

Lagrangian Coordination for Enhancing the Convergence of Analytical Target Cascading

Harrison M. Kim*

University of Illinois at Urbana-Champaign, Urbana, Illinois 61801

Wei Chen†

Northwestern University, Evanston, Illinois 60208

and

Margaret M. Wiecek‡

Clemson University, Clemson, South Carolina 29634

DOI: 10.2514/1.15326

Analytical target cascading is a hierarchical multilevel multidisciplinary design methodology. In analytical target cascading, top-level design targets (i.e., specifications) are propagated to lower-level design problems in a consistent and efficient manner. In this paper, a modified Lagrangian dual formulation and coordination for analytical target cascading are developed to enhance a formulation and coordination proposed earlier in the literature. The proposed approach guarantees all the properties established earlier and additionally offers new significant advantages. As established previously for the convex case, the proposed analytical target cascading coordination converges to a global optimal solution with corresponding optimal Lagrange multipliers in the dual space. The Lagrange multipliers can be viewed as the weights for deviations in analytical target cascading formulations. Thus the proposed coordination algorithm finds the optimal solution and the optimal weights for the deviation terms simultaneously. The enhancement allows for target cascading between levels, for the use of augmented Lagrangian to improve convergence of the coordination algorithm, and for prevention of unboundedness. A guideline to set the step size for subgradient optimization when solving the Lagrangian dual problem is also proposed.

Nomenclature

- C_{sys} = children subproblem set with respect to a system problem
- \mathbf{c} = penalty constants for the augmented terms
- $\mathbf{R}_{\text{sub}}^{\text{sub}}$ = subsystem responses at the subsystem level
- $\mathbf{R}_{\text{sub}}^{\text{sys}}$ = subsystem responses at the system level
- \mathbf{T}_{sys} = system targets
- $\mathbf{y}_{\text{sub}}^{\text{sub}}$ = subsystem linking variables at the subsystem level
- $\mathbf{y}_{\text{sub}}^{\text{sys}}$ = system linking variables
- $\mathbf{y}_{\text{sub}}^{\text{sys}}$ = subsystem linking variables at the system level
- α_k = step size for subsystem k
- $\lambda_k^{\mathbf{R}}$ = Lagrange multipliers associated with the deviations of responses of subsystem k
- $\lambda_k^{\mathbf{y}}$ = Lagrange multipliers associated with the deviations of linking variables of subsystem k
- ξ_k = subgradient for subsystem k

1. Introduction

ANALYTICAL target cascading (ATC) [1–5] is a hierarchical, multilevel, multidisciplinary methodology for large-scale system design. In ATC, a large-scale system is assumed to be hierarchically decomposed by available decomposition methods: physics-based, aspect-based, or model-based decomposition [6]. Hierarchical decision-making (i.e., cascading targets) is then

modeled as individual design optimization problems at each level, possibly in multiple subproblems. The focus of this paper is on developing a Lagrangian duality-based coordination algorithm that improves the convergence of ATC to enforce consistency between subproblems at multiple levels.

Under the target cascading paradigm, the original large-scale optimization problem is formulated as follows:

P_{original} :

$$\min_{\mathbf{x}} \|\mathbf{R}(\mathbf{x}) - \mathbf{T}\|_2^2 \quad (1)$$

subject to

$$\mathbf{g}(\mathbf{x}) \leq \mathbf{0}, \quad \mathbf{h}(\mathbf{x}) = \mathbf{0}$$

where the objective is to minimize the deviation between the targets \mathbf{T} and overall responses \mathbf{R} with respect to design variables \mathbf{x} , and the deviation is measured with the squared L_2 norm denoted by $\|\cdot\|_2^2$, for example,

$$\|\mathbf{x}\|_2^2 = \sum_i x_i^2$$

The inequality and equality design constraints are \mathbf{g} and \mathbf{h} , respectively.

With the ATC approach, the original problem is decomposed into multiple levels (two or more levels), that is, the constraints \mathbf{g} and \mathbf{h} and the design variables \mathbf{x} are decomposed into multiple levels (supersystem, system, subsystem, etc.). In the context of multidisciplinary design optimization (MDO), ATC can be viewed as a hierarchical multilevel formulation [1,5] as well as a coordination methodology [7]. Details of the ATC formulations are presented in Sec. II. Also, Allison et al. [8] compared ATC and collaborative optimization [9], an MDO approach, to highlight the similarities and differences in their use and applicability. In ATC, starting from the top-level targets, an optimization problem at the top level is solved to find an optimal design and subtargets for the immediate lower-level subsystems. At the next lower level, the objective is defined as the minimum deviation from the upper-level

Received 8 January 2005; revision received 24 May 2006; accepted for publication 2 June 2006. Copyright © 2006 by Wei Chen. Published by the American Institute of Aeronautics and Astronautics, Inc., with permission. Copies of this paper may be made for personal or internal use, on condition that the copier pay the \$10.00 per-copy fee to the Copyright Clearance Center, Inc., 222 Rosewood Drive, Danvers, MA 01923; include the code \$10.00 in correspondence with the CCC.

*Assistant Professor, Department of Industrial and Enterprise Systems Engineering; hmkim@uiuc.edu.

†Associate Professor, Department of Mechanical Engineering, 2145 Sheridan Road, Tech B224; weichen@northwestern.edu.

‡Professor, Department of Mathematical Sciences; wmalgor@clemson.edu.

targets with respect to the current responses and linking variables. Once the lower-level problem has been solved, the adjusted responses and linking variables are returned to the upper level so that the overall responses are adjusted based on the feasibility of lower-level design. This top-down and bottom-up approach follows the decision-making process in an organization.

Because the original design problem in Eq. (1) is assumed to be feasible [i.e., there exists a design satisfying all the constraints of Eq. (1)], from the computational aspect, one of the objectives of ATC is to enforce consistency between subproblem designs, indicating zero deviations in the subproblems. Thus an optimal ATC solution should enforce all the deviations to be less than the acceptable tolerance except for the overall response from the target in the top-level objective. Mathematically, the coordination sequence can have variations, for example, top-down and bottom-up [6], however, all of them should converge to the same optimal solution. Michelena et al. [10] analyzed ATC in the context of hierarchical overlapping coordination, but the repetition of the inner-loop coordination in case of multiple levels was a drawback in computational efficiency. Tosserams et al. [11] proposed to adopt the alternating directions method [12] to handle the inner loop, and the efficiency has been improved in their augmented Lagrangian formulation for the ATC.

In general, the weights associated with the deviation terms can be used to reflect the importance of minimizing the discrepancies between decomposed design elements. The conventional idea has been to set the weights for the ATC-introduced deviation terms as large values so that the deviation terms become smaller within the acceptable tolerance. However, large weights tend to cause numerical instability around the optimal solution as the deviation terms can be viewed as the external penalty terms [13]. Michalek and Papalambros [14] introduced a scheme to update the weight for each deviation term in the ATC as the solution process progresses by using the Karush–Kuhn–Tucker condition with a prespecified tolerance (i.e., allowable inconsistency). In their approach, as the prespecified tolerance becomes smaller, the weights become larger.

In this paper, a Lagrangian duality coordination algorithm is developed to improve the convergence of ATC and to enforce consistency between subproblems at multiple levels. This work follows upon the study by Lassiter et al. [15], who placed ATC in the historical context of Lagrangian decomposition and coordination methods for large-scale systems [16,17], proposed an ATC-based decomposition and coordination approach based on Lagrangian duality theory, and established its convergence. In the current paper, a revised Lagrangian-based ATC formulation and a solution process are proposed. Similar to the approach of Lassiter et al., the revised formulation guarantees that under convexity assumptions for the original design objective and constraints, the resulting Lagrangian coordination algorithm finds a unique global optimum as well as (optimal) weights for the ATC deviation terms. Additionally, the current study enhances the results in Lassiter et al. in four aspects:

1) A new decomposition scheme is proposed for the Lagrangian primal and dual problems, which results in the enhanced coordination algorithm.

2) In the proposed coordination algorithm, the adjacent-level problems are solved consecutively, which allows for passing targets between these levels and therefore is faithful to the fundamental ATC property of target cascading. On the contrary, in the approach of Lassiter et al. [15], the adjacent-level problems are solved simultaneously without target cascading.

3) The revised ATC formulation satisfies the Lagrangian separability property and therefore allows for the use of augmented Lagrangian, which was not possible before. Consequently, convergence of the proposed coordination process can be improved.

4) The revised ATC formulation prevents the subproblems from being unbounded, which, again, is not guaranteed in the approach of Lassiter et al. [15].

Two demonstration cases show that in the proposed Lagrangian coordination, the weights do not necessarily need to be large, and that for a convex ATC-decomposed problem there exists a unique set of weights to achieve the prespecified tolerance. Two aspects of the original problem, feasibility and ATC separability, are important

assumptions for the proposed approach. The feasibility assumption is concerned with the existence of a feasible solution that satisfies the feasible design space, although the overall responses \mathbf{R} may not achieve the targets \mathbf{T} perfectly. The separability assumption is concerned with the ATC-decomposed problem structure. The design constraints \mathbf{g} and \mathbf{h} are ATC-separable if subproblems do not share the same constraint. Common variables assume the form of responses and linking variables in ATC. In this paper, the goal is to find an ATC feasible/optimal solution while identifying the weights for the deviation terms introduced in the ATC formulation to match the responses and linking variables, not the weights for meeting the overall targets [i.e., the original objective in Eq. (1)].

The paper proceeds as follows. The traditional ATC formulation is revisited in Sec. II, whereas the Lagrangian dual ATC formulation is presented in Sec. III. In Sec. IV, the revised Lagrangian dual coordination for ATC is proposed and it is shown that this dual-solution process produces a solution that is also optimal for the original problem. The proposed algorithm is implemented in combination with the augmented Lagrangian and the weight issue is addressed. In Sec. V, the Lagrangian ATC formulation and coordination are demonstrated on two examples, convex and nonconvex, for which a global optimal solution is obtained and the corresponding weights for the deviation terms are found. The speed of convergence is also investigated with different step size update formulas. The paper is summarized in Sec. VI.

