Multidisciplinary Design Optimization of Dynamic Engineering Systems

James T. Allison and Daniel R. Herber

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

DOI: 10.2514/1.J052182

Dynamic engineering systems are playing an increasingly important role in society, especially as active and autonomous dynamic systems become more mature and prevalent across a variety of domains. Successful design of complex dynamic systems requires multidisciplinary analysis and design techniques. While multidisciplinary design optimization has been used successfully for the development of many dynamic systems, the established multidisciplinary design optimization formulations were developed around fundamentally static system models. General multidisciplinary design optimization approaches that address the specific needs of dynamic system design are still lacking. In this article, the use of multidisciplinary design optimization for dynamic system design is reviewed, associated challenges are identified, related efforts such as optimal control are discussed, and a vision for fully integrated design approaches is presented. Finally, a set of exciting new directions that provide an opportunity for fundamental work in multidisciplinary design optimization is laid out.

I. Introduction

DYNAMIC systems, or system state evolution through time, is an increasingly important aspect of systems designed by engineers. Most notably, “smart” engineering systems that are actively controlled via electronic feedback mechanisms are becoming exceedingly prevalent, and the dynamic behavior of these systems is core to system value (e.g., renewable energy systems and vehicle electrification). In addition, many groups now recognize the importance of autonomous [1] and semiautonomous [2] dynamic systems across several domains, including manufacturing and its impact on economic competitiveness [3,4]. Active and autonomous systems, however, pose special design challenges. Physical elements of active systems need to be designed differently than for passive dynamic systems. Physical dynamics and control systems should be designed in an integrated way to achieve the best possible system performance, and sometimes integrated design approaches are required to obtain feasible designs for especially demanding dynamic systems.
The value of integrated design approaches for mechatronic and other active systems has long been recognized [5,6], especially for systems with strong coupling between physical- and control-system design (e.g., flexible robotics [7–9]). From a dynamics perspective, the design of the physical elements of a system and its control system are tightly integrated, yet compartmentalized design processes developed for the creation of passive physical systems are still in widespread use. Often, a sequential design process is used where the physical system is designed first (often relying on legacy design objectives and processes for mechanical or other physical systems) followed by control-system design. Fixing physical design before moving on to control design reduces design flexibility, produces an artificially small feasible design domain, and (except in rare circumstances) produces suboptimal results. In extreme cases, engineers may be unable to find any design that meets system requirements using a sequential approach. Integrated approaches, however, can lead to system-optimal designs by exploiting synergy between physical- and control-design decisions and, in some cases, enable solution of previously unsolvable problems [10].

Adopting integrated dynamic system design approaches has a clear benefit, but it has proven to be challenging. Organizational and technical issues that drive the transition toward integrated design methods difficult. Even with extensive integration efforts, some system interactions may be overlooked, resulting in reduced system performance. For example, modern agricultural harvesters maintain header height within a narrow window in order to harvest crops effectively (a header is the component in front of the vehicle that is designed to harvest a particular crop). Difficulty in controlling header height is one factor that limits harvester speed. It was discovered recently that further performance improvements cannot be obtained via control-design changes [11]. The interaction between physical-system and control-system designs was not fully accounted for in the original vehicle design, and physical-system redesign is necessary to achieve better performance.

Overlooking important interactions in a dynamic system model used for design may also result in unexpected results and, in some cases, spectacular failures. For example, at the June 2000 opening of the Millennium Bridge in the United Kingdom, pedestrians excited lateral vibrations in this passive dynamic system. These vibrations required pedestrians to walk with a synchronized teetering motion to maintain stability, amplifying lateral vibrations and rendering the bridge unusable [12–17]. While designing the bridge, engineers accounted for the effect of pedestrians walking on the bridge, but they did not account for the effect of bridge motion on pedestrians. Limited redesign [18–21] was performed to attenuate lateral vibrations using active control, but more comprehensive integrated modeling and design processes early on may have prevented the need for expensive redesign. Similar unexpected behaviors that hamper dynamic system design efforts across a broad range of engineering domains may be avoided by developing models with more complete dynamic interactions and by developing and adopting improved design methodologies created specifically for dynamic systems.

