Proceedings of the DETC'00
ASME 2000 Design Engineering Technical Conferences
and Computers and Information Engineering Conference
Baltimore, Maryland, September 10-13, 2000

DETC 2000/DAC-14253

An Investigation into Modeling and Solution Strategies for Optimal Design and Control

Julie A. Reyer
Graduate Student
reyer@umich.edu

Panos Y. Papalambros
Professor
pyp@umich.edu

Optimal Design Laboratory
Department of Mechanical Engineering
University of Michigan
Ann Arbor, Michigan, 48109, U.S.A.
734.936.2624

Keywords: Optimal design, optimal control, monotonicity analysis, decomposition strategies, partitioning

ABSTRACT

Combined optimal embodiment design and control is increasingly necessary in designing modern artifacts. Several strategies for achieving such a system-level optimum were examined previously in the context of a case study. Since different strategies often lead to different results, a more comprehensive understanding of the relative advantages and disadvantages of these "concurrent design" processes is desirable. The present article explores this question using a very simple problem to demonstrate the complexities involved. Monotonicity analysis and a decomposition framework help explain the reasons for arriving at different results, even unexpectedly. Conclusions are drawn regarding the effectiveness of the various modeling formulations.

INTRODUCTION

Traditional embodiment design and control of an artifact or system has been performed in a sequential, if not altogether separate, manner. One creates a design and then designs a controller for that specific design. Optimization was naturally applied in the same sequence: optimize the design then optimize the control. The next logical improvement was to iteratively optimize the design then optimize the control, or combine them in some appropriate manner. Combined design and control has been studied for several systems, including structures and robotics [e.g., Messac and Turner 1984; Khot and Abhyankar 1993; Park and Asada 1994].

In structural design the first integration attempts were to optimize the traditional sequential analysis, namely, optimize around the entire sequence instead of around each part of the sequence. Another approach involved a bilevel optimization process: At each structural design point, a corresponding optimal controller was found using the analytical solution to a linear quadratic regulator (LQR) controller [Hale et al. 1985]. Of course, the most general approach is a true concurrent strategy, sometimes referred to as 'all at once,' where all objectives and constraints for both design and control are placed in a single optimization model and solved as one problem. For many engineering problems, the complexity and size of the combined problem could make this concurrent strategy impractical, unless special decomposition techniques could be used.

In an earlier paper [Reyer and Papalambros 1999], various strategies for optimizing combined design and control systems were studied in the context of a case study involving a direct current electric motor. These strategies reach different results and do not necessarily lead to the best design and control solution. The reasons for such results are not readily apparent. One might suspect that unexpected results come from failure of the numerical optimization algorithms or from the presence of multiple optima. A simple mathematical model could then be concocted to allow exploration of the conceptual issues involved without the added burden of possibly unreliable numerical processes.

In the present paper we introduce a very simple demonstration example, and then look at the various design and control strategies as applied to this simple problem. Using monotonicity analysis [Papalambros and Wilde 1988] we trace the reasons for reaching different solutions to the way coupling constraints are handled by the different strategies. Looking at the design and control problem in a formal hierarchical decomposition framework [as, e.g., in Wagner 1993] illuminates the modeling discrepancies among the strategies. The following sections describe the above points in detail.

DEMONSTRATION MODEL

A simple model is developed for use in the exploration of combined optimal embodiment design and control problems. The model is the simplest mathematical construct that contains all of the quantities of interest in the combined problem. A more complete rationale for the formulation and terminology used can be found in the cited reference [Reyer and Papalambros 1999]. The term “simple” below means that the quantity in question is local to the problem at hand (design or control), while the term “coupling” means that the relevant quantity appears in the other problem as well.

The Design Problem

The design problem for the demonstration model has several requirements: The design problem needs at least one design variable and one input coupling parameter; it must have an optimal solution and that solution must permit the existence of an output coupling parameter.

$$
\begin{aligned}
&\text{minimize } f = a_1 d \\
&\text{variable: } d \\
&\text{coupling parameter: } b \\
&\text{simple parameters: } \boldsymbol{a} = \{a_1, a_2, a_3\} \geq \mathbf{0} \\
&\text{subject to:} \\
&\qquad a_2 b + a_3 - d \leq 0 \\
&\text{output: } v = d^* \\
&\text{optimum: } d^* = a_2 b + a_3
\end{aligned}
\tag{1}
$$

The model in Equation (1) has one design variable d, one input coupling parameter b, several simple parameters $\boldsymbol{a}$, and one output coupling parameter v. The problem is linear with a single unique global optimal solution which is constraint bound (i.e., fully determined by active constraints).

The Control Problem

The control problem is an optimal gains one with finite-dimensional variables, thus allowing its treatment by nonlinear programming techniques rather than variational calculus. Essentially we assume that the controller configuration is decided.