II. ATC Formulation

For simplicity, ATC coordination for the bilevel formulation is considered, however, the proposed algorithm is not limited only to the bilevel case. In the case where the problem hierarchy is composed of more than two levels, the proposed algorithm can be repeatedly applied to any combinations of the two different sublevels of the overall hierarchy that constitute the entire coordination. Here it is assumed that under the ATC paradigm the original problem is decomposed into two levels: system and subsystem subproblems. The top system-level problem with the deviation terms in the objective is stated as follows:

$$P_{\text{sys}}: \min_{\mathbf{x}_{\text{sys}}, \mathbf{R}_{\text{sub}}^{\text{sys}}, \mathbf{y}_{\text{sub}}^{\text{sys}}} \|\mathbf{R}_{\text{sys}} - \mathbf{T}_{\text{sys}}\|_2^2 + \sum_{k \in C_{\text{sys}}} \|\mathbf{w}_k^{\mathbf{R}} \circ (\mathbf{R}_{\text{sub},k}^{\text{sys}} - \mathbf{R}_{\text{sub},k}^{\text{sub}})\|_2^2 + \sum_{k \in C_{\text{sys}}} \|\mathbf{w}_k^{\mathbf{y}} \circ (\mathbf{y}_{\text{sub},k}^{\text{sys}} - \mathbf{y}_{\text{sub},k}^{\text{sub}})\|_2^2 \quad (2)$$

where

$$\mathbf{R}_{\text{sys}} = \mathbf{R}_{\text{sys}}(\mathbf{x}_{\text{sys}}, \mathbf{R}_{\text{sub}}^{\text{sys}})$$

subject to

$$\mathbf{g}_{\text{sys}}(\mathbf{R}_{\text{sys}}, \mathbf{x}_{\text{sys}}) \leq \mathbf{0}, \quad \mathbf{h}_{\text{sys}}(\mathbf{R}_{\text{sys}}, \mathbf{x}_{\text{sys}}) = \mathbf{0}$$

where the objective is to minimize the deviation between the system targets \mathbf{T}_{sys} and the system responses \mathbf{R}_{sys} that are functions of the system design variables \mathbf{x}_{sys} and the subsystem responses $\mathbf{R}_{\text{sub}}^{\text{sys}}$. The deviation is measured by the squared L_2 norm. The design constraints are \mathbf{g}_{sys} and \mathbf{h}_{sys} , and the deviation constraints are defined with respect to the values of the subsystem responses $\mathbf{R}_{\text{sub}}^{\text{sub}}$ and the subsystem linking variables $\mathbf{y}_{\text{sub},k}^{\text{sys}}$, where k is an element of the children subproblem set C_{sys} . Note that $\mathbf{R}_{\text{sub}}^{\text{sys}}$ refer to the subsystem response values at the system level and $\mathbf{R}_{\text{sub}}^{\text{sub}}$ refer to the subsystem response values at the subsystem level, i.e., for the same physical entity, variables are duplicated at different levels. The deviations are multiplied by vectors of positive weights $\mathbf{w}_k^{\mathbf{R}}$ and $\mathbf{w}_k^{\mathbf{y}}$. The symbol \circ is used to define term-by-term multiplication of two vectors, for example,

$$(a_1, a_2, \dots, a_n) \circ (b_1, b_2, \dots, b_n) = (a_1 b_1, a_2 b_2, \dots, a_n b_n)$$

(see [14]). In Eq. (2), the responses and linking variables from the subsystem level are set as parameters with superscript sub. Also, the subscript k appearing in the responses and linking variables associates these vectors with subsystem k , with $k \in C_{\text{sys}}$. The bottom subsystem-level ATC subproblem for the k th element is formulated as follows:

$$P_{\text{sub},k}: \min_{\mathbf{x}_{\text{sub},k}, \mathbf{y}_{\text{sub},k}} \|\mathbf{w}_k^{\mathbf{R}} \circ (\mathbf{R}_{\text{sub},k}^{\text{sys}} - \mathbf{R}_{\text{sub},k}^{\text{sub}})\|_2^2 + \|\mathbf{w}_k^{\mathbf{y}} \circ (\mathbf{y}_{\text{sub},k}^{\text{sys}} - \mathbf{y}_{\text{sub},k}^{\text{sub}})\|_2^2 \quad (3)$$

where

$$\mathbf{R}_{\text{sub},k}^{\text{sub}} = \mathbf{R}_{\text{sub},k}^{\text{sub}}(\mathbf{x}_{\text{sub},k}, \mathbf{y}_{\text{sub},k}^{\text{sub}})$$

subject to

$$\mathbf{g}_{\text{sub},k}(\mathbf{R}_{\text{sub},k}^{\text{sub}}, \mathbf{x}_{\text{sub},k}, \mathbf{y}_{\text{sub},k}^{\text{sub}}) \leq \mathbf{0}, \quad \mathbf{h}_{\text{sub},k}(\mathbf{R}_{\text{sub},k}^{\text{sub}}, \mathbf{x}_{\text{sub},k}, \mathbf{y}_{\text{sub},k}^{\text{sub}}) = \mathbf{0}$$

where the response and linking variable targets from the system are set constant with superscript sys.

The all-in-one (AIO) problem obtained by assembling together the preceding system and subsystem subproblems is stated as follows:

$$P_{\text{AIO}}: \min_{\mathbf{x}_{\text{sys}}, \mathbf{R}_{\text{sub}}^{\text{sys}}, \mathbf{R}_{\text{sub}}^{\text{sub}}, \mathbf{x}_{\text{sub}}, \mathbf{y}_{\text{sub}}^{\text{sys}}, \mathbf{y}_{\text{sub}}^{\text{sub}}} f = \|\mathbf{R}_{\text{sys}} - \mathbf{T}_{\text{sys}}\|_2^2 + \sum_{k \in C_{\text{sys}}} \|\mathbf{w}_k^{\mathbf{R}} \circ (\mathbf{R}_{\text{sub},k}^{\text{sys}} - \mathbf{R}_{\text{sub},k}^{\text{sub}})\|_2^2 + \sum_{k \in C_{\text{sys}}} \|\mathbf{w}_k^{\mathbf{y}} \circ (\mathbf{y}_{\text{sub},k}^{\text{sys}} - \mathbf{y}_{\text{sub},k}^{\text{sub}})\|_2^2 \quad (4)$$

where

$$\mathbf{R}_{\text{sys}} = \mathbf{R}_{\text{sys}}(\mathbf{x}_{\text{sys}}, \mathbf{R}_{\text{sub}}^{\text{sys}}), \quad \mathbf{R}_{\text{sub},k}^{\text{sub}} = \mathbf{R}_{\text{sub},k}^{\text{sub}}(\mathbf{x}_{\text{sub},k}, \mathbf{y}_{\text{sub},k}^{\text{sub}}) \quad \forall k \in C_{\text{sys}}$$

subject to

$$\mathbf{g}_{\text{sys}}(\mathbf{R}_{\text{sys}}, \mathbf{x}_{\text{sys}}) \leq \mathbf{0}, \quad \mathbf{h}_{\text{sys}}(\mathbf{R}_{\text{sys}}, \mathbf{x}_{\text{sys}}) = \mathbf{0}, \quad \mathbf{g}_{\text{sub},k}(\mathbf{R}_{\text{sub},k}^{\text{sub}}, \mathbf{x}_{\text{sub},k}, \mathbf{y}_{\text{sub},k}^{\text{sub}}) \leq \mathbf{0} \quad \forall k \in C_{\text{sys}}, \quad \mathbf{h}_{\text{sub},k}(\mathbf{R}_{\text{sub},k}^{\text{sub}}, \mathbf{x}_{\text{sub},k}, \mathbf{y}_{\text{sub},k}^{\text{sub}}) = \mathbf{0} \quad \forall k \in C_{\text{sys}}$$

Note that at optimality of P_{AIO} , the deviation terms for the subsystem responses and linking variables become zero if the original problem is feasible, that is, there exists a feasible design for which

$$\mathbf{R}_{\text{sub},k}^* = \mathbf{R}_{\text{sub},k}^{\text{sys}} = \mathbf{R}_{\text{sub},k}^{\text{sub}} \quad \forall k \in C_{\text{sys}}$$

and

$$\mathbf{y}_{\text{sub},k}^* = \mathbf{y}_{\text{sub},k}^{\text{sys}} = \mathbf{y}_{\text{sub},k}^{\text{sub}} \quad \forall k \in C_{\text{sys}}$$

Hence, solving the P_{AIO} in Eq. (4) is equivalent to solving the original problem. Because of their behavior at optimality of the P_{AIO} , the deviation terms can be moved from the objective function down to the constraints as new (equality) deviation constraints in Eq. (5). At optimality, the deviation constraints should be satisfied (the deviations should become zero) regardless of the weights, thus the weights in the deviation constraints can be eliminated. Note that the norm for the deviation constraints has been dropped in the P_{AIO} , which is a departure from the conventional ATC formulation, but it will be adjusted by the augmented Lagrangian formulation (see Sec. IV.B).

$$P_{\text{AIO}}: \min_{\mathbf{x}_{\text{sys}}, \mathbf{R}_{\text{sub}}^{\text{sys}}, \mathbf{R}_{\text{sub}}^{\text{sub}}, \mathbf{x}_{\text{sub}}, \mathbf{y}_{\text{sub}}^{\text{sys}}, \mathbf{y}_{\text{sub}}^{\text{sub}}} f = \|\mathbf{R}_{\text{sys}} - \mathbf{T}_{\text{sys}}\|_2^2 \quad (5)$$

where

$$\mathbf{R}_{\text{sys}} = \mathbf{R}_{\text{sys}}(\mathbf{x}_{\text{sys}}, \mathbf{R}_{\text{sub}}^{\text{sys}}), \quad \mathbf{R}_{\text{sub},k}^{\text{sub}} = \mathbf{R}_{\text{sub},k}^{\text{sub}}(\mathbf{x}_{\text{sub},k}, \mathbf{y}_{\text{sub},k}^{\text{sub}}) \quad \forall k \in C_{\text{sys}}$$

subject to

$$\mathbf{g}_{\text{sys}}(\mathbf{R}_{\text{sys}}, \mathbf{x}_{\text{sys}}) \leq \mathbf{0}, \quad \mathbf{h}_{\text{sys}}(\mathbf{R}_{\text{sys}}, \mathbf{x}_{\text{sys}}) = \mathbf{0}, \quad \mathbf{g}_{\text{sub},k}(\mathbf{R}_{\text{sub},k}^{\text{sub}}, \mathbf{x}_{\text{sub},k}, \mathbf{y}_{\text{sub},k}^{\text{sub}}) \leq \mathbf{0} \quad \forall k \in C_{\text{sys}}, \quad \mathbf{h}_{\text{sub},k}(\mathbf{R}_{\text{sub},k}^{\text{sub}}, \mathbf{x}_{\text{sub},k}, \mathbf{y}_{\text{sub},k}^{\text{sub}}) = \mathbf{0} \quad \forall k \in C_{\text{sys}}, \quad \mathbf{R}_{\text{sub},k}^{\text{sys}} - \mathbf{R}_{\text{sub},k}^{\text{sub}} = \mathbf{0} \quad \forall k \in C_{\text{sys}}, \quad \mathbf{y}_{\text{sub},k}^{\text{sys}} - \mathbf{y}_{\text{sub},k}^{\text{sub}} = \mathbf{0} \quad \forall k \in C_{\text{sys}}$$

Under the feasibility assumption for the original problem, there should exist a feasible solution of the P_{AIO} with zero deviations and an ATC solution process should produce that solution even without assigning large weight values to each of the deviation terms. In this work, based on Lagrangian duality theory, an ATC solution process is proposed to find a P_{AIO} optimal solution with zero deviations as well as optimal weights for the deviation terms.