In addition to organizational issues, such as compartmentalized legacy design processes and difficulties in managing system interactions, several technical challenges exist as well. Optimization-based approaches for integrated dynamic system design require sophisticated system models. Ideally, these models provide accurate representation of full system dynamics, are validated and accurate [22,23], and are computationally efficient. Models for complete system design should also provide flexibility in both physical- and control-system design spaces [24], as well as incorporate multiple disciplines and important system interactions [25–29]. Models have greater utility in design when fidelity can be adjusted to accommodate different needs at different design phases [30]. Many of these modeling objectives are competing, so tradeoffs must be made. Extensive efforts have been made to improve the value of models in dynamic system design. For example, surrogate models that approximate high-fidelity models are computationally less expensive to evaluate, but the expense of sampling required to build surrogate models may be significant [31–34]. Model reduction techniques may also be used to reduce the computational expense of dynamic system models [34–37], but the resulting models have reduced accuracy. Scaling techniques have been investigated as a way of providing flexibility in the physical design space [38,39] without needing to develop physics-based models, but they are often valid only over a fairly limited domain.

Models need to account for complete system dynamics while allowing for changes in physical-system design. Often, models developed for control-system design provide outstanding predictions of system dynamics, but they are based on a fixed physical-system design. For example, engineers developing a control system for an electric motor can use a dynamic model based on measurable physical parameters, such as inductance or resistance. This is fine as long as the physical design does not change. If it does, the parameters must be identified and validated again. If another set of engineers is developing the physical design of a motor, they would need a different type of model that can predict system behavior based on independent physical design variables (variables that engineers have direct control over), such as geometric dimensions. Models that are developed for physical-system design allow for physical design changes, but they often are based on simplified dynamics or static analysis. Achieving full design space flexibility simultaneously with an accurate representation of system dynamics normally requires a substantial investment in model development, and it can be a bottleneck in the adoption of fully integrated dynamic system design.

While the transition to integrated design methods for dynamic systems may seem replete with obstacles, this sort of transition is not without precedent. The coupling between physical-system design and control-system design is analogous to the coupling between product design and manufacturing. In both cases, the conventional approach is sequential, which often proves to be restrictive and inefficient. Design for manufacture (DFM) has successfully addressed the coupling between product design and manufacturing by accounting for manufacturing needs during product development [40]. Achieving integrated design of dynamic systems will require an effort similar in magnitude to the effort that was needed to develop successful DFM methods.

Multidisciplinary design optimization (MDO) offers a solid framework for moving dynamic system design theory and practice toward a more fully integrated state. Traditionally, MDO work has aimed to integrate previously independent analysis and design activities to improve engineering system performance and reduce development costs [41,42]. In this article, we will review how MDO has been used in dynamic system design, explore opportunities for enhanced design capabilities based on MDO, and identify promising new directions for fundamental work in MDO.

While numerous dynamic systems have been designed using MDO methods, the established MDO formulations largely are based on static system analysis or black-box simulations, and they do not address system dynamics explicitly. As a result, the nuances of dynamic behavior are implicitly deemphasized, and challenges related to system dynamics must be addressed on a case-by-case basis. The ever-growing scale and complexity of modern dynamic systems are taxing conventional design methodologies. Fundamentally different MDO formulations that embrace time-dependent behavior and address system dynamics directly are needed to meet these demands and realize new dynamic system capabilities.

Here we define multidisciplinary dynamic system design optimization (MDSDO) as a branch of MDO that deals with systems where the evolution of system state through time is a critical element of performance; where multiple disciplines, energy domains, models, or subsystems must be integrated; and where the unique properties of dynamic systems are exploited to improve system performance and yield efficient problem solutions. Active systems (systems that use active control to govern behavior) are playing an increasingly important role in society, and they are a particularly important application of MDSDO.

