9, 3\}$,

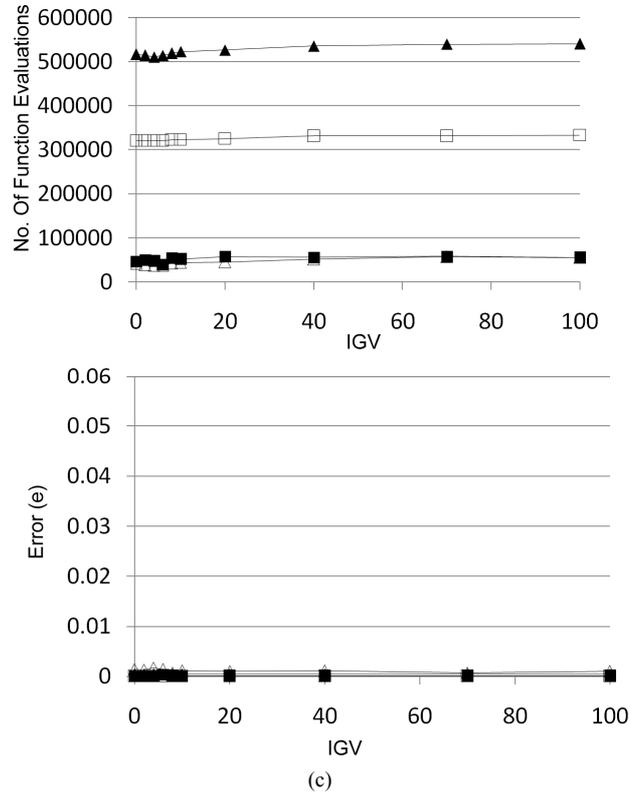
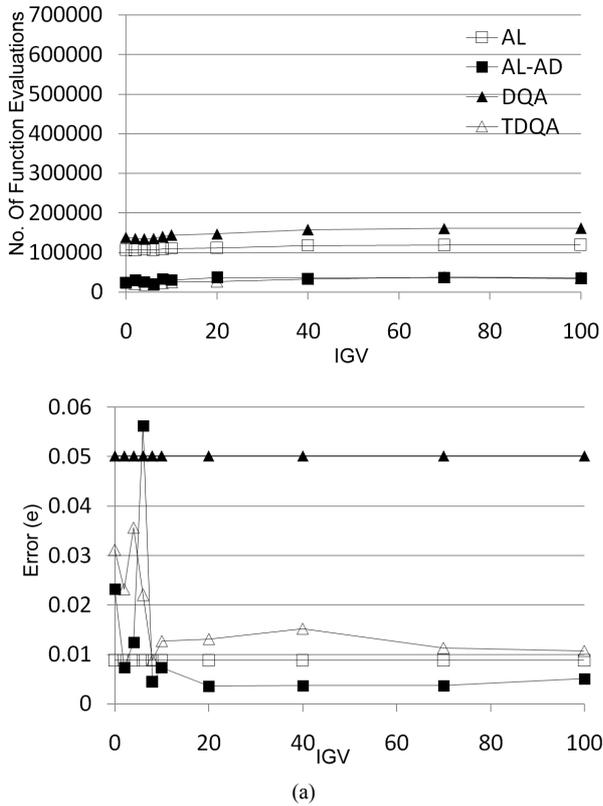
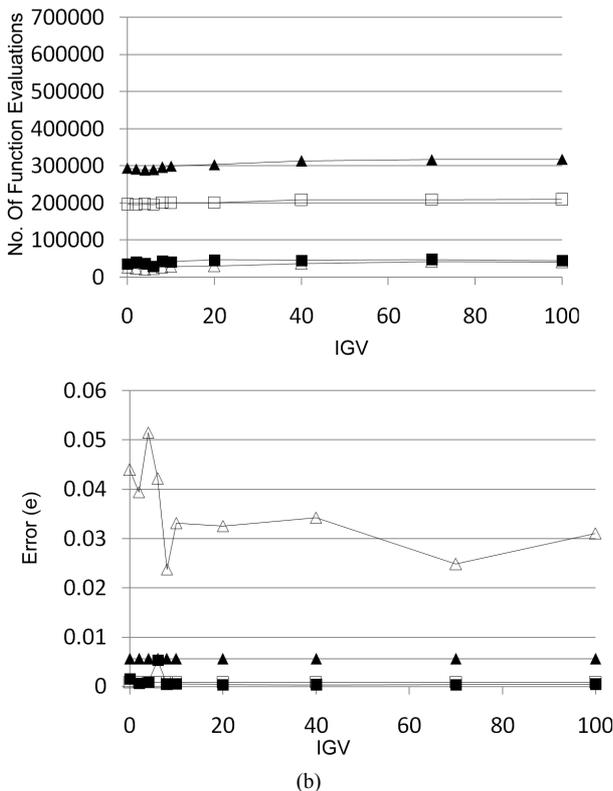