III. Dual ATC Formulation

The ATC coordination proposed in this paper is based on well-established theorems on Lagrangian duality [18] that are briefly reviewed in the Appendix. The local duality theorem presents the equivalence between solving the primal problem and the Lagrangian dual problem where local convexity is maintained. Partial duality states that the dual problem can be defined with respect to any subset of primal problem's constraints in the same case. Thus, in the proposed ATC solution process, a dual problem is formulated with respect to only the set of deviation constraints. Then the augmented Lagrangian is introduced by adding in the original objective function an additional penalty term that tends to "convexify" the Lagrangian. The penalty term can be viewed as the deviation term in the ATC formulation. In the proposed ATC coordination based on Lagrangian duality, the augmented Lagrangian approach is adopted to help achieve convergence. In this paper, convexity/affinity is assumed for the original design constraints \mathbf{g} and \mathbf{h} in the original problem. According to the Lagrangian duality theory, and in particular, the partial duality, the AIO problem in Eq. (5) is treated as the primal problem for which the AIO Lagrangian problem can be formulated by relaxing the deviation constraints, assigning Lagrangian multipliers to them and bringing them up to the objective function. The AIO Lagrangian problem (AIO-LP) is as follows:

$$\text{AIO-LP:} \min_{\mathbf{x}_{\text{sys}}, \mathbf{R}_{\text{sub}}^{\text{sys}}, \mathbf{R}_{\text{sub}}^{\text{sub}}, \mathbf{x}_{\text{sub}}, \mathbf{y}_{\text{sub}}^{\text{sys}}, \mathbf{y}_{\text{sub}}^{\text{sub}}} f_{\text{AIO}} = \|\mathbf{R}_{\text{sys}} - \mathbf{T}_{\text{sys}}\|_2^2 + \sum_{k \in C_{\text{sys}}} (\boldsymbol{\lambda}_k^{\mathbf{R}})^T (\mathbf{R}_{\text{sub},k}^{\text{sys}} - \mathbf{R}_{\text{sub},k}^{\text{sub}}) + \sum_{k \in C_{\text{sys}}} (\boldsymbol{\lambda}_k^{\mathbf{y}})^T (\mathbf{y}_{\text{sub},k}^{\text{sys}} - \mathbf{y}_{\text{sub},k}^{\text{sub}}) \quad (6)$$

where

$$\mathbf{R}_{\text{sys}} = \mathbf{R}_{\text{sys}}(\mathbf{x}_{\text{sys}}, \mathbf{R}_{\text{sub}}^{\text{sys}}), \quad \mathbf{R}_{\text{sub},k}^{\text{sub}} = \mathbf{R}_{\text{sub},k}^{\text{sub}}(\mathbf{x}_{\text{sub},k}, \mathbf{y}_{\text{sub},k}^{\text{sub}}) \quad \forall k \in C_{\text{sys}}$$

subject to

$$\mathbf{g}_{\text{sys}}(\mathbf{R}_{\text{sys}}, \mathbf{x}_{\text{sys}}) \leq \mathbf{0}, \quad \mathbf{h}_{\text{sys}}(\mathbf{R}_{\text{sys}}, \mathbf{x}_{\text{sys}}) = \mathbf{0}, \quad \mathbf{g}_{\text{sub},k}(\mathbf{R}_{\text{sub},k}^{\text{sub}}, \mathbf{x}_{\text{sub},k}, \mathbf{y}_{\text{sub},k}^{\text{sub}}) \leq \mathbf{0} \quad \forall k \in C_{\text{sys}}, \quad \mathbf{h}_{\text{sub},k}(\mathbf{R}_{\text{sub},k}^{\text{sub}}, \mathbf{x}_{\text{sub},k}, \mathbf{y}_{\text{sub},k}^{\text{sub}}) = \mathbf{0} \quad \forall k \in C_{\text{sys}}$$

where $\boldsymbol{\lambda}_k^{\mathbf{R}}$ and $\boldsymbol{\lambda}_k^{\mathbf{y}}$ denote the vectors of Lagrange multipliers associated with the deviations of responses and linking variables of

subsystem k , respectively. The Lagrange multipliers are variables of the Lagrangian dual problem (AIO–DP) of the form:

AIO–DP:

$$\max_{\lambda} \phi_{\text{AIO}}(\lambda) \quad (7)$$

where

$$\begin{aligned} \phi_{\text{AIO}}(\lambda) = & \min_{\mathbf{x}_{\text{sys}}, \mathbf{R}_{\text{sub}}^{\text{sys}}, \mathbf{R}_{\text{sub},k}^{\text{sub}}, \mathbf{x}_{\text{sub},k}^{\text{sub}}, \mathbf{y}_{\text{sub},k}^{\text{sub}}} f_{\text{AIO}} = \|\mathbf{R}_{\text{sys}} - \mathbf{T}_{\text{sys}}\|_2^2 \\ & + \sum_{k \in C_{\text{sys}}} (\lambda_k^{\text{R}})^T (\mathbf{R}_{\text{sub},k}^{\text{sys}} - \mathbf{R}_{\text{sub},k}^{\text{sub}}) + \sum_{k \in C_{\text{sys}}} (\lambda_k^{\text{y}})^T (\mathbf{y}_{\text{sub},k}^{\text{sys}} - \mathbf{y}_{\text{sub},k}^{\text{sub}}) \end{aligned}$$

where

$$\begin{aligned} \mathbf{R}_{\text{sys}} &= \mathbf{R}_{\text{sys}}(\mathbf{x}_{\text{sys}}, \mathbf{R}_{\text{sub}}^{\text{sys}}), \\ \mathbf{R}_{\text{sub},k}^{\text{sub}} &= \mathbf{R}_{\text{sub},k}^{\text{sub}}(\mathbf{x}_{\text{sub},k}, \mathbf{y}_{\text{sub},k}^{\text{sub}}) \quad \forall k \in C_{\text{sys}} \end{aligned}$$

subject to

$$\begin{aligned} \mathbf{g}_{\text{sys}}(\mathbf{R}_{\text{sys}}, \mathbf{x}_{\text{sys}}) &\leq \mathbf{0}, \quad \mathbf{h}_{\text{sys}}(\mathbf{R}_{\text{sys}}, \mathbf{x}_{\text{sys}}) = \mathbf{0}, \\ \mathbf{g}_{\text{sub},k}(\mathbf{R}_{\text{sub},k}^{\text{sub}}, \mathbf{x}_{\text{sub},k}, \mathbf{y}_{\text{sub},k}^{\text{sub}}) &\leq \mathbf{0} \quad \forall k \in C_{\text{sys}}, \\ \mathbf{h}_{\text{sub},k}(\mathbf{R}_{\text{sub},k}^{\text{sub}}, \mathbf{x}_{\text{sub},k}, \mathbf{y}_{\text{sub},k}^{\text{sub}}) &= \mathbf{0} \quad \forall k \in C_{\text{sys}} \end{aligned}$$

where

$$\begin{aligned} \lambda^T &= [(\lambda^{\text{R}})^T, (\lambda^{\text{y}})^T], \\ (\lambda^{\text{R}})^T &= [(\lambda_k^{\text{R}})^T \quad \forall k \in C_{\text{sys}}], \\ (\lambda^{\text{y}})^T &= [(\lambda_k^{\text{y}})^T \quad \forall k \in C_{\text{sys}}] \end{aligned}$$

The AIO–LP problem can be decomposed into the system and subsystems in a different way than that given in Lassiter et al. [15], who retained the system responses and linking variables as variables at the system level and kept their copies as variables at the subsystem level. On the contrary, we retain the responses and linking variables as well as their duplicates at each level. We treat one of the two entities as variables and the other entity as targets of one level and exchange the roles of the two entities at the other level. The system and subsystem problems are as follows:

SYS–LP:

$$\begin{aligned} \min_{\mathbf{x}_{\text{sys}}, \mathbf{R}_{\text{sub}}^{\text{sys}}, \mathbf{y}_{\text{sub}}^{\text{sub}}} f_{\text{sys}} &= \|\mathbf{R}_{\text{sys}} - \mathbf{T}_{\text{sys}}\|_2^2 + \sum_{k \in C_{\text{sys}}} (\lambda_k^{\text{R}})^T (\mathbf{R}_{\text{sub},k}^{\text{sys}} - \mathbf{R}_{\text{sub},k}^{\text{sub}}) \\ &+ \sum_{k \in C_{\text{sys}}} (\lambda_k^{\text{y}})^T (\mathbf{y}_{\text{sub},k}^{\text{sys}} - \mathbf{y}_{\text{sub},k}^{\text{sub}}) \end{aligned} \quad (8)$$

where

$$\mathbf{R}_{\text{sys}} = \mathbf{R}_{\text{sys}}(\mathbf{x}_{\text{sys}}, \mathbf{R}_{\text{sub}}^{\text{sys}})$$

subject to

$$\mathbf{g}_{\text{sys}}(\mathbf{R}_{\text{sys}}, \mathbf{x}_{\text{sys}}) \leq \mathbf{0}, \quad \mathbf{h}_{\text{sys}}(\mathbf{R}_{\text{sys}}, \mathbf{x}_{\text{sys}}) = \mathbf{0}$$

The k th subsystem problem is stated as follows:

SYS–LP _{k} :

$$\begin{aligned} \min_{\mathbf{R}_{\text{sub},k}^{\text{sub}}, \mathbf{x}_{\text{sub},k}, \mathbf{y}_{\text{sub},k}^{\text{sub}}} f_{\text{sub},k} &= (\lambda_k^{\text{R}})^T (\mathbf{R}_{\text{sub},k}^{\text{sys}} - \mathbf{R}_{\text{sub},k}^{\text{sub}}) \\ &+ (\lambda_k^{\text{y}})^T (\mathbf{y}_{\text{sub},k}^{\text{sys}} - \mathbf{y}_{\text{sub},k}^{\text{sub}}) \end{aligned} \quad (9)$$

where

$$\mathbf{R}_{\text{sub},k}^{\text{sub}} = \mathbf{R}_{\text{sub},k}^{\text{sub}}(\mathbf{x}_{\text{sub},k}, \mathbf{y}_{\text{sub},k}^{\text{sub}})$$

subject to

$$\begin{aligned} \mathbf{g}_{\text{sub},k}(\mathbf{R}_{\text{sub},k}, \mathbf{x}_{\text{sub},k}, \mathbf{y}_{\text{sub},k}) &\leq \mathbf{0}, \\ \mathbf{h}_{\text{sub},k}(\mathbf{R}_{\text{sub},k}, \mathbf{x}_{\text{sub},k}, \mathbf{y}_{\text{sub},k}) &= \mathbf{0} \end{aligned}$$

The corresponding Lagrangian dual problems for the system and subsystem problems are stated as follows:

SYS–DP:

$$\max_{\lambda} \phi_{\text{SYS}}(\lambda^{\text{R}}, \lambda^{\text{y}}) \quad (10)$$

where

$$\begin{aligned} \phi_{\text{SYS}}(\lambda^{\text{R}}, \lambda^{\text{y}}) &= \min_{\mathbf{x}_{\text{sys}}, \mathbf{R}_{\text{sub}}^{\text{sys}}, \mathbf{y}_{\text{sub}}^{\text{sub}}} f_{\text{sys}} \\ &= \|\mathbf{R}_{\text{sys}} - \mathbf{T}_{\text{sys}}\|_2^2 + \sum_{k \in C_{\text{sys}}} (\lambda_k^{\text{R}})^T (\mathbf{R}_{\text{sub},k}^{\text{sys}} - \mathbf{R}_{\text{sub},k}^{\text{sub}}) \\ &+ \sum_{k \in C_{\text{sys}}} (\lambda_k^{\text{y}})^T (\mathbf{y}_{\text{sub},k}^{\text{sys}} - \mathbf{y}_{\text{sub},k}^{\text{sub}}) \end{aligned}$$

where

$$\mathbf{R}_{\text{sys}} = \mathbf{R}_{\text{sys}}(\mathbf{x}_{\text{sys}}, \mathbf{R}_{\text{sub}}^{\text{sys}})$$

subject to

$$\mathbf{g}_{\text{sys}}(\mathbf{R}_{\text{sys}}, \mathbf{x}_{\text{sys}}) \leq \mathbf{0}, \quad \mathbf{h}_{\text{sys}}(\mathbf{R}_{\text{sys}}, \mathbf{x}_{\text{sys}}) = \mathbf{0}$$

SUB–DP _{k} :

$$\max_{\lambda_k} \phi_{\text{SUB},k}(\lambda_k^{\text{R}}, \lambda_k^{\text{y}}) \quad (11)$$

where

$$\begin{aligned} \phi_{\text{SUB},k}(\lambda_k^{\text{R}}, \lambda_k^{\text{y}}) &= \min_{\mathbf{R}_{\text{sub},k}^{\text{sub}}, \mathbf{x}_{\text{sub},k}, \mathbf{y}_{\text{sub},k}^{\text{sub}}} f_{\text{sub},k} \\ &= (\lambda_k^{\text{R}})^T (\mathbf{R}_{\text{sub},k}^{\text{sys}} - \mathbf{R}_{\text{sub},k}^{\text{sub}}) + (\lambda_k^{\text{y}})^T (\mathbf{y}_{\text{sub},k}^{\text{sys}} - \mathbf{y}_{\text{sub},k}^{\text{sub}}) \end{aligned}$$

where

$$\mathbf{R}_{\text{sub},k}^{\text{sub}} = \mathbf{R}_{\text{sub},k}^{\text{sub}}(\mathbf{x}_{\text{sub},k}, \mathbf{y}_{\text{sub},k}^{\text{sub}})$$

subject to

$$\begin{aligned} \mathbf{g}_{\text{sub},k}(\mathbf{R}_{\text{sub},k}, \mathbf{x}_{\text{sub},k}, \mathbf{y}_{\text{sub},k}) &\leq \mathbf{0}, \\ \mathbf{h}_{\text{sub},k}(\mathbf{R}_{\text{sub},k}, \mathbf{x}_{\text{sub},k}, \mathbf{y}_{\text{sub},k}) &= \mathbf{0} \end{aligned}$$

where

$$(\lambda_k)^T = [(\lambda_k^{\text{R}})^T, (\lambda_k^{\text{y}})^T] \quad \forall k \in C_{\text{sys}}$$

Note that the responses and linking variables with superscript sys become variable when solving the SYS–LP, whereas those with sub become constant. Similarly, when solving SUB–LP, those with superscript sub become variable, whereas those with superscript sys become constant. In other words, when solving SYS–LP, the values from the subsystem problems become parameters, and when solving SUB–LP, the values from the system problem become parameters.