We acknowledge the extensive work related to dynamic system design performed in fields such as optimal control [43,44], robotics [45], structural dynamics [46,47], and cyberphysical systems [48]. Each of these areas tackles an important piece of the larger dynamic system design problem. Here we aim to bring these and other
disciplines together, explore their complementary relationships, offer a high-level perspective regarding the future of dynamic system design, and develop a comprehensive vision for MDSDO. Section II discusses more deeply the elements of MDSDO, including multidisciplinary analysis, optimal physical-system design, optimal control, and integrated dynamic system design methods. Section III reviews how MDO has been applied so far to dynamic system design. Sec. IV outlines research directions required to build more complete theory and tools for MDO-based dynamic system design, and Sec. V offers concluding remarks.

II. Multidisciplinary Dynamic System Design Optimization

Here we are concerned with the design optimization of engineering systems where the evolution of system state through time is central to the functionality or value of a system. These dynamic engineering systems may be passive (time variation due only to natural system dynamics) or active (controlled via electronic feedback). In most real dynamic engineering systems, multiple interacting energy domains are involved, such as mechanical, thermal, electronic, hydraulic, etc., so multidisciplinary analysis (MDA) is required for successful design. These MDA models may be used to support physical- or control-system design decisions, or support integrated design approaches that consider physical- and control-system design decisions simultaneously. In this section, the current state of each of these topics will be reviewed.

A. Multidisciplinary Analysis of Dynamic Systems

In performing MDA for dynamic systems, we aim to capture the effects of each energy domain on the dynamics of the other domains. For example, when analyzing robotic systems, we must account for the coupling between mechanical and electrical dynamics. Otherwise, an independent mechanical system model might predict incorrectly that the electric actuator state could change instantly. Suppose we have two energy domains, $a$ and $b$, with the respective state variable trajectories $\xi_a(t)$ and $\xi_b(t)$. The dynamic response of each domain can be modeled using its own set of governing differential equations, but a more complete multidisciplinary model accounts for the dynamic interaction between $a$ and $b$. The resulting coupled system of differential equations is:

$$\dot{\xi}_a(t) = f_a(\xi_a(t), \xi_b(t), u_a(t), t)$$

$$\dot{\xi}_b(t) = f_b(\xi_a(t), \xi_b(t), u_b(t), t)$$

(1)

(2)

where $f_a(\cdot)$ and $f_b(\cdot)$ are the derivative functions for each domain, and $u_a(t)$ and $u_b(t)$ are control inputs that are present if the system is actively controlled. Observe that $\xi_a(t)$ depends on $\xi_b(t)$, and $\xi_b(t)$ depends on $\xi_a(t)$. The fully integrated system model can be represented more compactly if we define $\xi(t) = [\xi_a(t), \xi_b(t)]^T$, $u = [u_a(t), u_b(t)]^T$, and $f = [f_a(t), f_b(t)]^T$:

$$\dot{\xi}(t) = f(\xi(t), u(t), t)$$

(3)

The multidisciplinary dynamic system model given in Eq. (3) may be constructed using one of several well-developed strategies, such as bond graph modeling for lumped-parameter system models [29] or high-fidelity multiphysics models for systems involving continuum mechanics [26–28]. If a commercial tool that integrates all the desired domains is unavailable, then a possible solution is to integrate separate software tools. For example, a block-diagram modeling environment appropriate for control-system modeling may be incorporated with a multibody dynamics model of a mechanical system using a cosimulation approach [49].

In the development of dynamic engineering systems, we would often like to impose constraints on state trajectories (i.e., path constraints) for design requirements or modeling expedience. Adding an equality constraint to a system of differential equations without adding a new state variable produces a system of differential algebraic equations (DAEs) [50]. A DAE in semisimple form is:

$$\xi(t) = f(\xi(t), \gamma(t), u(t), t)$$

$$0 = g(\xi(t), \gamma(t), u(t), t)$$

(4a)

(4b)

where $f(\cdot)$ is an algebraic constraint, and $\gamma(t)$ is the algebraic variable (i.e., its time derivative $\dot{\gamma}(t)$ does not appear in the equations). Most DAE algorithms require that Eq. (4b) can be solved for $\gamma(t)$ (i.e., the Jacobian of the algebraic constraint must not be singular). A DAE that satisfies this requirement is an index-1 DAE, where the index identifies the number of differentiations required to transform a DAE into an ordinary differential equation.