Figure 12. Cost and error trends from different solutions of Problem 2 using $\beta = 2$ for AL and DQA with (a) $\tau = 10^{-2}$, (b) $\tau = 10^{-3}$, and (c) $\tau = 10^{-4}$.

Table 2. List of IG values for linking variables in Problem 3.

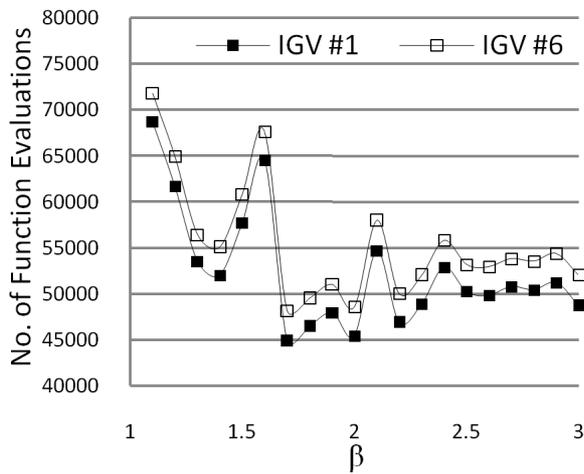
IGV	$[x_1, x_2, x_3]$
1	[2.3305, 1.9514, -0.4775]
2	[0.3635, 1.9144, -0.8961]
3	[-0.0929, 0.0739, 2.1134]
4	[5.0429, -2.6988, -3.2525]
5	[1.3226, 1.4993, 7.4720]
6	[8.1345, -3.6255, 5.3449]



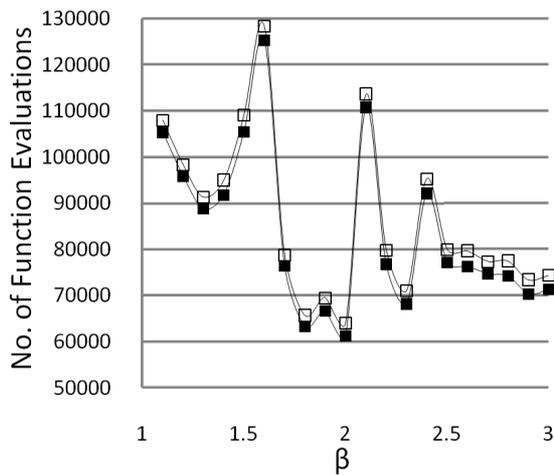
whereas for AL-AD and TDQA, $\beta = 1$. The termination tolerances were set to $\tau = \{10^{-2}, 10^{-3}, 10^{-4}\}$.

Figures 13 and 14 show that the computational cost changes greatly with variations in β value and that the fluctuations are more pronounced for the smaller τ values. **Figure 15** shows that error in AL is slightly dependent on β just for $\tau = 10^{-2}$ and it nearly disappears for $\tau = 10^{-3}$ and $\tau = 10^{-4}$. The error in DQA is more dependent on β than AL.