IV. Lagrangian Dual Coordination Algorithm for ATC

In this section it is first shown that the proposed ATC coordination is equivalent to solving the AIO problem within a framework of Lagrangian duality. In this framework, the all-in-one dual problem (AIO–DP) is solved, which implicitly solves the AIO–LP and also the AIO problem. In the proposed ATC dual coordination, SYS–DP and SUB–DPs are solved, which implicitly solves SYS–LP and SUB–LPs and also the AIO problem. The coordination is achieved as a result of an iterative process in which the Lagrange multipliers are

updated in a systematic way [12,13,19]. As a result, the coordination algorithm also finds the optimal weights for the deviation terms. Subgradient optimization [12] is used for solving SYS-DP and SUB-DPs. The deviation terms resulted from the Lagrangian relaxation are augmented to follow the original ATC formulation while maintaining the convergence of the coordination.

A. Equivalence Between the ATC Dual Coordination and the AIO Problem

The goal of this section is to show that the proposed ATC dual coordination is equivalent to solving the ATC-decomposed AIO problem, that is, the coordination leads to an optimal solution of the AIO problem due to the partial duality. Following the theorem on Lagrangian duality, if the AIO problem has a local optimal solution with the optimal value f^* , so does the AIO-DP, that is, there exists λ^* , a local optimal solution of the AIO-DP, whose value $\phi_{\text{AIO}}(\lambda^*)$ is equal to f^* . In the case where the objective and constraint functions are convex/affine, there exists a global optimal solution of the AIO problem and a corresponding global optimal solution (i.e., optimal Lagrange multipliers) of the AIO-DP with the same objective values.

A subgradient-based solution process for solving the AIO-DP starts with choosing an initial Lagrange multiplier vector $\lambda^0 = (\lambda^{R,0}, \lambda^{y,0})$. The AIO-LP is then solved with respect to the design variables $\mathbf{x}_{\text{sys}}, \mathbf{R}_{\text{sub}}^{\text{sys}}, \mathbf{R}_{\text{sub}}^{\text{sub}}, \mathbf{x}_{\text{sub}}, \mathbf{y}_{\text{sub}}$, and $\mathbf{y}_{\text{sub}}^{\text{sub}}$ for the fixed $\lambda^{R,0}$ and $\lambda^{y,0}$. To update the Lagrange multiplier, the subgradient of $\phi_{\text{AIO}}(\lambda)$ with respect to λ is calculated according to the following:

$$\begin{pmatrix} \xi^{\mathbf{R},i} \\ \xi^{\mathbf{y},i} \end{pmatrix} = \begin{pmatrix} \frac{\partial \phi_{\text{AIO}}}{\partial \lambda^{\mathbf{R},i}} \\ \frac{\partial \phi_{\text{AIO}}}{\partial \lambda^{\mathbf{y},i}} \end{pmatrix} = \begin{pmatrix} \mathbf{R}_{\text{sub}}^{\text{sys},i} - \mathbf{R}_{\text{sub}}^{\text{sub},i} \\ \mathbf{y}_{\text{sub}}^{\text{sys},i} - \mathbf{y}_{\text{sub}}^{\text{sub},i} \end{pmatrix} \quad (12)$$

where the index i denotes the i th iteration. Based on the subgradient information, the Lagrange multiplier vector is updated as follows:

$$\begin{pmatrix} \lambda^{\mathbf{R},i+1} \\ \lambda^{\mathbf{y},i+1} \end{pmatrix} = \begin{pmatrix} \lambda^{\mathbf{R},i} \\ \lambda^{\mathbf{y},i} \end{pmatrix} + \alpha \begin{pmatrix} \xi^{\mathbf{R},i} \\ \xi^{\mathbf{y},i} \end{pmatrix} \quad (13)$$

where α , the step size, is a positive scalar. With a proper step size, the AIO-DP should converge to an optimal solution λ^* whose value $\phi_{\text{AIO}}(\lambda^*)$ is equal to f^* , the optimal value of the AIO problem. Setting a proper step size is investigated in Sec. V.

The proposed dual coordination for bilevel ATC is given by the following algorithm.

1. Algorithm: Bilevel ATC Dual Coordination

- 1) Set $i = 0$. Choose initial Lagrange multipliers $\lambda_k^{\mathbf{R},i}$ and $\lambda_k^{\mathbf{y},i}$ for the deviation terms, initial subsystem responses, and linking variables $\mathbf{R}_{\text{sub},k}^{\text{sub},i}$ and $\mathbf{y}_{\text{sub},k}^{\text{sub},i} \forall k \in C_{\text{sys}}$.
- 2) Solve SYS-LP with respect to $\mathbf{x}_{\text{sys}}, \mathbf{R}_{\text{sub},k}^{\text{sys},i}$ and $\mathbf{y}_{\text{sub},k}^{\text{sys},i} \forall k \in C_{\text{sys}}$.
- 3) Pass optimal $\mathbf{R}_{\text{sub},k}^{\text{sys},i}$ and $\mathbf{y}_{\text{sub},k}^{\text{sys},i}$ to SUB-LP $_k \forall k \in C_{\text{sys}}$.
- 4) Solve SUB-LP $_k$ with respect to $\mathbf{x}_{\text{sub},k}^i, \mathbf{R}_{\text{sub},k}^{\text{sub},i}$ and $\mathbf{y}_{\text{sub},k}^{\text{sub},i} \forall k \in C_{\text{sys}}$.
- 5) Calculate the deviations for the optimal responses and linking variables

$$\begin{pmatrix} \xi_k^{\mathbf{R},i} \\ \xi_k^{\mathbf{y},i} \end{pmatrix} = \begin{pmatrix} \mathbf{R}_{\text{sub},k}^{\text{sys},i} - \mathbf{R}_{\text{sub},k}^{\text{sub},i} \\ \mathbf{y}_{\text{sub},k}^{\text{sys},i} - \mathbf{y}_{\text{sub},k}^{\text{sub},i} \end{pmatrix} \quad \forall k \in C_{\text{sys}} \text{ if } \|(\xi_k^{\mathbf{R},i}, \xi_k^{\mathbf{y},i})\| < (\text{tolerance}), \text{ stop}$$

- 6) Pass optimal $\mathbf{R}_{\text{sub},k}^{\text{sub},i}$ and $\mathbf{y}_{\text{sub},k}^{\text{sub},i} \forall k \in C_{\text{sys}}$ to SYS-LP.
- 7) Update the Lagrange multipliers

$$\begin{pmatrix} \lambda_k^{\mathbf{R},i+1} \\ \lambda_k^{\mathbf{y},i+1} \end{pmatrix} = \begin{pmatrix} \lambda_k^{\mathbf{R},i} \\ \lambda_k^{\mathbf{y},i} \end{pmatrix} + \alpha_k \begin{pmatrix} \xi_k^{\mathbf{R},i} \\ \xi_k^{\mathbf{y},i} \end{pmatrix} \quad \forall k \in C_{\text{sys}}$$

where $\alpha_k > 0$ is a step size.

- 8) Set $i = i + 1$ and go to step 2.

The proposed algorithm is different from algorithm 1 proposed by Lassiter et al. [15]. In the latter, the system and subsystem problems

are solved independently (or in parallel) for the responses and linking variables without copying their optimal values from one problem to the other problem. This is equivalent to solving their overall Lagrangian dual problem with the deviation constraints whose terms are both of the optimization variables. These constraints cannot be treated with the augmented Lagrangian approach because squaring their difference makes the augmented Lagrangian nonseparable. The newly proposed algorithm eliminates this major difficulty. Because the system and subsystem problems are solved consecutively and the optimal values of responses and linking variables are passed from one system to the other, the deviation constraints in the associated AIO-DP problem in Eq. (7) have only one term as the optimization variable and the other term constant. These constraints therefore produce a separable augmented Lagrangian (see Sec. IV.B).

The next result shows that solving the dual subproblems SYS-DP and SUB-DP $_k \forall k \in C_{\text{sys}}$ is equivalent to solving the AIO-DP.

2. Proposition: ATC Dual Coordination Solves the AIO Problem

Proof: We first show that solving the system-level Lagrangian problem and the subsystem-level Lagrangian subproblems with appropriately updated Lagrange multipliers is equivalent to solving the AIO-DP problem. In the coordination algorithm, the Lagrange multipliers for the deviation terms in Eq. (8) are first set as $\lambda^{R,0}$ and $\lambda^{y,0}$. The system and subsystem Lagrangian subproblems are then solved *sequentially*, that is, after solving the system-level problem SYS-LP with respect to its primal variables, the process proceeds with solving the subsystem-level subproblems with respect to their primal variables. In case there exists more than one subproblem, they can be solved simultaneously as the variables at the subsystem level are separable. After solving SYS-LP and SUB-LPs for each subsystem problem, the subgradient of the associated dual problem objective function is calculated with respect to the Lagrange multipliers. In the i th iteration, this subgradient is

$$\begin{pmatrix} \xi_k^{\mathbf{R},i} \\ \xi_k^{\mathbf{y},i} \end{pmatrix}_{\text{sub}} = \begin{pmatrix} \frac{\partial \phi_{\text{sub},k}}{\partial \lambda_k^{\mathbf{R},i}} \\ \frac{\partial \phi_{\text{sub},k}}{\partial \lambda_k^{\mathbf{y},i}} \end{pmatrix} = \begin{pmatrix} \mathbf{R}_{\text{sub},k}^{\text{sys},i} - \mathbf{R}_{\text{sub},k}^{\text{sub},i} \\ \mathbf{y}_{\text{sub},k}^{\text{sys},i} - \mathbf{y}_{\text{sub},k}^{\text{sub},i} \end{pmatrix} \quad (14)$$

where the responses and linking variables from the system level $\mathbf{R}_{\text{sub},k}^{\text{sys},i}$ and $\mathbf{y}_{\text{sub},k}^{\text{sys},i}$ are fixed parameters found earlier in the same iteration as optimal solutions for the SYS-LP. The right-hand side of Eq. (14), calculated for all subsystems, is identical to the subgradient of the AIO-DP in Eq. (12) and to the subgradient calculated in step 5 of the algorithm. By solving the system and subsystem problems sequentially, the subgradient is partially updated for the current design values at each level, and the calculation in step 5 yields the subgradient of the AIO-DP problem. Based on the subgradient, the Lagrange multipliers are updated in step 7 of the algorithm, which is again equivalent to the update for the AIO-DP in Eq. (13). Thus the sequence of solving the system-level problem and the subsystem-level subproblems is equivalent to solving the AIO-DP. Finally, by Lagrangian duality theory, solving the AIO-DP is equivalent to solving the AIO-LP and also the AIO, QED

B. Implementation with the Augmented Lagrangian

The ATC-introduced deviation terms for responses and linking variables in the objective of Eq. (6) are defined without a norm, which is a departure from the traditional ATC formulation. The type of norm for the deviation terms can significantly affect the convergence of the ATC dual coordination algorithm. Conventionally for ATC, the squared $L-2$ norm has been used. In this section, it is shown that the squared $L-2$ norm may cause numerical instability and at the same time it is demonstrated that augmented deviation terms introduced in the proposed ATC dual formulation prevent this instability.