Inequality path constraints are sometimes needed to model a dynamic system design problem (e.g., temperature, position, or force limits). If these bounds are reached, then algebraic inequality constraints become active. Constraints may enter or exit activity multiple times during a simulation. An ODE can be transformed into a DAE when an inequality path constraint becomes active. Imposing a new algebraic relationship like this reduces a system’s degrees of freedom. For every lost degree of freedom, a state variable must become an algebraic variable, i.e., a variable completely determined by state variables via the algebraic constraint. In active systems, control inputs normally become the algebraic variable since they are independent, while state variables must satisfy physics [51]. As additional inequality constraints become active, the DAE index may increase, increasing solution difficulty.

B. Optimal Control-System Design

The value of MDA models extends beyond the analysis of existing systems; they are important for optimal design of new systems. Simulations of MDA models predict system behavior given its specifications, but they can also be used for the inverse task (i.e., design): identifying a system specification that produces desired behavior. Physical-system design and control-system design both contribute to overall dynamic behavior in actively controlled systems. While physical- and control-system design normally are tightly coupled, they are often treated separately in conventional sequential design processes. Here we begin our exploration of dynamic system design with a brief review of optimal control, a design approach that aims to identify the control design that produces the best possible system performance.

In optimal control, a control design is sought that optimizes a cost function, often of this form:

$$\phi(\xi(t), u(t), t) = \psi(\xi(T_f), t_f) + \int_0^{T_f} L(\xi(t), u(t), t) \, dt$$

(5)

where $\psi(\cdot)$ and $L(\cdot)$ are the Mayer (terminal cost) and Lagrange (running cost) terms, and $T_f$ is the length of the time horizon considered in the design problem. If both terms are present, the function is often termed a Bolza objective. The problem is an open-loop control problem if the control trajectory $u(t)$ is the optimization variable. Observe that optimal control problems are solved with respect to an infinite-dimensional control trajectory, as opposed to a finite-dimensional optimization vector used in typical MDO formulations.

A classical approach for solving optimal control problems is to apply optimality conditions, such as Pontryagin’s maximum principle (PMP) [43,44], to identify the optimal control trajectory $u(t)$ that minimizes $\phi(\cdot)$. If an analytical solution to the optimality conditions cannot be found, the resulting boundary value problem (BVP) often can be solved numerically. This approach is an optimize-then-discretize approach, since a BVP obtained via optimality conditions is discretized and then solved ([52] p. 310).

In most practical implementations, we need to design a feedback controller (and often we need to design an observer to estimate states that cannot be measured directly). A simple form of feedback control is a full-state feedback regulator, where the control input is defined as $u(t) = -K\xi(t)$. Assuming this control structure, the optimal control problem may be solved with respect to the gain matrix $K$ instead of
the control trajectory. In addition, if the system model is linear and time invariant, i.e., the derivative function can be written as:

$$\dot{\xi}(t) = f(\xi(t), u(t)) = A\xi(t) + Bu(t)$$  (6)

and if $f(\cdot)$ is quadratic, then a closed-form solution for $u(t) = -K\cdot \xi(t)$ can be derived. The resulting optimal control law is known as a linear quadratic regulator (LQR) ([53] p. 337).

Optimal control approaches based on PMP are known as indirect methods. Direct methods are an alternative approach where optimal control problems are discretized first and then transcribed to a nonlinear programming (NLP) formulation. In other words, an infinite-dimensional optimal control problem is transcribed to a finite-dimensional NLP [52,54,55]. Direct transcription (DT) is a family of discretize-then-optimize methods for optimal control that use this strategy ([52], [56] p. 178). DT is a special case of the all-at-once (AAO) MDO formulation [57], also known as simultaneous analysis and design [58]. In DT, an NLP algorithm simultaneously solves the system state equations and the system optimization problem, eliminating the need for forward simulation. This is accomplished by applying a numerical integration method (such as collocation [50,52]) to convert differential state equations to a system of algebraic equations and discretizing the state and control trajectories. The resulting algebraic equations, known as defect constraints $[\xi(\cdot)]$, are posed as equality constraints in the optimization problem, and the discretized state and control trajectories are treated as optimization variables. If $U$ is a matrix where row $i$ is the control vector at time step $i$ (i.e., $u_i$) and $\Xi$ is a matrix where row $i$ is the state vector at time step $i$ (i.e., $\xi_i$), then the following is a DT formulation for optimal control:

$$\begin{align*}
\min_{U, \Xi} & \sum_{i=0}^{n-1} L(u_i, \xi_i)h_i \\
\text{subject to} & \quad \xi(U, \Xi) = 0
\end{align*}$$  (7)

Here, only the Lagrange cost term is included in the objective, $h_i$ is the time-step size at step $i$, and $n_i$ is the number of time steps. Note that incorporating $U$ and $\Xi$ as optimization variables increases problem dimension.

Historically, DT has been applied only to open-loop optimal control (trajectory optimization in particular [56,59–62]), but recently, it has been extended to more general dynamic system design problems including nonlinear feedback control design [10,63,64], dynamic system design without control [65], and integrated physical-system and control-system design (co-design) [24,66,67]. DT is related closely to other discretize-then-optimize techniques, such as pseudospectral methods [60,63,68–74], adjoint state methods [75–77], and temporal spectral element methods [78,79], as well as model predictive control [80]. A number of commercial [81–86] and open-source [69,87,88] DT software implementations are available.

The defect constraints in DT are solved simultaneously, meaning that forward simulation is not required to obtain state trajectories. Higher-order implicit quadrature methods are normally impractical for forward simulation, but they work well when the resulting defect constraints are instead solved simultaneously. When these higher-order methods are used with DT, high solution accuracy can be maintained even with large time steps, and larger step sizes reduce optimization problem dimension [52,62,89].

Even with large time steps, DT optimization problems still have much higher dimension than other discretize-then-optimize approaches. Why then would one consider using DT? First, DT optimization problems have special structures that promote efficient computation: in some cases, even exponential convergence [68]. Optimization variables appear explicitly in constraint functions, making sensitivities easier to compute, and when analytical derivatives are impractical to obtain, the Jacobian sparsity pattern enables efficient application of sparse finite differences [56]. In addition, defect constraints are independent, enabling fine-grained parallel computing. For linear dynamic systems, DT problem formulations are often either quadratic or linear programs, allowing for especially efficient problem solution.

Another reason to consider DT is the ability to impose inequality constraints on trajectories, something that generally cannot be done with indirect methods. Equality path constraints are also easily included, extending applicability to DAE systems. DT also works well on challenging singular optimal control problems, and it is often successful at maintaining numerical stability when solving highly nonlinear problems [56,67,90].

DT possesses the unique property that system dynamics are represented directly in the optimization formulation, and it offers one promising direction for development of MDO formulations for dynamic system design. DT will be revisited in greater detail in Sec. IV after additional context is developed, with particular emphasis on extension of DT to co-design applications.

Dynamic engineering systems are often multidisciplinary, and MDA techniques are needed to model interactions across multiple energy domains [cf. Eq. (3)]. In practice, multidisciplinary dynamic system models are used widely, but design is usually limited to a single discipline. Even if interactions across multiple domains are modeled with great sophistication, design efforts that concentrate on dynamic system performance typically address control design only. For example, many aeroelasticity design studies (e.g., [91–93]) use advanced multidisciplinary models that capture complicated aerodynamic and structural interactions, but the physical system is held fixed while the control system is designed. This addresses only part of the dynamic system design problem. Physical-system design has an important, if not dominant, influence on system dynamics. By some definitions of MDO, optimal control studies do not constitute MDO, even if MD is used, since the design component involves only one discipline. A systems-oriented approach incorporates multidisciplinary design in addition to MDA.