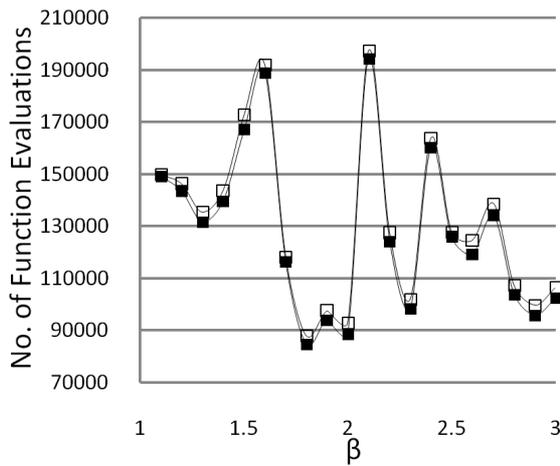
Figure 16 indicates that the computational cost dependency on IG is negligible; the changes in computational cost are lower than 5% for all the methods. The computational cost for AL and DQA, especially for DQA, changes significantly while the error is nearly identical



(a)



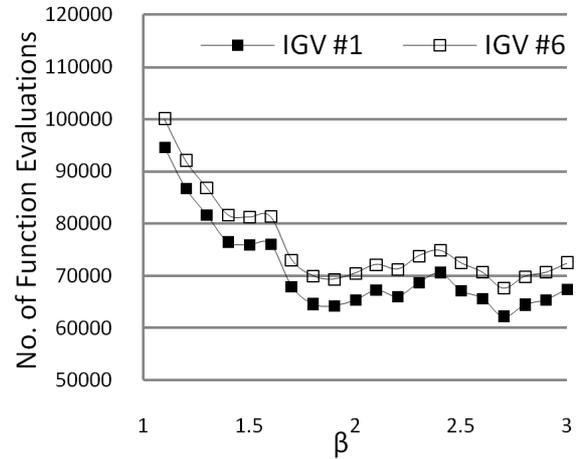
(b)



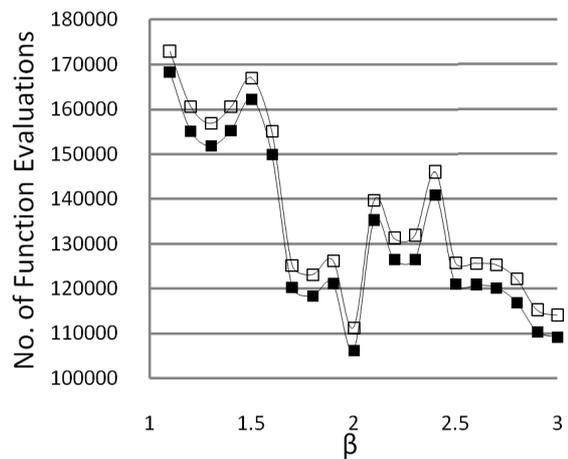
(c)

Figure 13. Cost trends for AL-based solution of Problem 3 using (a) $\tau = 10^{-2}$, (b) $\tau = 10^{-3}$, and (c) $\tau = 10^{-4}$.

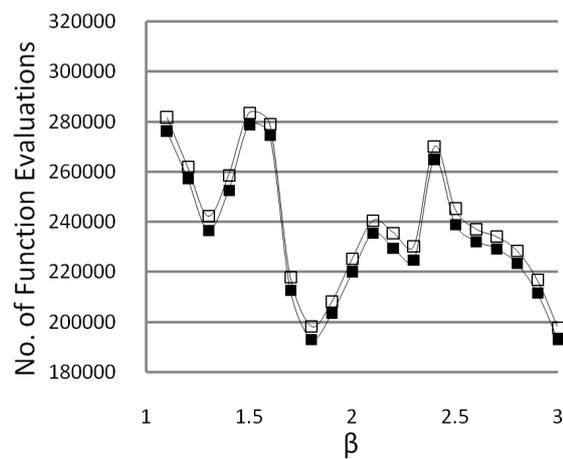
for tighter tolerances. Also, dependency of the error on IGV in AL-AD and TDQA is observable at $\tau = 10^{-2}$ and vanishes for tighter tolerances.



(a)



(b)



(c)

Figure 14. Cost trends for DQA-based solution of Problem 3 using (a) $\tau = 10^{-2}$, (b) $\tau = 10^{-3}$, and (c) $\tau = 10^{-4}$.