It is observed that the coordination algorithm for the AIO problem avoids numerical instability around the optimal solution when the norm has been dropped in the deviation terms. For example, the squared $L-2$ norm makes the Lagrange multipliers associated with the relaxed deviation constraints large. Consider a simple case of

system-level problem with no feasibility constraint written as follows:

$$\min_{x_1} f_{\text{sys}} = (x_1 - T)^2 + \lambda(x_1 - x_1^l)^2$$

where x_1^l is the lower system response. The stationarity condition comes with

$$\begin{aligned} \frac{\partial f_{\text{sys}}}{\partial x_1} &= \frac{\partial}{\partial x_1} (x_1 - T)^2 + 2\lambda(x_1 - x_1^l) = 0, \\ \lambda &= -\frac{1}{2} \left(\frac{\partial}{\partial x_1} (x_1 - T)^2 \right) (x_1 - x_1^l)^{-1} \end{aligned}$$

With the zero deviation for the lower-level response, that is, $x_1 = x_1^l$, the Lagrange multiplier corresponding to the relaxed deviation constraint becomes negative infinity assuming that the target T is not attainable. On the other hand, by dropping the squared L -2 norm for the deviation term, the system-level problem becomes

$$\min_{x_1} f_{\text{sys}} = (x_1 - T)^2 + \lambda(x_1 - x_1^l)$$

The corresponding stationary condition is

$$\frac{\partial f_{\text{sys}}}{\partial x_1} = \frac{\partial}{\partial x_1} (x_1 - T)^2 + \lambda = 0, \quad \lambda = -\frac{\partial}{\partial x_1} (x_1 - T)^2$$

which prevents the corresponding Lagrange multiplier from becoming infinity. However, dropping the squared L -2 norm does not follow the traditional ATC formulation exactly and makes subsystem-level subproblem objectives in the ATC formulation unbounded. This difficulty will be resolved with the augmented Lagrangian.

For the proposed Lagrangian dual coordination, the augmented Lagrangian formulation is adopted to further convexify the Lagrangian. The penalty constants \mathbf{c} in the augmented terms in Eq. (16) must be positive. One instance is to define them as

$$\mathbf{c}_k^{\mathbf{R}} = (c_{k,1}^{\mathbf{R}}, \dots, c_{k,p(k)}^{\mathbf{R}}) = (\sqrt{|\lambda_{k,1}^{\mathbf{R}}|}, \dots, \sqrt{|\lambda_{k,p(k)}^{\mathbf{R}}|})$$

and

$$\mathbf{c}_k^{\mathbf{y}} = (c_{k,1}^{\mathbf{y}}, \dots, c_{k,r(k)}^{\mathbf{y}}) = (\sqrt{|\lambda_{k,1}^{\mathbf{y}}|}, \dots, \sqrt{|\lambda_{k,r(k)}^{\mathbf{y}}|})$$

where $p(k)$ and $r(k)$ are the numbers of responses and linking variables for subsystem k , respectively.

AIO-LP:

$$\begin{aligned} \min_{\mathbf{R}_{\text{sub}}, \mathbf{R}_{\text{sub}}, \mathbf{x}_{\text{sub}}, \mathbf{y}_{\text{sub}}} f_{\text{AIO}} &= \|\mathbf{R}_{\text{sys}} - \mathbf{T}_{\text{sys}}\|_2^2 \\ &+ \sum_{k \in C_{\text{sys}}} (\boldsymbol{\lambda}_k^{\mathbf{R}})^T (\mathbf{R}_{\text{sub},k}^{\text{sys}} - \mathbf{R}_{\text{sub},k}^{\text{sub}}) + \sum_{k \in C_{\text{sys}}} (\boldsymbol{\lambda}_k^{\mathbf{y}})^T (\mathbf{y}_{\text{sub},k}^{\text{sys}} - \mathbf{y}_{\text{sub},k}^{\text{sub}}) \\ &+ \sum_{k \in C_{\text{sys}}} \|(\mathbf{c}_k^{\mathbf{R}}) \circ (\mathbf{R}_{\text{sub},k}^{\text{sys}} - \mathbf{R}_{\text{sub},k}^{\text{sub}})\|_2^2 \\ &+ \sum_{k \in C_{\text{sys}}} (\mathbf{c}_k^{\mathbf{y}}) \circ (\mathbf{y}_{\text{sub},k}^{\text{sys}} - \mathbf{y}_{\text{sub},k}^{\text{sub}})\|_2^2 \end{aligned} \quad (15)$$

where

$$\begin{aligned} \mathbf{R}_{\text{sys}} &= \mathbf{R}_{\text{sys}}(\mathbf{x}_{\text{sys}}, \mathbf{R}_{\text{sub}}^{\text{sys}}), \\ \mathbf{R}_{\text{sub},k}^{\text{sub}} &= \mathbf{R}_{\text{sub},k}^{\text{sub}}(\mathbf{x}_{\text{sub},k}, \mathbf{y}_{\text{sub},k}^{\text{sub}}) \quad \forall k \in C_{\text{sys}} \end{aligned}$$

subject to

$$\begin{aligned} \mathbf{g}_{\text{sys}}(\mathbf{R}_{\text{sys}}, \mathbf{x}_{\text{sys}}) &\leq \mathbf{0}, \quad \mathbf{h}_{\text{sys}}(\mathbf{R}_{\text{sys}}, \mathbf{x}_{\text{sys}}) = \mathbf{0}, \\ \mathbf{g}_{\text{sub},k}(\mathbf{R}_{\text{sub},k}^{\text{sub}}, \mathbf{x}_{\text{sub},k}, \mathbf{y}_{\text{sub},k}^{\text{sub}}) &\leq \mathbf{0} \quad \forall k \in C_{\text{sys}}, \\ \mathbf{h}_{\text{sub},k}(\mathbf{R}_{\text{sub},k}^{\text{sub}}, \mathbf{x}_{\text{sub},k}, \mathbf{y}_{\text{sub},k}^{\text{sub}}) &= \mathbf{0} \quad \forall k \in C_{\text{sys}} \end{aligned}$$

where the symbol $\|\cdot\|_2^2$ indicates the squared L -2 norm.

The modified SYS-LP and SUB-LP are shown as follows:

SYS-LP:

$$\begin{aligned} \min_{\mathbf{R}_{\text{sub}}, \mathbf{R}_{\text{sub}}, \mathbf{y}_{\text{sub}}} f_{\text{sys}} &= \|\mathbf{R}_{\text{sys}} - \mathbf{T}_{\text{sys}}\|_2^2 + \sum_{k \in C_{\text{sys}}} (\boldsymbol{\lambda}_k^{\mathbf{R}})^T (\mathbf{R}_{\text{sub},k}^{\text{sys}} - \mathbf{R}_{\text{sub},k}^{\text{sub}}) \\ &+ \sum_{k \in C_{\text{sys}}} (\boldsymbol{\lambda}_k^{\mathbf{y}})^T (\mathbf{y}_{\text{sub},k}^{\text{sys}} - \mathbf{y}_{\text{sub},k}^{\text{sub}}) \\ &+ \sum_{k \in C_{\text{sys}}} \|(\mathbf{c}_k^{\mathbf{R}}) \circ (\mathbf{R}_{\text{sub},k}^{\text{sys}} - \mathbf{R}_{\text{sub},k}^{\text{sub}})\|_2^2 \\ &+ \sum_{k \in C_{\text{sys}}} (\mathbf{c}_k^{\mathbf{y}}) \circ (\mathbf{y}_{\text{sub},k}^{\text{sys}} - \mathbf{y}_{\text{sub},k}^{\text{sub}})\|_2^2 \end{aligned} \quad (16)$$

where

$$\mathbf{R}_{\text{sys}} = \mathbf{R}_{\text{sys}}(\mathbf{x}_{\text{sys}}, \mathbf{R}_{\text{sub}}^{\text{sys}})$$

subject to

$$\mathbf{g}_{\text{sys}}(\mathbf{R}_{\text{sys}}, \mathbf{x}_{\text{sys}}) \leq \mathbf{0}, \quad \mathbf{h}_{\text{sys}}(\mathbf{R}_{\text{sys}}, \mathbf{x}_{\text{sys}}) = \mathbf{0}$$

SUB-LP:

$$\begin{aligned} \min_{\mathbf{R}_{\text{sub},k}, \mathbf{y}_{\text{sub},k}} f_{\text{sub},k} &= (\boldsymbol{\lambda}_k^{\mathbf{R}})^T (\mathbf{R}_{\text{sub},k}^{\text{sys}} - \mathbf{R}_{\text{sub},k}^{\text{sub}}) \\ &+ (\boldsymbol{\lambda}_k^{\mathbf{y}})^T (\mathbf{y}_{\text{sub},k}^{\text{sys}} - \mathbf{y}_{\text{sub},k}^{\text{sub}}) + \|(\mathbf{c}_k^{\mathbf{R}}) \circ (\mathbf{R}_{\text{sub},k}^{\text{sys}} - \mathbf{R}_{\text{sub},k}^{\text{sub}})\|_2^2 \\ &+ \|(\mathbf{c}_k^{\mathbf{y}}) \circ (\mathbf{y}_{\text{sub},k}^{\text{sys}} - \mathbf{y}_{\text{sub},k}^{\text{sub}})\|_2^2 \end{aligned} \quad (17)$$

where

$$\mathbf{R}_{\text{sub},k}^{\text{sub}} = \mathbf{R}_{\text{sub},k}^{\text{sub}}(\mathbf{x}_{\text{sub},k}, \mathbf{y}_{\text{sub},k}^{\text{sub}})$$

subject to

$$\mathbf{g}_{\text{sub},k}(\mathbf{R}_{\text{sub},k}^{\text{sub}}, \mathbf{x}_{\text{sub},k}, \mathbf{y}_{\text{sub},k}^{\text{sub}}) \leq \mathbf{0}, \quad \mathbf{h}_{\text{sub},k}(\mathbf{R}_{\text{sub},k}^{\text{sub}}, \mathbf{x}_{\text{sub},k}, \mathbf{y}_{\text{sub},k}^{\text{sub}}) = \mathbf{0}$$