The demonstration example must have at least one control variable and one input coupling parameter. The optimal solution must exist and permit the existence of an output coupling parameter. The statement of this control problem is given in Equation (2). The model has one control variable p, one input coupling parameter v, several simple parameters $\boldsymbol{u}$, and an output coupling parameter b. Due to the problem’s simplicity the control strategy is equivalent to both a proportional feedback and an LQR controller. The control optimum is based on the LQR solution and has an analytical form.

$$
\begin{aligned}
&\text{minimize } J = \int_0^{t_f} (x^T u_1 x + z^T u_2 z)\,\mathrm{d}t \\
&\qquad = \frac{u_5^{\,2}(u_1 + u_2 p^2)}{2(v - u_3 p)}[e^{2(v - u_3 p)t_f} - 1] \\
&\text{variable: } k \\
&\text{coupling parameter: } v \\
&\text{simple parameters: } \boldsymbol{u} = \{u_1, u_2, u_3, u_4, u_5, t_f\} \geq \mathbf{0} \\
&\text{subject to:} \\
&\qquad \dot{x}(t) = vx + u_3 z = (v - u_3 p)x \Rightarrow x(t) = e^{(v - u_3 p)t} \\
&\qquad y = u_4 x \\
&\qquad z = -px \\
&\qquad x(t = 0) = u_5 \\
&\text{output: } b = x(t = 1) = e^{(v - u_3 p^*)} \\
&\text{optimum: } p^* = (u_1 v + \sqrt{(u_1^2 v^2 + u_1 u_2 u_3^2)})/u_3^2
\end{aligned}
\tag{2}
$$

The Combined Problem

The combined embodiment design and control problem integrates the preceding design and control models, Equation 3.

$$
\begin{aligned}
&\text{minimize } \phi = w_1 f + w_2 J \\
&\text{where: } f = a_1 d \\
&\quad \text{and } J = \int_0^{t_f} (x^T u_1 x + z^T u_2 z)\,\mathrm{d}t \\
&\qquad = \frac{u_5^{\,2}(u_1 + u_2 p^2)}{2(v - u_3 p)}[e^{2(v - u_3 p)t_f} - 1] \\
&\text{variables: } d, k \\
&\text{coupling quantities: } b, v \\
&\text{simple parameters: } \boldsymbol{a} = \{a_1, a_2, a_3\} \geq \mathbf{0} \\
&\qquad \boldsymbol{u} = \{u_1, u_2, u_3, u_4, u_5, t_f\} \geq \mathbf{0} \\
&\text{subject to:} \\
&\qquad \dot{x}(t) = vx + u_3 z = (v - u_3 p)x \Rightarrow x(t) = e^{(v - u_3 p)t} \\
&\qquad y = u_4 x \\
&\qquad z = -px \\
&\qquad x(t = 0) = u_5 \\
&\quad l: -p \leq 0 \\
&\quad g: a_2 b + a_3 - d \leq 0 \\
&\quad \beta: e^{(v - u_3 p^*)} - b \leq 0 \\
&\quad \gamma: d - v \leq 0
\end{aligned}
\tag{3}
$$

The combined objective is the weighted sum of the design and control objectives, representing the Pareto solution of the two problems. The combined model has two variables *d* and *p*, two coupling quantities *b* and *v*, two sets of simple parameters ***a*** and ***u***. The model includes four inequality constraints: one from the control side *l*; one from the design side *g*, and two for coupling β and γ.

These coupling constraints are *directed* as explained in the next section. "Direction" means that they can be viewed as equality constraints that are equivalent to *in*equality constraints with the proper sense of the inequality sign (direction), which are active at the optimum [Papalambros and Wilde 1988]. This is a critical issue in understanding how the different strategies work and will be explored in more detail further below.

SOLVING THE COMBINED PROBLEM

We briefly review the different solution strategies that we will use to solve the combined problem. Then we discuss the issue of directing the coupling constraints, which will turn out to be an important element in our subsequent interpretation of the numerical results.

Solution Strategies

The combined optimization strategies are typical formulations for the solution of systems involving both an embodiment design problem and a control problem. These strategies are explained more fully in Reyer and Papalambros [1999]. The traditional method of optimizing the design then optimizing the control is called the *Single Pass Strategy.* Since the single pass formulation requires that one set of coupling quantities be fixed, an improvement is to modify those parameters via the *Iterative Strategy* — repetitively designing and controlling until the coupling quantities agree. The *Fixed Input Parameter Strategy* optimizes the entire design and control system while fixing one set of coupling parameters. The *Concurrent Strategy* or "All At Once" optimizes the system treating all coupling quantities as variables. The *Bilevel Strategy* treats the optimal control as a subproblem and optimizes the system assuming that the optimal system will have optimal controller gains.

Directing Coupling Constraints

Coupling constraints perform a critical role in the combined problem. Here we will examine their origins, their directions, and their effects. The coupling constraints stem from the acceptability conditions that are applied in the traditional, sequential, single pass, design-first-then-control strategies. In the single pass strategy the corresponding input and output coupling parameters must coincide in an acceptable manner. For example, in a direct current motor, the designed power needs to be larger than the amount of power to which the controlled motor is subjected. The direction of the equation for acceptability needs to be examined for each coupling parameter, but is generalized to Equation (4).

Acceptability Condition: $$\boldsymbol{b}_{out} - \boldsymbol{b}_{in} \le \mathbf{0} \qquad (4)$$

	d	b_{in}	b_{out}
minimize $f = a_1 d$	+		
subject to: $a_2 b_{in} + a_3 - d \le 0$	-	+	
control output: $b_{out} - e^{(v - u_3 k^*)} = 0$			+/-
coupling constraint: $b_{out} - b_{in}$ (≤ or ≥ ?) 0		-/+	+/-

Figure 1. Monotonicities of Design Problem

Since our demonstration example is a mathematical one, there is no physical interpretation to help direct the coupling constraints. However, a rigorous method can be employed using the principles of monotonicity analysis [Papalambros and Wilde 1988].