3. Conclusions

The numerical behavior of the analytical target cascading (ATC) method was investigated for multilevel optimization of hierarchical systems based on different solution

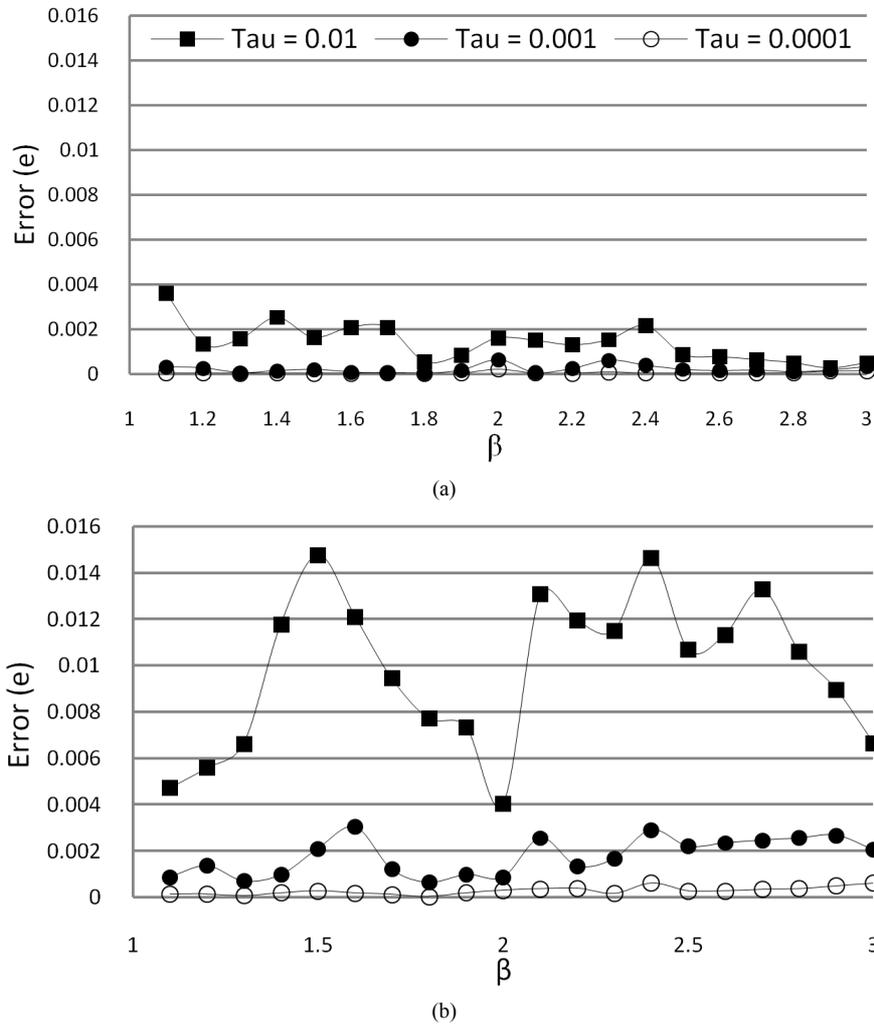


Figure 15. Error trends for (a) AL and (b) DQA solutions of Problem 3 using IGV #3 with $\tau = \{10^{-2}, 10^{-3}, 10^{-4}\}$.

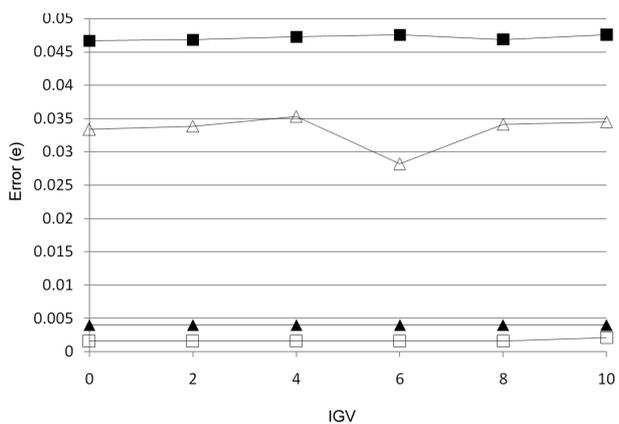
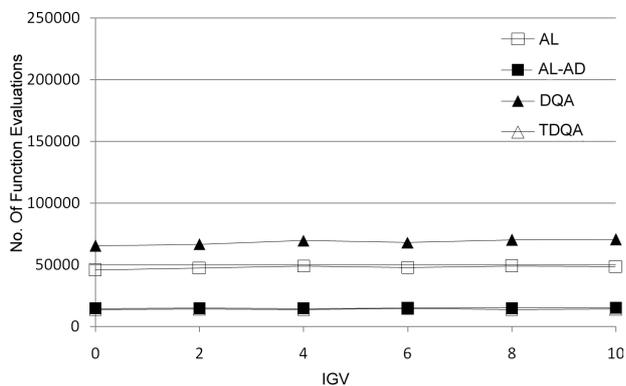
strategies. The strategies considered included Augmented Lagrangian with method of multipliers (AL), Augmented Lagrangian with Alternating Direction method of multipliers (AL-AD), Diagonal Quadratic Approximation (DQA), and Truncated Diagonal Quadratic Approximation (TDQA). Three example problems were used to examine the effects of penalty parameter updating coefficient β and convergence tolerance τ on the computational cost and solution accuracy. In addition, the effect of initial guessed values (IGV) for the response/target and linking variables was also investigated.