These additional terms with the norm make the SYS-LP and SUB-LPs follow the conventional ATC formulation by adding squared L -2 norm terms and prevent the SUB-LPs from being unbounded while at the same time they keep the augmented Lagrangian of Eq. (15) separable. When updating the Lagrange multipliers by the subgradient calculation in Eq. (14), the augmented subgradient includes the squared term of the deviation due to the penalty terms introduced in Eqs. (16) and (17):

$$\begin{aligned} \begin{pmatrix} \boldsymbol{\xi}_k^{\mathbf{R},i} \\ \boldsymbol{\xi}_k^{\mathbf{y},i} \end{pmatrix} &= \begin{pmatrix} \frac{\partial \phi_{\text{AIO}}}{\partial \boldsymbol{\lambda}_k^{\mathbf{R},i}} \\ \frac{\partial \phi_{\text{AIO}}}{\partial \boldsymbol{\lambda}_k^{\mathbf{y},i}} \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{R}_{\text{sub},k}^{\text{sys},i} - \mathbf{R}_{\text{sub},k}^{\text{sub},i} \\ \mathbf{y}_{\text{sub},k}^{\text{sys},i} - \mathbf{y}_{\text{sub},k}^{\text{sub},i} \end{pmatrix} + \begin{pmatrix} \nabla_{\boldsymbol{\lambda}_k^{\mathbf{R},i}} (\|(\mathbf{c}_k^{\mathbf{R}}) \circ (\mathbf{R}_{\text{sub},k}^{\text{sys}} - \mathbf{R}_{\text{sub},k}^{\text{sub}})\|_2^2) \\ \nabla_{\boldsymbol{\lambda}_k^{\mathbf{y},i}} (\|(\mathbf{c}_k^{\mathbf{y}}) \circ (\mathbf{y}_{\text{sub},k}^{\text{sys}} - \mathbf{y}_{\text{sub},k}^{\text{sub}})\|_2^2) \end{pmatrix} \\ &\approx \begin{pmatrix} \mathbf{R}_{\text{sub},k}^{\text{sys},i} - \mathbf{R}_{\text{sub},k}^{\text{sub},i} \\ \mathbf{y}_{\text{sub},k}^{\text{sys},i} - \mathbf{y}_{\text{sub},k}^{\text{sub},i} \end{pmatrix} \quad \forall k \in C_{\text{sys}} \end{aligned} \quad (18)$$

Because the deviation terms $(\mathbf{R}_{\text{sub},k}^{\text{sys}} - \mathbf{R}_{\text{sub},k}^{\text{sub}})$ and $(\mathbf{y}_{\text{sub},k}^{\text{sys}} - \mathbf{y}_{\text{sub},k}^{\text{sub}})$ tend to become very small (i.e., close to zero), the squared terms can be dropped in Eq. (18) in numerical implementation so that step 5 of the algorithm remains valid.

V. Demonstration Examples

In this section, two example problems are solved by the proposed ATC Lagrangian dual coordination: a convex quadratic programming problem and a nonconvex geometric programming problem. It is demonstrated that for each problem, the solution process converges to an optimal solution with the corresponding Lagrange multipliers that are identical to those for the original problem. Although the proposed method is based on the convexity assumption, the geometric programming example demonstrates the efficiency of the proposed method even in nonconvex cases.

A. Convex Quadratic Programming Problem

The original convex minimization problem with a unique global optimal solution is stated as follows:

P_{original} :

$$\min_{x_1, x_2, \dots, x_{18}} x_1^2 + x_2^2 \quad (19)$$

subject to

$$\begin{aligned} g_1 &= -x_3 + x_4 - x_5 + 2 \leq 0, & g_2 &= x_5 - x_6 - x_7 + 1 \leq 0, \\ g_3 &= x_8 + x_9 - x_{11} + 1 \leq 0, & g_4 &= -x_8 + x_{10} - x_{11} + 1 \leq 0, \\ g_5 &= x_{11} - x_{12} - x_{13} + 2 \leq 0, & g_6 &= x_{11} + x_{12} - x_{14} + 1 \leq 0, \\ h_1 &= x_3 + x_4 + x_5 - x_1 + 1 = 0, \\ h_2 &= x_5 + x_6 + x_7 - x_2 + 1 = 0, \\ h_3 &= x_8 - x_9 - x_{10} + x_{11} - x_3 + 1 = 0, \\ h_4 &= x_{11} + x_{12} + x_{13} + x_{14} - x_6 = 0, & 0 &\leq x_3, x_4, \dots, x_{18} \end{aligned}$$

For ATC demonstration, copies of variables x_3 , x_6 , and x_{11} are introduced in the equality constraints h_5, \dots, h_8 . A global optimal solution to the preceding problem is found at

$$\mathbf{x}^* = (5, 10, 4, 0, 0, 9, 0, 1, 0, 0, 2, 0, 4, 3, 4, 2, 9, 2)$$

and the corresponding Lagrange multipliers for the deviation constraints h_5, \dots, h_8 are $\boldsymbol{\lambda}^* = (10, 20, 60, 60)$. Following the AIO-LP in Eq. (15), the deviation constraints h_5, \dots, h_8 can be relaxed and moved to the objective function as follows:

$P_{\text{AIO-LP}}$:

$$\begin{aligned} \min_{x_1, x_2, \dots, x_{18}} & x_1^2 + x_2^2 + \lambda_1 h_5 + \lambda_2 h_6 + \lambda_3 h_7 + \lambda_4 h_8 \\ & + |\lambda_1|(h_5)^2 + |\lambda_2|(h_6)^2 + |\lambda_3|(h_7)^2 + |\lambda_4|(h_8)^2 \end{aligned} \quad (20)$$

subject to

$$\begin{aligned} g_1 &= -x_3 + x_4 - x_5 + 2 \leq 0, & g_2 &= x_5 - x_6 - x_7 + 1 \leq 0, \\ g_3 &= x_8 + x_9 - x_{11} + 1 \leq 0, & g_4 &= -x_8 + x_{10} - x_{11} + 1 \leq 0, \\ g_5 &= x_{11} - x_{12} - x_{13} + 2 \leq 0, & g_6 &= x_{11} + x_{12} - x_{14} + 1 \leq 0, \\ h_1 &= x_3 + x_4 + x_5 - x_1 + 1 = 0, \\ h_2 &= x_5 + x_6 + x_7 - x_2 + 1 = 0, \\ h_3 &= x_8 - x_9 - x_{10} + x_{11} - x_3 + 1 = 0, \\ h_4 &= x_{11} + x_{12} + x_{13} + x_{14} - x_6 = 0, & 0 &\leq x_3, x_4, \dots, x_{18} \end{aligned}$$

Following ATC formulations in Eqs. (16) and (17), the original problem is decomposed into one system-level subproblem and two subsystem-level subproblems as shown in Fig. 1. Two subsystems share a linking variable that is denoted as x_{16} in subsystem 1 and x_{18} in subsystem 2, that is, $x_{16} = x_{18}$. The linking variable is coordinated at the system level and denoted as x_{11} at that level, that is, $x_{11} = x_{16} = x_{18}$ at optimality. The linking variable is duplicated in each of the element design problems: the subsystem linking variable coordinated at the system level takes $y_{\text{sub}}^{\text{sys}} = x_{11}$, the subsystem

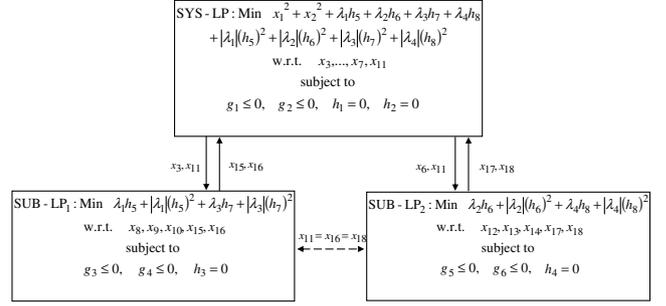


Fig. 1 The ATC-decomposed problem of the original problem.

linking variable in subsystem 1 takes $y_{\text{sub}}^{\text{sub1}} = x_{16}$, and the subsystem linking variable in subsystem 2 takes $y_{\text{sub}}^{\text{sub2}} = x_{18}$. Similarly, the subsystem responses at the system level are denoted as x_3 and x_6 and those at the subsystem level are denoted as x_{15} and x_{17} , respectively. In other words, for the same subsystem response, a duplicate variable is created at the system level: $R_{\text{sub1}}^{\text{sys}} = x_3$, $R_{\text{sub1}}^{\text{sub1}} = x_{15}$, $R_{\text{sub2}}^{\text{sys}} = x_6$, and $R_{\text{sub2}}^{\text{sub2}} = x_{17}$.

Following the dual ATC formulation, the equality constraints h_5, \dots, h_8 are relaxed and moved to the objective function at the system-level primal problem SYS-LP. The squared L_2 norm deviation terms are also added in the objective function of each subproblem according to formulations in Eqs. (16) and (17). According to the proposed coordination algorithm, the SYS-LP is solved with the initial values of the Lagrange multipliers. After solving SYS-LP in Eq. (16), the values for x_3 , x_6 , and x_{11} are cascaded to SUB-LP₁ and SUB-LP₂ in Eq. (17). The two subsystem problems are solved to achieve the minimum deviations from the target values from the system-level problem and the values for x_{15} , x_{16} , x_{17} , and x_{18} , as well as for the local variables, are obtained. After solving the subproblems at the system and subsystem levels, the Lagrange multipliers are updated based on the subgradient obtained according to Eq. (18).

The optimal solution obtained by the ATC Lagrangian dual coordination is shown in Tables 1–4. The coordination algorithm converges to the global optimal solution with the corresponding Lagrange multipliers that are assigned as weights to the deviation terms in the objective functions. The algorithm starts with $\boldsymbol{\lambda}_0 = [5, 15, 50, 50]$, $\mathbf{x}_0 = [0, 0, \dots, 0]$, $m = 5$ with the step size update following Eq. (21).

The proposed coordination algorithm is tested for several cases with different initial Lagrange multipliers and step sizes. Each of the cases shows different convergence history, however, they all converge to the same optimal solution. The step size for updating the Lagrange multipliers in step 7 of the algorithm is updated as follows [12]:

$$\alpha_k^i = \frac{1+m}{i+m} \cdot \frac{1}{\|\boldsymbol{\xi}_k^i\|} \quad \forall k \in C_{\text{sys}} \quad (21)$$

where i is the iteration number and m is a positive integer. In implementing $\|\boldsymbol{\xi}_k^i\|$, the L_2 norm was used. The preceding step size update is simple and, in contrast to other update methods, does not require prior knowledge of the final optimal objective value $\phi_{\text{AIO}}(\boldsymbol{\lambda}^*)$ of the dual problem. When the final optimal objective

Table 1 Summary of the ATC Lagrangian dual coordination for the system-level subproblem

ATC level	ATC variable/parameter	Type (variable: V ; parameter: P)	Variable/parameter	ATC optimal solution	Global optimal solution
System	\mathbf{R}_{sys}	Response (V)	x_1, x_2	(5.0, 9.99)	(5.0, 10.0)
	\mathbf{x}_{sys}	Local variable (V)	x_4, x_5, x_7	(0.0, 0)	(0.0, 0)
	$\mathbf{R}_{\text{sub}}^{\text{sys}} = (R_{\text{sub1}}^{\text{sys}}, R_{\text{sub2}}^{\text{sys}})$	Subsystem responses (V)	x_3, x_6	(4.0, 8.99)	(4.0, 9.0)
	$\mathbf{y}_{\text{sub}}^{\text{sys}}$	Subsystem linking variable (V)	x_{11}	2.0	2.0
	$\mathbf{R}_{\text{sub}}^{\text{sub}} = (R_{\text{sub1}}^{\text{sub}}, R_{\text{sub2}}^{\text{sub}})$	Subsystem response (P)	x_{15}, x_{17}	(4.0, 8.99)	(4.0, 9.0)
	$\mathbf{y}_{\text{sub}}^{\text{sub}} = (y_{\text{sub1}}^{\text{sub}}, y_{\text{sub2}}^{\text{sub}})$	Subsystem linking variable (P)	x_{16}, x_{18}	(2.0, 2.0)	(2.0, 2.0)