Consider the design problem and its monotonicities, Figure 1. For the design problem, the only variable, *d*, has a positive monotonicity in the objective function. According to the First Monotonicity Principle the design constraint is required to be active and the problem is solved. For the combined strategies, however, the coupling *parameters*, ignored in the separate sequential analysis, become coupling *variables*. As variables the coupling quantities must also be examined in monotonicity analysis. The design constraint, required to be active to bound *d*, bounds the coupling variable, b_{in}, from above. The Second Monotonicity Principle indicates that, if an optimum exists, there must be another active constraint that bounds b_{in} from below. Hence, the coupling constraint for *b* in this problem must be as in Equation (5).

Coupling Constraint: $$b_{out} - b_{in} \le 0 \qquad (5)$$

The control output equality may either become a directed equality $b_{out} - e^{(v - u_3 p^*)} \equiv > 0$ or be removed from the problem via substitution so that $e^{(v - u_3 p^*)} - b_{in} \le 0$.

In general, monotonicity analysis can be used to determine the direction of the coupling constraints. Once properly directed, the coupling constraints become an important part of the combined optimal embodiment design and control solution. The coupling constraints guarantee feasibility of the final optimal solution, even when the corresponding input and output coupling parameters do not exactly match. Activity of inactivity of the coupling constraints will be used to interpret the numerical solutions presented in the next section.

SOLUTION OF DEMONSTRATION MODEL

We will first present numerical results for each design strategy and then show how these results can be confirmed and

Table 1: Demonstration Model: Numerical Results

Quantity	Sequential					Combined					
	Single Pass				Itera-tive	Fixed Input Parameters				Con-current	Bilevel
ϕ	2.023	2.015	1.993	2.120	1.993	1.931	1.915	1.993	2.120	1.915	1.993
w_1, w_2	N/A	N/A	N/A	N/A	N/A	0.500	0.500	0.500	0.500	0.500	0.500
f	1.230	1.225	1.208	1.300	1.208	1.075	1.109	1.208	1.300	1.109	1.208
J	2.816	2.806	2.777	2.940	2.777	2.786	2.720	2.777	2.940	2.720	2.777
g	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
l	-2.816	-2.806	-2.777	-2.940	-2.777	-3.665	-3.325	-2.777	-2.940	-3.325	-2.777
β	-0.025	-0.019	0.000	-0.106	0.000	0.000	0.000	0.000	-0.106	0.000	0.000
γ	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
b_{in}	0.075	0.109	0.208	0.300	0.208	0.075	0.109	0.208	0.300	0.109	0.208
$b_{in}^{\dagger}$	0.230	0.225	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
b_{out}	0.205	0.206	0.208	0.194	0.208	0.075	0.109	0.208	0.194	0.109	0.208
d	1.230	1.225	1.208	1.300	1.208	1.075	1.109	1.208	1.300	1.109	1.208
p	2.816	2.806	2.777	2.940	2.777	3.665	3.3246	2.777	2.940	3.325	2.777
v	1.230	1.225	1.208	1.300	1.208	1.075	1.109	1.208	1.300	1.109	1.208

interpreted analytically using monotonicity analysis.

Numerical Solutions

Complete numerical results from optimization under each strategy are shown in Table 1. The combined objective value, ϕ, is stated for all strategies, though the sequential techniques consider only separate design and control objectives. The value of the coupling parameter input to the design analysis is b_{in}. The quantity $b_{in}^{\dagger}$ in the single pass strategy stems from an iteration after the initial value does not pass the acceptability condition. The value of b output from the control analysis is b_{out}. The remaining variables and coupling quantity are d, p and v.

The concurrent strategy solution is the best, with the lowest combined objective value. The bilevel and iterative strategy solutions match and are second best. The single pass is the worst. The fixed input parameter solutions can match any of the other solutions, depending on the value of b_{in}. Activity of constraints changes depending on the solution strategy. Both the single pass and fixed input strategies find the coupling constraint β inactive while the other methods maintain its activity.

Three strategies, iterative, concurrent, and bilevel, indicate activity of the design and coupling constraints, g, β, and γ. Though these strategies have the same active constraints, their final optimal solutions do not match; the concurrent method provides a better (lower) minimum than the other strategies. From a system analysis viewpoint, the fact that these methods have the same active constraints, satisfy the optimality conditions, but give different answers, creates severe difficulties. No information is available from the iterative and bilevel optimizations to indicate that their solutions could be improved upon. Moreover, the concurrent strategy results in the lowest control objective, but since the concurrent formulation does not require optimal gains and the bilevel method does, one would assume that the bilevel method should have a better optimal value for the control objective. Since that is not the case, one might suppose that the a priori restriction that the controller be optimal must somehow prevent the bilevel method from finding the system optimum.

Though the numerical solutions permit these observations, a more detailed analysis of the model formulations will shed more light on the causes of such discrepancies.

Analytical Solutions

The analytical solutions presented here confirm the numerical results: the strategies in fact do *not* yield the same optimal solutions. Several hypotheses are formulated to explain the discrepancies found when solving the problem with the five different optimization strategies.