The results showed that although the computational cost in the AL and DQA methods is influenced by the value of β , it does not follow a specific ascending/descending pattern. The computational cost dependency on β is generally higher with increasing the convergence tolerance. Although previous studies recommend $\beta > 1$ and $2 < \beta < 3$, the results found here indicate that $1 < \beta < 2$ is also acceptable and that no single value of β can be suggested to reduce the computational cost in all the

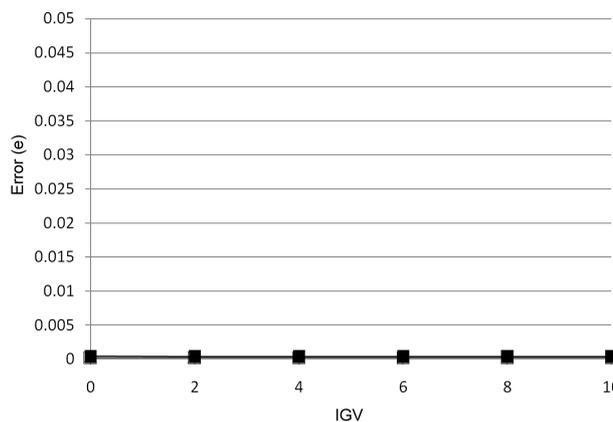
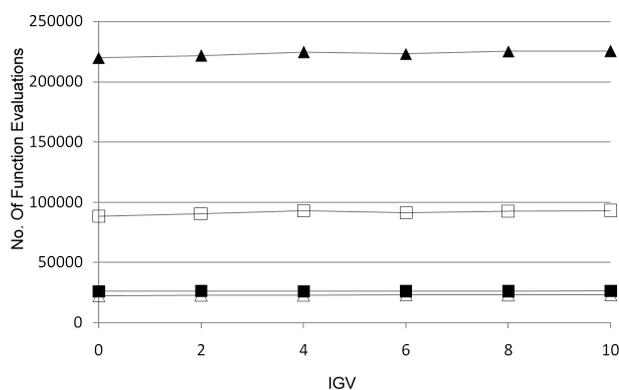
ATC-based optimization problems and solution strategies. The results also showed that the relationship between the computational cost and β is not dependent on the IGV as best noted in the results of the DQA method.

In terms of solution accuracy, AL and DQA results depend on the β value irrespective of the selected IGV. With higher β values, better accuracy is obtained with AL while the behavior is different for DQA. The dependency of solution accuracy on β is reduced with tighter tolerance values. Comparison of the DQA and AL results indicate that AL is more stable in terms of accuracy whereas DQA needs to have a tighter tolerance to obtain reasonable accuracy, although a tighter tolerance causes significant changes in the computational cost. In the absence of optimum β for computational cost and accuracy, the AL method appears to be more reliable than DQA.