### C. Optimal Physical-System Design

The role of physical-system dynamics should be a core consideration in dynamic engineering system design. In other words, the onus of optimizing dynamic system performance rests also on engineers designing physical elements of the system, and not just the control-system engineers. When making physical-system design decisions, we should include comprehensive treatment of system dynamics if we want to capitalize on passive dynamics. Many physical-system design optimization efforts, however, incorporate simplified system dynamics (such as steady-state or pseudostatic models) or static analysis that neglects dynamic effects altogether. Design objectives are often approximations of actual dynamic system performance metrics (e.g., mass [94] or gravity balance [95]). While these simplifications are sufficient in some cases, performance can be improved by using more complete dynamic models when designing physical systems, and improved models will also enhance the ability to design more challenging dynamic systems.

While comprehensive dynamic models are in use, they are normally developed for control design, and they do not allow for physical design changes. The next generation of system models needs to incorporate realistic dynamics while providing flexibility in the physical design space, i.e., we need models of the form:

$$\dot{\xi}(t) = f(\xi(t), \mathbf{x}_p, t)$$  (8)

where $\mathbf{x}_p$ is a vector of physical-system (or plant) design variables. Models of this type require more development effort than models with a fixed physical-system design [e.g., Eq. (3)].

Consider the following passive physical dynamic system design optimization problem:

$$\begin{align*}
\min_{\mathbf{x}_p} & \quad \phi(\xi(t), \mathbf{x}_p, t) \\
\text{subject to} & \quad g_p(\xi(t), \mathbf{x}_p) \leq 0 \\
& \quad \dot{\xi}(t) = f(\xi(t), \mathbf{x}_p, t) = 0
\end{align*}$$  (9)

Here, the system objective function is minimized with respect to $\mathbf{x}_p$ only. We also introduce a new function, $g_p(\cdot)$, that quantifies physical-system constraints such as stress, deflection, or geometric requirements ([96] p. 13). This function depends on both physical
design and state variables, accounting fully for the influence of dynamic response on physical design requirements. Other more simplified design formulations neglect direct dependence of plant constraints on $x(t)$. The objective function used here is the overall system objective that depends on dynamic response, as opposed to a static or simplified dynamic physical design proxy objective. Variants of the formulation given in Eq. (9) have been studied; Wang and Arora reviewed methods for solving this class of problems [65]. These formulations were based on DT, where states were discretized and treated as independent variables (another example of DT applied to passive system design is presented in [97]). This class of problem formulation has been used extensively for design optimization of structures while accounting for dynamic considerations. Kang et al. provided a review of approaches for optimizing structures subject to transient loads [47], and Barthelemy and Hafika reviewed approximation methods relevant to this problem class [98]. Finally, others have used this formulation approach in designing more general mechanical systems [99–101].

D. Optimal Dynamic System Design

Approaches for optimizing the physical- and control-system designs of dynamic systems separately were just reviewed. Independent solution approaches are required to capitalize on the synergistic relationship between physical- and control-system designs. Conventional sequential system design approaches [6,46,102–105] only account partially for coupling between physical-system (plant) and control-system design decisions, producing suboptimal results [106]. In sequential design, control design is performed after plant design is complete. If optimization is employed for each task, the sequential approach consists of solving Eq. (9) to obtain the optimal plant design $x_{p^*}$, which is then used as the basis for solving the optimal control problem [minimizing $\phi_{c}^*(t)$] from Eq. (5)] to obtain the optimal control trajectory $u_{c}(t)$. This process is illustrated in Fig. 1.

While this sequential process often produces feasible system design, better methods exist. As actively controlled systems become more complex and performance requirements more stringent, sequential system design may fall short, motivating the use of more integrated design methods.

The sequential system design process illustrated in Fig. 1 represents the case where plant design is based on passive dynamics. For example, in designing a passive–active automotive suspension [107], we may start by designing the passive suspension (i.e., mechanical linkage, spring, damper, etc.) to optimize comfort and handling, but with no control actuation. The resulting plant design may then be used as a basis for designing the active control system (holding plant design fixed). Ideally, we use the same system objective function $\phi(t)$ in both design phases. Often, due to legacy design practices for physical systems, an alternative plant-design objective function $\phi_{p}(t)$ is used instead (e.g., mass). The effects of using separate plant and control objective functions are discussed next.