Constraint activity is identified a priori using the monotonicity principles mentioned in the previous section. In addition to these principles, some constraints can be declared active

Table 2: Demonstration Model: Analytical Solutions

Strategy	Constraints				Analytical Objectives		Systems, with active constraints removed, reduce to the following:
	g	β	γ	J	f^*	J^*	
Single Pass	a 1	i 3	a 3	4	$b+1$	$J = \frac{1+(b+1+\sqrt{b^2+2b+2})}{-2\sqrt{b^2+2b+2}}(e^{-20\sqrt{b^2+2b+2}}-1)$	fully solved
Iterative	a 1	a 3	a 3	4	$b+1$	$J = \frac{1+(b+1-\ln(b))}{2\ln(b)}(b^{-20}-1)$	$\ln(b) = -\sqrt{b^2+2b+2}$
Fixed, b≤0.208	a 1	a 1	a 1		$b+1$	$J = \frac{1+(b+1-\ln(b))}{2\ln(b)}(b^{-20}-1)$	fully solved
Fixed, b>0.208	a 1	i	a 1	4	$b+1$	$J = \frac{1+(b+1+\sqrt{b^2+2b+2})}{-2\sqrt{b^2+2b+2}}(e^{-20\sqrt{b^2+2b+2}}-1)$	fully solved
Bilevel	a 1	a 2	a 1	4	$b+1$	$J = \frac{1+(b+1-\ln(b))}{2\ln(b)}(b^{-20}-1)$	$p = b+1-\ln(b)$ $= b+1+\sqrt{b^2+2b+2}$ $\Rightarrow \ln(b) = -\sqrt{b^2+2b+2}$
Concurrent	a 1	a 2	a 1		interior solution		$\underset{b}{\text{minimize}} \quad \phi = w_1 f + w_2 J$ $f = b+1$ $J = \frac{(1+(b+1-\ln(b))^2)(b^{20}-1)}{2\ln(b)}$ Subject to: l: $\ln(b) - b - 1 \leq 0$

Notes: a: active constraint, i: inactive constraint
1: First Monotonicity Principle, 2: Second Monotonicity Principle, 3: activity from strategy definition, 4: LQR Solution

due to the nature of the optimization strategies. From optimal control theory an analytical method for designing LQRs is available [Lewis and Symos 1995]. Using monotonicity principles, LQR solutions and strategy definitions, the analytical solutions for the example are derived and given in Table 2.

The single pass strategy had the worst numerical solution, and that solution depended upon an initial coupling parameter value. The analysis seeks to confirm those conclusions. Since the single pass strategy follows the sequential design-first-then- control methodology, constraint activity and optima are valid for the design and control separately, but not necessarily for the combined system. The design problem is linear with the objective increasing in d, and so constraint g must be active to provide a bound on d, linking the value of d to the input coupling parameter b. The remaining coupling constraint β is not typically active but it must be checked for feasibility. The control problem is a standard LQR formulation that has an analytical interior solution. The resulting optimum is a function of the input coupling parameter b. The strategy will fail when the value of the input parameter is not the correct one for the system as a whole.

Analysis for the iterative strategy is similar to that for the single pass one. Now the coupling constraints g, β and γ are always active. Simultaneous solution of the resulting set of equations yields the strategy optimum. The extra active constraint removes the dependence of the solution on an input parameter. The iterative solution strategy will fail if no simultaneous solution of the active constraints exists. Note that the activity for β and γ is dictated by the definition of the strategy rather than by the problem. For general combined design and control problems such simultaneous solutions may not be feasible and the strategy will fail.

The fixed input parameter strategy improves on the single pass method. Instead of sequentially optimizing the design first then the control, optimization is applied to the entire design and control system. The analytical solution proves that, like the single pass strategy, the optimum depends on the value of the input coupling parameter. Examination of the model reveals that variables d and v are always increasing in the objective. Constraints g and γ, decreasing in d and v respectively, must be active to bound d and v, respectively. The entire problem then reduces to Equation (6). In the reduced problem the objective ϕ'

$$\underset{p}{\text{minimize}} \quad \phi' = w_1(b+1) + w_2\left(\frac{(1+p^2)}{2(b+1-p)}\left[e^{20(b+1-p)} - 1\right]\right)$$

subject to: (6)

$$x(t) = e^{(v-p)t}$$

$$l: -p \le 0$$

$$\beta: e^{(b+1-p)} - b \le 0$$

contains the summation of a function of p and a constant, $w_1(b+1)$. The constant term may be ignored in the optimization and the problem reduces to a typical LQR formulation with one constraint, β. The constraint is active when b is less than 0.208 (in which case the LQR and the β-constrained solutions are equivalent) and inactive otherwise. When the constraint is inactive, the standard LQR minimization determines variable p. The resulting optimum for the fixed input parameter strategy, regardless of whether β is active, depends on the input parameter b. Based on the analytical solutions, the fixed input parameter solutions fall into three categories: $b > 0.208$, $b = 0.208$ and $b < 0.208$, see Table 1 for a numerical comparison. When $b > 0.208$, the solution should match the single pass strategy solution. When $b = 0.208$, the fixed input parameter solution should match the iterative solution. When $b < 0.208$, the solution does not relate to the sequential solutions. In fact, when $b < 0.208$, the β coupling constraint indicates that the single pass and iterative solutions are infeasible. The analysis proves that the improvement of the optimal solution from the fixed input parameter strategy, as compared to the single pass and iterative strategies, depends on the value of the input parameter.

The benefits of the bilevel strategy center around the existence of analytical solutions for the optimal control subproblem. For certain optimal control strategies an analytical solution exists and the combined problem becomes readily solvable. In the example here the LQR method gives an equation for the interior solution to the optimal gain problem, p^*. In the remaining problem, variables d and v are both increasing in the objective. The First Monotonicity Principle demands that constraints g and γ must be active to bound variables d and v, respectively. Variable b is increasing in the active constraint g and therefore requires constraint β to be active. The three active constraint equations and the analytical solution from the optimal gains problem form a system of equations that must be solved simultaneously to find the optimal values for the four variables.