Table 2 Summary of the ATC Lagrangian dual coordination for the subsystem-level subproblem 1

ATC level	ATC variable/parameter	Type (variable: V ; parameter: P)	Variable/parameter	ATC optimal solution	Global optimal solution
Subsystem 1	R_{sub1}^{sys}	Subsystem response target (P)	x_3	4.0	4.0
	R_{sub1}^{sub}	Subsystem 1 response (V)	x_{15}	4.0	4.0
	y_{sub1}^{sys}	Subsystem 1 linking variable target (P)	x_{11}	2.0	2.0
	y_{sub1}^{sub}	Subsystem 1 linking variable (V)	x_{16}	2.0	2.0
	x_{sub1}	Subsystem 1 local variable (V)	x_8, x_9, x_{10}	(1.0,0,0)	(1.0,0,0)

Table 3 Summary of the ATC Lagrangian dual coordination for the subsystem-level subproblem 2

ATC level	ATC variable/parameter	Type (variable: V ; parameter: P)	Variable/parameter	ATC optimal solution	Global optimal solution
Subsystem 2	R_{sub2}^{sys}	Subsystem 2 response target (P)	x_6	8.99	9.0
	R_{sub2}^{sub}	Subsystem 2 response (V)	x_{17}	8.99	9.0
	y_{sub2}^{sys}	Subsystem 2 linking variable target (P)	x_{11}	2.0	2.0
	y_{sub2}^{sub}	Subsystem 2 linking variable (V)	x_{18}	2.0	2.0
	x_{sub2}	Subsystem 2 local variable (V)	x_{12}, x_{13}, x_{14}	(4,0,3.0)	(0,4,0,3.0)

Table 4 Optimal Lagrange multipliers for the deviation terms

Lagrange multiplier	ATC optimal solution	Global optimal solution
λ_1	9.98	10.0
λ_2	20.0	20.0
λ_3	60.0	60.0
λ_4	60.0	60.0

value is known (or can be estimated), an alternative step size update method in Eq. (22) can be used [12].

$$\alpha_k^i = [\phi_{AIO}(\lambda_k^*) - \phi_{AIO}(\lambda_k^i)] \cdot \frac{1}{\|\xi_k^i\|} \quad \forall k \in C_{sys} \quad (22)$$

Figure 2 plots the deviation vs primal objective function for the AIO–LP problem in Eq. (15), that is, it plots the sum of all deviations in the objective function of Eq. (15) against the objective value of the primal problem given in Eq. (5). During the iteration process, the deviation becomes smaller until it reaches the zero value, that is, system and subsystem designs are consistent with zero deviation for which the corresponding primal and dual objective value of 125 is obtained.

Minimization of the primal objective value and the deviation can be modeled as a biobjective minimization problem for which Eq. (15) represents the related weighted-sum formulation in which the weight of the first (primal) objective is always equal to one and the weight of the other objective (the sum of all four deviation terms) is related to the values of Lagrange multipliers λ_k^R with $k \in C_{sys}$ and λ_k^Y with $k \in C_{sys}$. Because all the Lagrange multipliers are positive, so are the weights used when solving the weighted-sum problem in every iteration of the proposed algorithm. Hence, during the execution of the algorithm, the Pareto curve of the biobjective problem is generated (Fig. 2) and each iteration point corresponds to one of the Pareto solutions related to the AIO–LP problem. Additionally, as the iterations progress, the primal optimal solution with zero deviation is obtained as a result of the proposed algorithm. The initial values are $\lambda^0 = [1, 1, 1, 1]$, $x^0 = [0, 0, \dots, 0]$, $m = 100$.

B. Nonconvex Geometric Programming Problem

In the second example, the proposed Lagrangian coordination based on the convexity assumption of the objective function and constraints is expanded to a nonconvex case. Although being nonconvex, the example has a unique solution. The original geometric programming problem is stated as follows:

$P_{original}$:

$$\min_{x_1, x_2, \dots, x_{18}} x_1^2 + x_2^2 \quad (23)$$

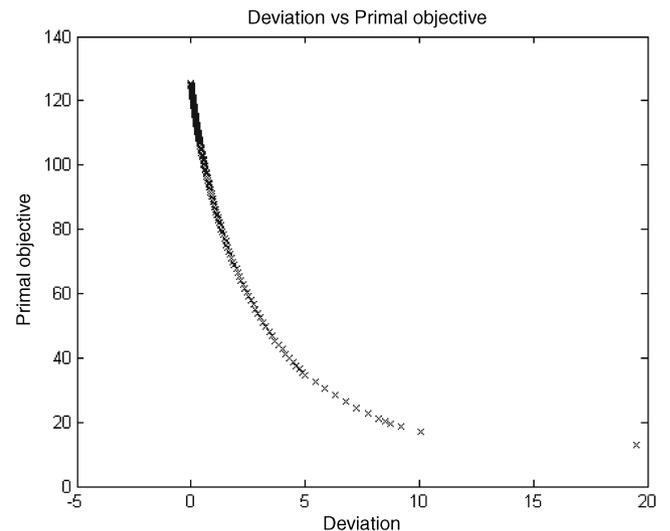


Fig. 2 Deviation vs primal objective Pareto curve for the AIO–LP in Eq. (15) of the example problem.

Table 5 Optimal design with the weight update scheme and with the dual coordination

Design variables	Optimal design	Weight update scheme [14]		Lagrangian dual coordination	
		Final design	Error (%)	Final design	Error (%)
x_1	2.84	2.75	-2.93	2.84	0.12
x_2	3.09	2.58	-16.66	3.07	-0.49
x_3	2.36	2.25	-4.43	2.36	0.19
x_4	0.76	0.76	0.01	0.76	0.04
x_5	0.87	0.88	1.11	0.87	-0.02
x_6	2.81	2.21	-21.46	2.80	-0.60
x_7	0.94	0.99	5.28	0.94	0.07
x_8	0.97	0.97	-0.55	0.97	0.05
x_9	0.87	0.90	4.02	0.86	-0.46
x_{10}	0.80	0.82	3.05	0.79	-0.39
x_{11}	1.30	1.14	-12.73	1.29	-0.68
x_{12}	0.84	0.84	-0.08	0.84	0.00
x_{13}	1.76	1.52	-13.59	1.75	-0.52
x_{14}	2.84	2.75	-2.93	2.84	0.12

subject to

$$\begin{aligned}
 g_1 &= \frac{(x_3^{-2} + x_4^2)}{x_5^2} \leq 1, & g_2 &= \frac{(x_5^2 + x_6^{-2})}{x_7^2} \leq 1, \\
 g_3 &= \frac{(x_8^2 + x_9^2)}{x_{11}^2} \leq 1, & g_4 &= \frac{(x_8^{-2} + x_{10}^2)}{x_{11}^2} \leq 1, \\
 g_5 &= \frac{(x_{11}^2 + x_{12}^{-2})}{x_{13}^2} \leq 1, & g_6 &= \frac{(x_{11}^2 + x_{12}^2)}{x_{14}^2} \leq 1, \\
 h_1 &= \frac{(x_3^2 + x_4^{-2} + x_5^2)}{x_1^2} = 1, & h_2 &= \frac{(x_5^2 + x_6^2 + x_7^2)}{x_2^2} = 1, \\
 h_3 &= \frac{(x_8^2 + x_9^{-2} x_{10}^{-2} + x_{11}^2)}{x_3^2} = 1, \\
 h_4 &= \frac{x_{11}^2 + x_{12}^2 + x_{13}^2 + x_{14}^2}{x_6} = 1
 \end{aligned}$$

Notice that polynomial functions in the objective and constraints are represented in the posynomial form, which indicates that the original problem has a unique global minimum [20]. A global optimal solution is found at

$$\mathbf{x} = (2.84, 3.09, 2.36, 0.76, 0.87, 2.81, 0.94, 0.97, 0.87, 0.80, 1.30, 0.84, 1.76, 1.55)$$

Following the two-level decomposition [1], the same as the one depicted in Fig. 1, the original problem is decomposed after introducing the following additional equality constraints:

$$\begin{aligned}
 h_5 &= x_{15} - x_3 = 0, & h_6 &= x_{17} - x_6 = 0, \\
 h_7 &= -x_{11} + x_{16} = 0, & h_8 &= x_{18} - x_{11} = 0
 \end{aligned} \tag{24}$$

The optimal Lagrangian multipliers associated with these equality constraints are 4.25, 5.53, 7.68, and 7.68.

According to the AIO-LP in Eq. (15), the deviation constraints h_5, \dots, h_8 can be relaxed and moved to the objective function as follows:

$P_{\text{AIO-LP}}$:

$$\begin{aligned}
 \min_{x_1, x_2, \dots, x_{18}} & x_1^2 + x_2^2 + \lambda_1 h_5 + \lambda_2 h_6 + \lambda_3 h_7 + \lambda_4 h_8 \\
 & + |\lambda_1|(h_5)^2 + |\lambda_2|(h_6)^2 + |\lambda_3|(h_7)^2 + |\lambda_4|(h_8)^2
 \end{aligned} \tag{25}$$

subject to

$$\begin{aligned}
 g_1 &= \frac{(x_3^{-2} + x_4^2)}{x_5^2} \leq 1, & g_2 &= \frac{(x_5^2 + x_6^{-2})}{x_7^2} \leq 1, \\
 g_3 &= \frac{(x_8^2 + x_9^2)}{x_{11}^2} \leq 1, & g_4 &= \frac{(x_8^{-2} + x_{10}^2)}{x_{11}^2} \leq 1, \\
 g_5 &= \frac{(x_{11}^2 + x_{12}^{-2})}{x_{13}^2} \leq 1, & g_6 &= \frac{(x_{11}^2 + x_{12}^2)}{x_{14}^2} \leq 1, \\
 h_1 &= \frac{(x_3^2 + x_4^{-2} + x_5^2)}{x_1^2} = 1, & h_2 &= \frac{(x_5^2 + x_6^2 + x_7^2)}{x_2^2} = 1, \\
 h_3 &= \frac{(x_8^2 + x_9^{-2} x_{10}^{-2} + x_{11}^2)}{x_3^2} = 1, \\
 h_4 &= \frac{x_{11}^2 + x_{12}^2 + x_{13}^2 + x_{14}^2}{x_6} = 1
 \end{aligned}$$

The coordination process starts at the system level by solving the Lagrangian primal problem, followed by solving two subsystem Lagrangian primal problems. First, the SYS-LP is solved for the initial values of the Lagrange multipliers. As shown in Fig. 1, after solving SYS-LP in Eq. (23), the values for $x_3, x_6,$ and x_{11} are cascaded to SUB-LP₁ and SUB-LP₂. The two subsystem problems are solved to achieve the minimum deviations from the target values from the system-level problem and the values for $x_{15}, x_{16}, x_{17},$ and x_{18} , as well as for the local variables, are obtained. After solving the subproblems at the system and subsystem levels, the Lagrange multipliers are updated based on the subgradient obtained following Eq. (18).