By moving the IGV farther away from the corresponding values at the point of optimum, all methods required more function calls, as expected. While the so

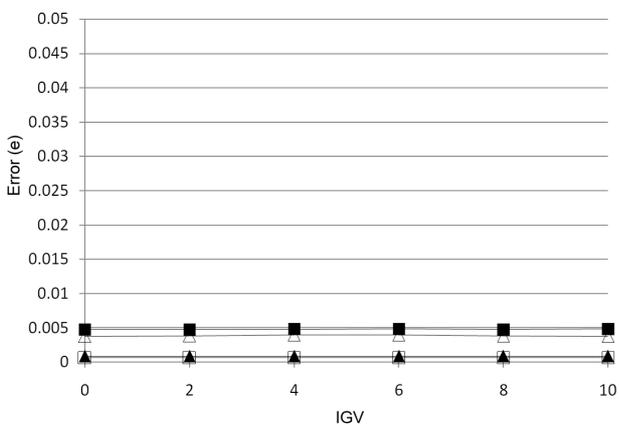
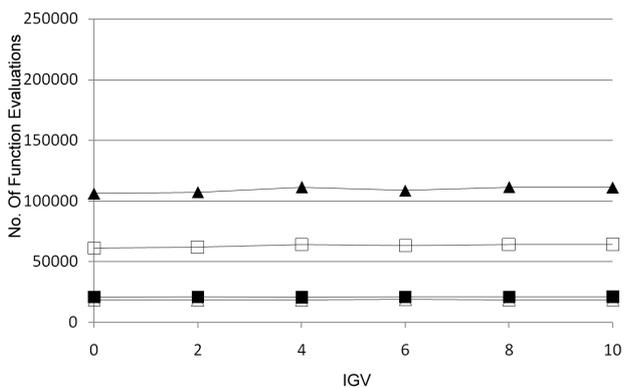


(a)



(c)

Figure 16. Cost and error trends from different solutions of Problem 3 using $\beta = 2$ for AL and DQA with (a) $\tau = 10^{-2}$, (b) $\tau = 10^{-3}$, and (c) $\tau = 10^{-4}$.



(b)

lution accuracy in AL and DQA was not influenced by the choice of IGV, the trend was quite the opposite for AL-AD and TDQA as they both had great dependency on IGV. The inner loop convergence requirement is more costly for AL and DQA than TDQA and AL-AD. Furthermore, the increase in computational cost for AL-AD and TDQA is much greater than AL and DQA when IGV is farther away from the optimum, but TDQA and AL-AD still show better performance. AL-AD and TDQA need tighter termination tolerances to have better accuracy.

In summary, the τ and β values have greater effect on AL and DQA solutions than the other two coordination strategies and they are not influenced by IGV. Hence, in using AL and DQA, appropriate values for these two parameters can enhance both solution accuracy and computational cost. In contrast, the computational cost and accuracy of AL-AD and TDQA are greatly dependent on the IGV.

As part of the future work, the computational characteristics of a newly developed approach based on the exponential method of multipliers within the framework of ATC will be investigated.

4. Acknowledgements

The funding provided for this research by the National Science Foundation under Grant No. CMMI-0826547 is gratefully acknowledged.