The analytical solution confirms the numerical results, where the bilevel solution matches the iterative one. The analytical solution does, however, provide further insight. In the iterative solution activity was required for all coupling constraints by the strategy's definition. In the bilevel solution constraint activity was determined using monotonicity analysis to bound specific variables with individual constraints. This distinction does not affect the solution of the demonstration example, but it could be important in more complicated problems. If a system has an inactive coupling constraint, the iterative solution would be prohibited from matching the bilevel solution.

Comparing the bilevel strategy with the fixed input parameter strategy shows that the only instance where the solutions would match is when $b = 0.208$, the point when the optimal gain problem is optimized. The bilevel solution will be worse than the fixed input parameter strategy when b is less than 0.208 and better when b is greater than 0.208. The bilevel analytical solution provides further indications that the strategy does not and cannot find the true system optimum.

The concurrent strategy achieved the best numerical optimum, as expected. The strategy expands on the fixed input parameter one by changing the input coupling parameter into a coupling variable. The concurrent formulation, like the fixed input parameter one, has a single combined objective, where the variables d and v are both increasing and k is not monotonic. Constraint γ bounds v from below. Constraint g provides a lower bound on d and has a nonobjective variable b, which is increasing. In order to bound b, constraint β must be active. Using these active constraints to eliminate variables d, p and v, a new one-dimensional optimization problem results, Table 2.

This reduced problem does not have an apparent analytical solution. The LQR methodology used in the previous strategies does not apply due to the variable term $w_1(b+1)$ in the objective. The l constraint cannot be identified as active by monotonicity, since the objective is non monotonic. Furthermore, the constraint does not have a solution when it is active ($l = 0$). Thus, the solution to the reduced problem is an interior one with the same value as that found earlier numerically. Since this method represents optimization of the entire system and the solution satisfies the optimality conditions, its results is taken as the system's true optimum.

Graphical Interpretation

Graphical representations often provide insights into small optimization problems. The problem has four variables. Practical graphical representation may be done in three dimensions with two variables. To that end, two variables, b and v, are eliminated from the problem using constraints g and γ, which are active in all solution strategies.

Figure 2 depicts the reduced problem's design objective, control objective, and the β coupling constraint. The design objective is constant for all values of the control variable p and increases proportionally with d. The control objective increases proportionally with the design variable and is concave in the control variable. The β constraint is feasible as both d and p increase. (Note that the axes in the figure are reversed for viewing purposes.)

In the combined problem the two objectives and constraint are incorporated into a single problem. Figure 3 shows the feasible region of the equally weighted combined objective. The minimum of this function is marked in the figure and matches the concurrent strategy minimum.

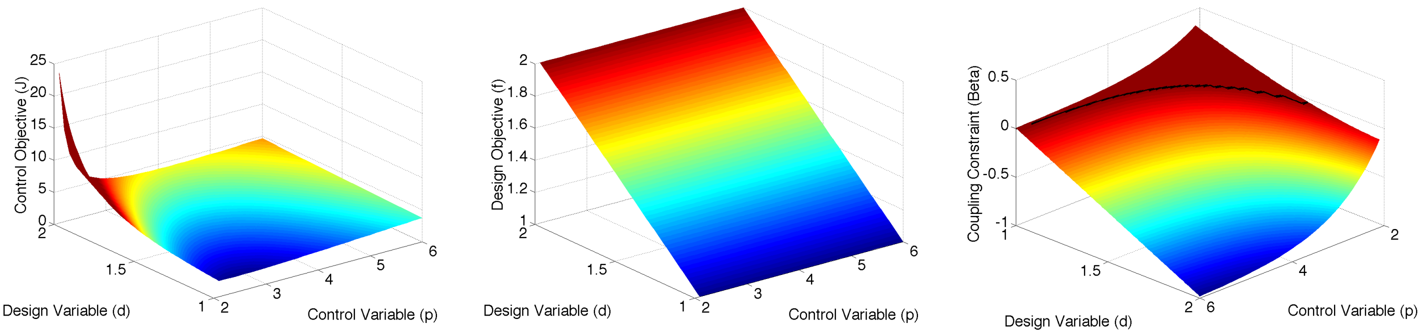


Figure 2. Control Objective, Design Objective and β Coupling Constraint

Figure 4 presents a two-dimensional graph of the results. The parallel horizontal lines represent isometric lines of the design objective. The convex lines are isometric contours of the control objective with values decreasing as d decreases. Two optimal functions are shown: the minimum of the control objective with respect to the variable p and parameter d and the minimum of the design objective with respect to variable d and parameter p. Constraint β is colinear with the design objective, with the feasible region above the line as shown in the figure. Finally, the locations of the solutions from the various strategies are marked.

Notably, all of the solutions except for that from the concurrent strategy fall on the minimum control objective line, while the concurrent solutions, even at the extreme points, do not occur on that line. The single pass solution can fall anywhere on the feasible part of the minimum control objective line, depending on the input coupling parameter, b_{in}. The iterative solution must always occur at the intersection of the minimum design objective and minimum control objective lines. The bilevel solution must always occur on the minimum control objective line. The fixed input parameter solution for this problem actually forms a locus of all possible 'optimal' solution values – for $b_{in} \geq 0.208$ the solution occurs on the feasible minimum control objective line, and for $b_{in} \leq 0.208$ the solution falls on the part of the minimum design objective line that lies to the right of the minimum control objective line. The concurrent solutions occur on the minimum design objective line from the point where $w_1 = 0$ and to the right.