Performance of the dual coordination was compared with two other approaches in the literature to validate the proposed method with respect to two criteria: accuracy of the solution and efficiency of

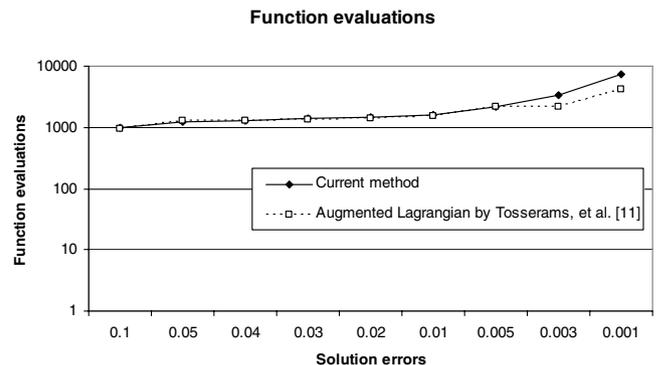


Fig. 3 Comparison of function evaluations.

the coordination process. For the accuracy comparison, the method was compared with the weight update scheme of Michalek and Papalambros [14]. Table 5 reports the optimal solution of this example obtained with the scheme of Michalek and Papalambros and the dual coordination, for which the tolerance for termination of the coordination process was set to 0.01. Notice that the proposed coordination outperforms the weight update scheme and that the higher errors from the scheme are due to the discrepancy between the weights (27.72, 14.61, 16.64, and 16.64) and the Lagrangian multipliers (4.25, 5.53, 7.68, and 7.68) for the deviation constraints h_5, \dots, h_8 . These results confirm that the current coordination algorithm finds the optimal solution as well as the optimal weights, which correspond to the Lagrange multipliers, with a higher accuracy than the weight update scheme of Michalek and Papalambros.

For the efficiency comparison, the current coordination process was compared with that recently proposed by Tosserams et al. [11] and based on the augmented Lagrangian relaxation. They showed that their relaxation outperforms the weight update scheme [14] with respect to the computational effort (i.e., number of function evaluations) by order of 10 to 100. The current coordination algorithm presented in this paper is compared with that in Tosserams et al. to validate the efficiency. Figure 3 shows that the numbers of function evaluations are similar in the order of magnitude, when using the current method as well as the augmented Lagrangian relaxation by Tosserams et al., However, more numerical studies are needed to compare the two approaches in more detail. Furthermore, for an extended hierarchy of more than two levels, Tosserams et al. proposed adopting the alternating direction method [11] to reduce the inner-loop computational effort. This alternating direction method can also be incorporated with the current coordination to maintain computational efficiency in multilevel cases.

VI. Conclusion

In this paper, an ATC formulation and coordination earlier proposed in the literature are conceptually and computationally enhanced. The enhanced ATC formulation also employs Lagrange multipliers associated with the deviation constraints as the weights for the deviation terms whereas the Lagrangian duality theory guarantees that the proposed algorithm successfully converges to an optimal solution of the original convex/affine design problem. Similarly, the coordination algorithm finds the optimal weights for the deviation terms, which, contrary to the conventionally large weight values for the deviation terms, need not necessarily be large. The algorithm uses subgradient optimization to solve the Lagrangian dual problem and also the ATC-decomposed primal problem whose solution is proven to be identical with the all-in-one solution. Thus the proposed algorithm encompasses two desirable features: it is specialized for ATC-decomposed problems and maintains the convergence characteristics of the Lagrangian dual problem.

The conceptual enhancement involves a different decomposition scheme of the primal and dual problems, which results in target cascading between levels and, in turn, allows for the use of augmented Lagrangian. The latter improves computational efficiency of the coordination algorithm as well as makes the method suitable for nonconvex problems. A guideline to impose a norm for the ATC-introduced deviation terms is proposed, using the augmented Lagrangian dual formulation. The use of norm prevents the algorithm from numerical instability around the optimal solution due to large weights for the deviation terms.

Subgradient optimization is used to update the Lagrange multipliers (i.e., weights for the deviation terms) as the solution process progresses. There can be variations in the convergence behavior depending on the step size update formula, the initial starting values for the design variables, and the Lagrange multipliers. However, two demonstration cases show that the proposed algorithm consistently converges to the optimal solution of the original problem and either outperforms or performs at least as well as other methods. Although the algorithm assumes convexity/affinity of the responses and constraint functions, it can be extended to solving realistic engineering problems in nonconvex feasible spaces as

shown in the geometric programming example problem. Furthermore, in one of the demonstration examples, the analogy is drawn between the coordination algorithm and generation of Pareto optimal solutions of the biobjective problem implied by the ATC formulation.

Future work may involve adopting different dual optimization schemes such as cutting plane methods [12] for improving efficiency and robustness of the algorithm.

Appendix: Local Duality and Augmented Lagrangian

I. Local Duality Theorem

Suppose that the problem
Minimize

$$f(\mathbf{x}) \quad (\text{A1})$$

subject to

$$\mathbf{h}(\mathbf{x}) = \mathbf{0}$$

has a local solution at \mathbf{x}^* with corresponding value r^* and Lagrange multiplier $\boldsymbol{\lambda}^*$. Suppose also that \mathbf{x}^* is a regular point of the constraints and that the corresponding Hessian of the Lagrangian $\mathbf{L}^* = \mathbf{L}(\mathbf{x}^*)$ is positive definite and the dual function ϕ is defined by the equation

$$\phi(\boldsymbol{\lambda}) = \min_{\mathbf{x}} [f(\mathbf{x}) + \boldsymbol{\lambda}^T \mathbf{h}(\mathbf{x})] \quad (\text{A2})$$

Then the dual problem
Maximize

$$\phi(\boldsymbol{\lambda}) \quad (\text{A3})$$

has a local solution at $\boldsymbol{\lambda}^*$ with corresponding value r^* and \mathbf{x}^* as the point corresponding to $\boldsymbol{\lambda}^*$ in the definition of ϕ .

II. Augmented Lagrangian Function

The augmented dual Lagrangian function for the equality constrained problem given by Eq. (8) is defined as the function

$$l_c(\mathbf{x}, \boldsymbol{\lambda}) = f(\mathbf{x}) + \boldsymbol{\lambda}^T \mathbf{h}(\mathbf{x}) + \frac{1}{2}c|\mathbf{h}(\mathbf{x})|^2 \quad (\text{A4})$$

for some positive constant c . From the view of duality theory, the augmented Lagrangian is simply the standard Lagrangian for the problem
minimize

$$f(\mathbf{x}) + \frac{1}{2}c|\mathbf{h}(\mathbf{x})|^2 \quad (\text{A5})$$

subject to $h(\mathbf{x}) = \mathbf{0}$.

This problem is equivalent to the original problem, because the addition of the term $\frac{1}{2}c|\mathbf{h}(\mathbf{x})|^2$ to the objective does not change the optimal solution point or the optimal Lagrange multiplier.

Acknowledgements

This work has been supported at Northwestern University by the NSF grant no. DMII 0335880, the Ford University Research Program, and the exploratory NSF research grant no. DMI-0503737 in collaboration with the University of Michigan. Margaret M. Wiecek's research has been supported in part by the Automotive Research Center, a U.S. Army TACOM Center of Excellence for Modeling and Simulation of Ground Vehicles at the University of Michigan, and by the National Science Foundation, grant no. DMS-0425768. The authors appreciate useful discussions with L. F. P. Etman and S. Tosserams at the Technische Universiteit Eindhoven and helpful comments from anonymous reviewers that improved the quality of the manuscript.

References

- [1] Kim, H. M., Michelena, N. F., and Papalambros, P. Y., and Jiang, T., "Target Cascading in Optimal System Design," *Journal of mechanical design*, Vol. 125, No. 3, 2003, pp. 474–480.
- [2] Kim, H. M., Rideout, D. G., Papalambros, P. Y., and Stein, J. L., "Analytical Target Cascading in Automotive Vehicle Design," *Journal of Mechanical Design*, Vol. 125, No. 3, 2003, pp. 481–489.
- [3] Kim, H. M., "Target Cascading in Optimal System Design," Ph.D. Thesis, Univ. of Michigan, Ann Arbor, MI, 2001.
- [4] Kokkolaras, M., Fellini, R., Kim, H. M., Michelena, N. F., and Papalambros, P. Y., "Extension of the Target Cascading Formulation to the Design of Product Families," *Structural and Multidisciplinary Optimization*, Vol. 24, No. 4, 2002, pp. 293–301.
- [5] Michelena, N., Kim, H. M., and Papalambros, P. Y., "A System Partitioning and Optimization Approach to Target Cascading," *International Conference on Engineering Design*, Vol. 24, Technische Universität München, Munich, Germany, 1999, pp. 1109–1112.
- [6] Michelena, N. F., and Papalambros, P. Y., "A Hypergraph Framework for Optimal Model-Based Decomposition of Design Problems," *Computational Optimization and Applications*, Vol. 8, No. 2, 1997, pp. 173–196.
- [7] Michelena, N., Park, H., and Papalambros, P. Y., "Convergence Properties of Analytical Target Cascading," *AIAA Journal*, Vol. 41, No. 5, 2003, pp. 897–905.
- [8] Allison, J. T., Kokkolaras, M., and Papalambros, P. Y., "On the Use of Analytical Target Cascading and Collaborative Optimization for Complex System Design," *6th World Congress on Structural and Multidisciplinary Optimization*, International Society for Structural and Multidisciplinary Optimization, cosponsored by AIAA and the Brazilian Society of Mechanical Sciences and Engineering, Rio de Janeiro, 2005.
- [9] Braun, R. D., "Collaborative Optimization: An Architecture For Large-Scale Distributed Design," Ph.D. Thesis, Stanford Univ., Stanford, CA, 1996.
- [10] Michelena, N., Papalambros, P. Y., Park, H., and Kulkarni, D., "Hierarchical Overlapping Coordination for Large-Scale Optimization by Decomposition," *AIAA Journal*, Vol. 37, 1999, pp. 890–896.
- [11] Tosserams, S., Etman, L. F. P., Papalambros, P. Y., and Rooda, J. E., "An Augmented Lagrangian Relaxation for Analytical Target Cascading using the Alternating Directions Method of Multipliers," *Structural and Multidisciplinary Optimization*, Vol. 31, No. 3, Springer, Berlin, 2006, pp. 176–189.
- [12] Bertsekas, D. P., *Nonlinear Programming*, 2nd ed., Athena Scientific, Belmont, MA, 1999.
- [13] Papalambros, P. Y., and Wilde, D., *Principles of Optimal Design: Modeling and Computation*, 2nd ed., Cambridge Univ. Press, New York, 2000.
- [14] Michalek, J. J., and Papalambros, P. Y., "An Efficient Weighting Update Method to Achieve Acceptable Consistency Deviation in Analytical Target Cascading," *Journal of Mechanical Design*, Vol. 127, No. 2, 2005, pp. 206–214.
- [15] Lassiter, J., Wiecek, M., and Andrighetti, K., "Lagrangian Coordination and Analytical Target Cascading: Solving ATC-Decomposed Problems with Lagrangian Duality," *Optimization and Engineering*, Vol. 6, No. 3, 2005, pp. 361–381.
- [16] Everett, H., "Generalized Lagrange Multiplier Method for Solving Problems of Optimum Allocation of Resources," *Operations Research*, Vol. 11, No. 3, 1963, pp. 399–417.
- [17] Lasdon, L. S., "Duality and decomposition in mathematical programming," *IEEE Transactions on Systems Science and Cybernetics*, Vol. 4, No. 2, 1968, pp. 86–100.
- [18] Luenberger, D., *Linear and Nonlinear Programming*, 2nd Ed., Addison Wesley Longman, Reading, MA, 1984.
- [19] Hiriart-Urruty, J.-B., and Lemarechal, C., *Convex Analysis and Minimization Algorithms*, Springer-Verlag, Berlin, 1993.
- [20] Bightler, C. S., and Phillips, D. T., *Applied Geometric Programming*, Wiley, New York, 1976.

A. Messac
Associate Editor