REFERENCES

- [1] N. F. Michelena, H. M. Kim and P. Y. Papalambros, "A System Partitioning and Optimization Approach to Target Cascading," *Proceedings of the 12th International Conference on Engineering Design*, Munich, 24-26 August 1999, pp. 1109-1112.
- [2] H. M. Kim, N. F. Michelena, P. Y. Papalambros and T. Jiang, "Target Cascading in Optimal System Design," *ASME Transaction on Journal of Mechanical Design*, Vol. 125, 2003, pp. 474-480. [doi:10.1115/1.1582501](https://doi.org/10.1115/1.1582501)
- [3] H. M. Kim, N. F. Michelena, P. Y. Papalambros and T. Jiang, "Target Cascading in Optimal System Design," *Proceedings of the 26th Design Automation Conference*, Baltimore, 10-13 September 2000, pp. 293-302.
- [4] H. M. Kim, M. Kokkolaras, L. S. Louca, G. J. Delagrammatikas, N. F. Michelena, Z. S. Filipi, P. Y. Papalambros, J. L. Stein and D. N. Assanis, "Target Cascading in Automotive Vehicle Design: A Class 6 Truck Study," *International Journal of Vehicle Design*, Vol. 29, No. 3, 2002, pp. 199-225. [doi:10.1504/IJVD.2002.002010](https://doi.org/10.1504/IJVD.2002.002010)
- [5] H. M. Kim, D. G. Rideout, P. Y. Papalambros and J. L. Stein, "Analytical Target Cascading in Automotive Vehicle Design," *Proceedings of ASME Design Engineering Technical Conference and Computers and Information in Engineering Conference*, Pittsburgh, 9-12 September 2001, pp. 661-670.
- [6] N. F. Michelena, H. Park and P. Y. Papalambros, "Convergence Properties of Analytical Target Cascading," *AIAA Journal*, Vol. 41, No. 5, 2003, pp. 897-905. [doi:10.2514/2.2025](https://doi.org/10.2514/2.2025)
- [7] N. Tzevelekos, M. Kokkolaras, P. Y. Papalambros, M. F. Hulshof, L. F. P. Etman and J. E. Rooda, "An Empirical Local Convergence Study of Alternative Coordination Schemes in Analytical Target Cascading," *Proceedings of the 5th World Congress on Structural and Multidisciplinary Optimization*, Venice, 19-23 May 2003, pp. 1-6.
- [8] J. J. Michalek and P. Y. Papalambros, "An Efficient Weighting Update Method to Achieve Acceptable Inconsistency Deviation in Analytical Target Cascading," *Journal of Mechanical Design*, Vol. 127, No. 2, 2005, pp. 206-214. [doi:10.1115/1.1830046](https://doi.org/10.1115/1.1830046)
- [9] J. J. Michalek and P. Y. Papalambros, "Weights, Norms, and Notation in Analytical Target Cascading," *Journal of Mechanical Design*, Vol. 127, 2005, pp. 499-501. [doi:10.1115/1.1862674](https://doi.org/10.1115/1.1862674)
- [10] S. Tosserams, "Analytical Target Cascading: Convergence Improvement by Subproblem Post-Optimality Sensitivities," MS Thesis, Eindhoven University of Technology, Eindhoven, 2004.
- [11] D. P. Bertsekas, "Nonlinear Programming," 2nd Edition, Athena Scientific, Belmont, 2003.
- [12] J. B. Lassiter, M. M. Wiecek and K. R. Andrighetti, "Lagrangian Coordination and Analytical Target Cascading: Solving ATC-Decomposed Problems with Lagrangian-Duality," *Optimization and Engineering*, Vol. 6, No. 3, 2005, pp. 361-381. [doi:10.1007/s11081-005-1744-4](https://doi.org/10.1007/s11081-005-1744-4)
- [13] V. Y. Blouin, J. Lassiter, M. Wiecek and G. M. Fadel, "Augmented Lagrangian Coordination for Decomposed Design Problems," *6th World Congresses of Structural and Multidisciplinary Optimization*, Rio de Janeiro, 30 May-3 June, 2005, pp. 1-10.
- [14] H. M. Kim, W. Chen and M. M. Wiecek, "Lagrangian Coordination for Enhancing the Convergence of Analytical Target Cascading," *AIAA Journal*, Vol. 44, No. 10, 2006, pp. 2197-2207. [doi:10.2514/1.15326](https://doi.org/10.2514/1.15326)
- [15] S. Tosserams, L. F. P. Etman, P. Y. Papalambros and J. E. Rooda, "An Augmented Lagrangian Relaxation for Analytical Target Cascading Using the Alternating Directions Method of Multipliers," *Structural and Multidisciplinary Optimization*, Vol. 31, No. 3, 2006, pp. 176-189. [doi:10.1007/s00158-005-0579-0](https://doi.org/10.1007/s00158-005-0579-0)
- [16] Y. Li, Z. Lu and J. J. Michalek, "Diagonal Quadratic Approximation for Parallelization of Analytical Target Cascading," *Journal of Mechanical Design*, Vol. 130, 2008, pp. 051402-1-051402-11.
- [17] W. Wang, V. Y. Blouin, M. Gardenghi, M. M. Wiecek, G. M. Fadel and B. Sloop, "A Cutting Plane Method For Analytical Target Cascading With Augmented Lagrangian Coordination," *Proceedings of the ASME International Design Engineering Technical Conferences & Computers and Information in Engineering Conference IDETC/CIE*, Montreal, 15-18 August 2010, pp. 791-800.