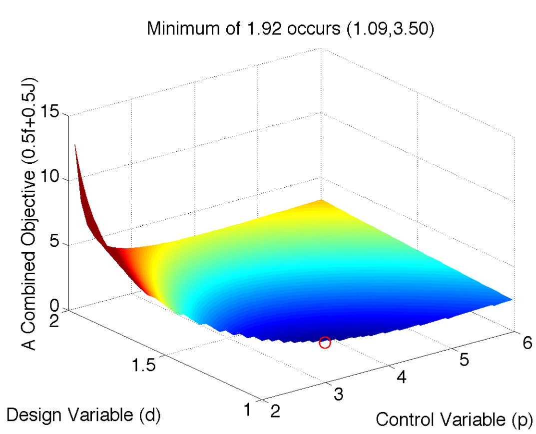


Figure 3. Feasible Combined Objective

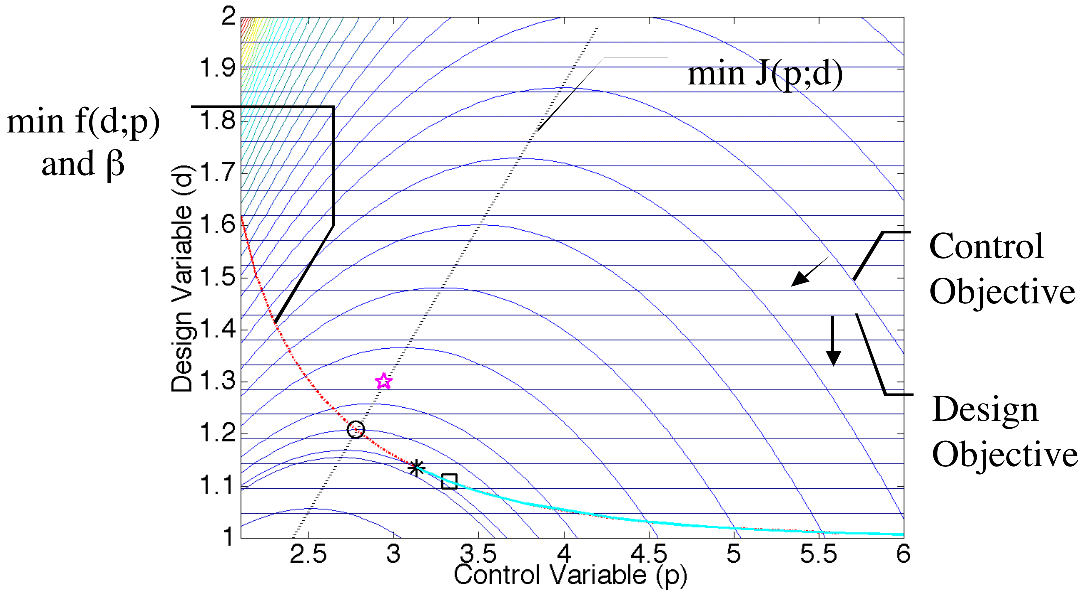
A DECOMPOSITION VIEWPOINT

In this section design decomposition – the development of a partition and coordination strategy – is applied to the concurrent strategy. As mentioned earlier, although the concurrent strategy is conceptually the best, it is also likely to lead to large intractable optimization problems. The design and control problems, solved together, with the coupling quantities, could combine to form a problem with a large number of variables and constraints. Decomposition presents an opportunity to divide and solve large problems rigorously, and it may provide a way to solve these more complex concurrent problems. Furthermore, the other four strategies can be viewed as special cases of the partitioned concurrent problem. A decomposition framework can provide further insight into when, if ever, these other strategies can find the true system optimum.

Partition Synthesis

A decomposition strategy first requires a partitioned problem. The development of a partition requires examination of the relationship between the functions and variables in a system model. Wagner [Wagner 1993]developed a useful way to visualize these relationships with a matrix called a functional dependence table (FDT). In the FDT, the rows represent the model functions and the columns signify the model variables. Functions either depend on a variable, denoted with ■, or do not. The functional dependence table for the demonstration problem appears in Figure 5(a).

Many partitions exist for this model. A good strategy is one that minimizes the coupling between the partitioned subproblems. Additionally, for combined embodiment design and control problems, the availability of specialized methods for



☆ : Single Pass and Fixed Input Parameter, b_{in}=0.3 ❋ : AAO, $w_1 = 0$

❍ : Iterative and OGS ❑: AAO, $w_1 = w_2$

Figure 4. Design and Control Variable Space

solving separate optimal design and optimal control problems induce a strong preference for keeping the design side of the problem separate from the control side. Krishnamachari [Krishnamachari, 1996] developed a methodology for choosing optimal hierarchical partitions while enforcing preferences on what functions should be kept together within a subproblem. Using this methodology leads to the partition in Figure 5(b).

The partitioned problem has two linking variables, the coupling variables *b* and *v*, and two subproblems, one for design and one for control, Figure 6. In the hierarchical representation the master problem has two variables, *b* and *v*, three constraints, *g*, β and γ, and the weighted combined objective. The design subproblem has one variable *d*, two constraints, *g* and γ, and the design objective *f*. Note that in the design subproblem, the control part of the system objective *J*, and the constraints *l* and β, are eliminated, since the design variable cannot affect them. Similarly, the control subproblem has one variable *p*, two constraints *l* and β, and the control objective *J*. Solving this hierarchically partitioned concurrent problem is a new optimization strategy, hereafter called the *Partitioning Strategy.*

(a)

	b	*d*	*p*	*v*
f		■		
J			■	■
g	■	■		
l			■	
β	■		■	■
γ		■		■

(b)

	b	*v*	*d*	*p*
f			■	
J		■		■
g	■		■	
g		■	■	
b	■	■		■
l				■

Figure 5. Functional Dependency Tables for (a)Unpartitioned and (b)Partitioned Problem

Solution of a partitioned problem requires a coordination strategy. For a rudimentary problem, such as the current example, most coordination strategies will converge to a minimum. General partitioned combined embodiment design and control problems, however, may not converge so easily and will require more careful consideration for selecting a good coordination strategy.

The partitioned problem is solved with a standard sequential quadratic programming algorithm. Though this coordination strategy is not particularly efficient, (Table 3 shows a numerical comparison of the effort needed for convergence), it does find the same solution as the original non-partitioned problem. In the next section this decomposition strategy will be compared to the other four strategies in further pursuit of an explanation of the differences in the results obtained by them.

STRATEGIES VIEWED AS DECOMPOSITIONS

Our earlier discussion demonstrated that only one of the five combined embodiment design and control strategies found the optimal system solution. This section proposes a reason for the failure of the other strategies: the strategies are missing parts

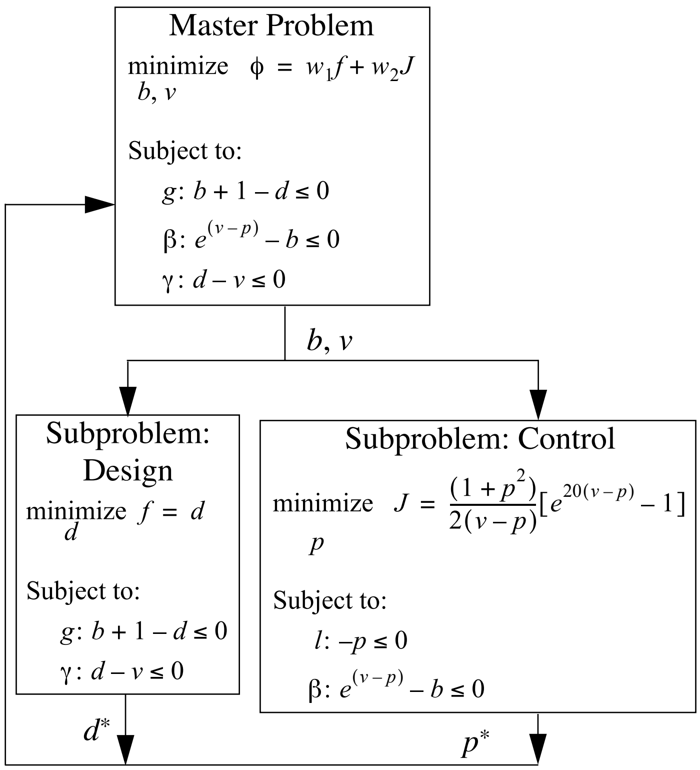


Figure 6. Partitioned concurrent strategy

of the system optimization problem. The partitioning strategy is compared with the other four, single pass, iterative, fixed input parameter, and bilevel, to determine the missing components in the formulations. The discussion describes the missing modeling elements and examines how adding these elements would affect the complexity of the strategy.

Single Pass Strategy

The single pass strategy does not and cannot achieve the system optimum. One might expect that the deficiency in the strategy is the a priori fixing of one coupling quantity — generally the input to the design problem, b. Table 1 indicates otherwise. Even when b is set to the value of the system optimum, the single pass strategy does not achieve the system optimum.

Comparison with the partitioning strategy explains why the single pass strategy does not achieve the system optimum, even with a perfect b_{in}. The single pass model maintains the design and control as separate problems. These problems closely resemble the design and control subproblems in the partitioning strategy. However, the separate subproblems are missing key constraints, namely, the coupling constraints β and γ. Instead, the constraints are imposed on the *results* of the subproblem optimization. As such, when the β constraint is active from a system perspective, the result of the constraint being ignored during optimization is that the constraint may be violated — see the results in Table 1. A similar difficulty would

Table 3: Comparison of Number of Calculations in Partitioned Problem

	AAO	Partitioned
Main function calls	134	25
Main gradient calls	25	9
Total calls to design analysis	134	251
Total calls to control analysis	134	125

occur when the γ constraint, forced to be active by the strategy, is inactive from a system viewpoint. The forced activity and inactivity for the coupling constraints β and γ, enforced outside the optimization model, is an important disadvantage of the strategy.

The single pass strategy does not have the equivalent of a master problem. Without a master problem the strategy cannot be used to examine the trade-offs between the design and control subproblems. No consideration is given to the possibility that a small concession in one may allow a large improvement in the other. Combined with the requirement of fixing one coupling quantity, this lack of a master problem to coordinate the optimization of the two subproblems, explains why the single pass strategy does not achieve system-optimal results.

Iterative Strategy

As another sequential method, the iterative strategy is similar to the single pass one. There is no master problem and the coupling constraints are enforced outside the subproblem optimization, even though the coupling quantity b is allowed to vary.

In the example, both coupling constraints *are* active at the system optimum. One might assume that the system optimum is achievable when the correct constraint activities are identified, but that does not turn out to be true. The lack of a master problem objective might be an explanation of failure, but that is also not the case. Indeed, it is the definition of the optimization problem that leads to different results. Each subproblem must be optimized with no coordination between them. In the example, the optimal design subproblem solution is at the system optimum but the optimal control subproblem is not. If the optimal control subproblem is solved, the optimal system problem cannot achieve the system optimum. Thus, the requirement in the iterative strategy that the designed artifact have both an optimal design and an optimal control causes the method to miss the system optimum.

Fixed Input Parameter Strategy

A comparison to the partitioning strategy does not gain much information about the fixed input parameter strategy. This strategy is the same as the concurrent one, with the coupling quantity b fixed. The fixed input parameter strategy is capable of

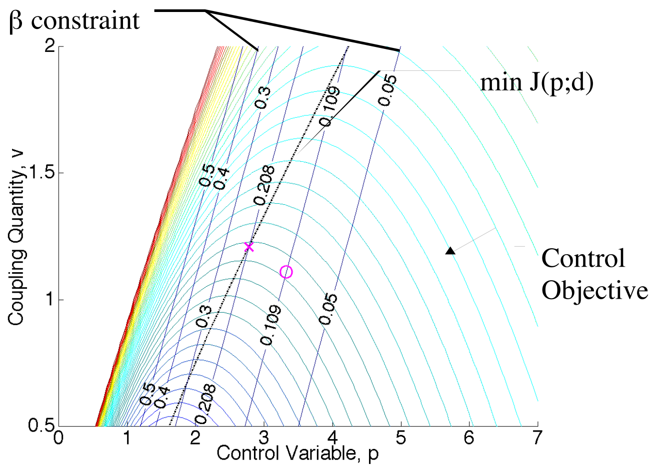


Figure 7. Control Subproblem

achieving the system optimum (see Table 1), when the fixed parameter is fixed at its system optimum value. Though formal optimization is the correct way to determine the best value for the parameter, a sampling of a few points could lead to a significant improvement in the solution without the overhead of an entire optimization. The fixed input parameter strategy's achievements are only limited by the value to which the parameter is fixed.

Bilevel Strategy

The bilevel strategy closely resembles the partitioning strategy. Though the design subproblem does not exist in the bilevel strategy, the information from that subproblem is incorporated into the master problem. The control subproblem, while similar, does not coincide exactly. The bilevel strategy's control subproblem does not include the coupling constraint β. With the coupling constraint considered only at the master problem level, it cannot bound the bilevel solution. Thus, for any value of v, a parameter to the control subproblem, p is found on the line "min $J(p;d)$," Figure 7. If it falls to the left of the system constraint β, that point is not feasible at the system level. The bilevel solution will generally not match the system optimum when the coupling constraint β is active.

SUMMARY

The combined design and control problem was modeled by a multiobjective formulation that includes optimal control as an optimal gains problem. The concurrent (all at once) solution was taken as the correct strategy for finding the true system optimum. Using a very simple demonstration example we have shown that several other strategies will not always yield the true system optimum.

The discrepancies were traced to identifying the correct activity of the coupling constraints, as well as to the actual formulation of the combined problem implied by each alternate strategy. Monotonicity analysis and a decomposition framework allowed us to analyze the situation rigorously and derive a deeper insight. Although a general theory does not readily emerge from this limited study, the results here point to potentially profitable directions of future research. Moreover, the results here show that the combined design and control problem presents significant challenges even in its simplest form, so system design should be conducted with an appropriate degree of caution.

ACKNOWLEDGMENT

This research has been partially supported by the Automotive Research Center, a US Army Center of Excellence at the University of Michigan. This support is gratefully acknowledged.

REFERENCES

Hale, A. L., Lisowski, R. J., and Dahl, W. E., 1985, "Optimal Simultaneous Structural and Control Design of Maneuvering Flexible Spacecraft," *Journal of Guidance, Control, and Dynamics*, vol. 8-1, pp. 86-93.

Khot, N. S. and Abhyankar, N. S., 1993, "Integrated Optimum Structural and Control Design," in *Structural Optimization: Status and Promise*, vol. 150, Progress in Astronautics and Aeronautics, (M. P. Kamat, Ed.), Washington, D.C.: American Institute of Aeronautics and Astronautics, pp. 743-767.

Krishnamachari, R., 1996, *A Decomposition Synthesis Methodology for Optimal Systems Design*, Doctoral Thesis, University of Michigan, Department of Mechanical Engineering, Ann Arbor.

Lewis, F. L. and Symos, V. L., 1995, *Optimal Control*, John Wiley & Sons, New York.

Messac, A. and Turner, J. D., 1984, "Dual Structural-Control Optimization of Large Space Structures," *AIAA Dynamics Specialists Conference*, Paper 84-1042, Palm Springs CA.

Papalambros, P. Y. and Wilde, D. J., 1988, *Principles of Optimal Design*, Cambridge University Press, New York.

Park, J.-H. and Asada, H., 1994, "Concurrent Design Optimization of Mechanical Structure and Control for High Speed Robots," *Journal of Dynamic Systems, Measurement, and Control*, vol. 116, pp. 344-356.

Reyer, J. A. and Papalambros, P. Y., 1999, "Optimal Design and Control of an Electric DC Motor," *ASME Design Engineering Technical Conferences*, paper number DAC-8599, Las Vegas, Nevada, September 12-15.

Wagner, T., 1993, *A General Decomposition Methodology for Optimal System Design*, Doctoral Thesis, University of Michigan, Department of Mechanical Engineering, Ann Arbor.