We can classify plant-design objectives into five types:

Case 1: The first type of plant objective accounts only for passive plant dynamics, and is based on the original system objective function $\phi(x(t), x_{p})$.

Case 2: The second type of plant objective accounts only for passive plant dynamics, and is based on an approximation of the original system objective function $\phi(x(t), x_{p})$.

Case 3: The third type of plant objective does not account for plant dynamics (i.e., a static analysis), and is based on an approximation of the original system objective function $\phi(x(t))$.

Case 4: The fourth type of plant objective accounts for active plant dynamics, and is based on the original system objective function $\phi(x(t), u(t), x_{p})$.

Case 5: The fifth type of plant objective accounts for active plant dynamics, and is based on an approximation of the original system objective function $\phi(x(t), u(t), x_{p})$.

Case 1 corresponds to Fig. 1. Case 2 involves an approximate system objective function that still depends on the complete system dynamics, meaning that a dynamic system simulation is required to evaluate $\phi(\cdot)$ but is still limited because it does not incorporate active control. In addition, some systems cannot be simulated without active control, so cases 1 and 2 are not always available options. For example, an active automotive suspension [24] may be simulated in a passive mode, but a robotic manipulator requires control actuation to simulate. Bowling et al., however, did introduce “dynamic capability” equations based on system dynamics that guide physical robot design toward improved active dynamic performance without requiring control design [108].

Case 3 is a more significant simplification where static- or frequency-based analysis eliminates the need for simulation. The objective depends only on plant design. Ravichandran et al. presented an example of a case 3 sequential design approach for reducing energy consumption of a counterbalanced robotic manipulator [95]. For very slow pseudostatic movements, energy consumption is approximately minimized if the manipulator is designed to have perfect gravity balance (i.e., any manipulator position can be held with zero actuation torque). Using gravity balance as the plant-design objective, $\phi_{p}^*(t)$ simplifies the problem, but it is inaccurate for high-speed motions [109]. Allison presented a more complete formulation that produces system-optimal results for high-speed counterbalanced manipulators [21]. Trivedi et al. also used a case 3 approach for soft robotic manipulator design where a static model was used for physical–system optimization [105].

The commonly used case 2 and case 3 objectives arise when separate objectives for plant and control design are specified. Several researchers have asserted that co-design problems are fundamentally multiojective [110–112]. While co-design problems may indeed be multiojective because of intrinsic tradeoffs in the system (e.g., cost vs performance), a problem is not automatically multiojective because it is a co-design problem. When separate plant and control objectives are used, the plant objective is often an approximation of the real system objective (e.g., gravity balance approximating energy efficiency). Separate plant objectives may also be used because of legacy design processes. For example, when physical design is performed in isolation, using a plant objective that is not directly connected to dynamics or active control (such as mass or other static measures) is a logical choice. These legacy design paradigms, however, are firmly established. Abandoning familiar design objectives and adopting objectives that more accurately reflect overall system purpose may be challenging when working to adopt an integrated systems design approach. Part of designing with a holistic systems perspective is to develop system components that, when combined together, produce the best overall system behavior, as opposed to optimizing the components individually. Integrated system design requires consistent use of the same system objective across all system elements (or the same set of objectives if the system design problem is inherently multiojective). Cases 1 and 4 are examples of approaches that use a common objective.

Cases 4 and 5 are fundamentally different from the others in that the objective depends explicitly on control design. Control input is considered during plant design, but it is held fixed. Incorporating the effects of active control improves solution quality. It also opens up the possibility of an iterated sequential design process where, after completing a single pass of sequential design, we can feed $u_{c}(t)$ back into the plant-design problem and iterate. Pil and Asada demonstrated a modified form of case 3 that allows for iteration and incorporates physical prototyping guided by control-design sensitivity data [113], and Padula et al. introduced a three-stage...
iterated sequential method that includes plant-, control-, and system-level design [114]. The iterated sequential method based on case 4 is a special case of the block coordinate descent (BCD) optimization method [115] p. 272. To understand this connection with BCD, suppose we have optimization problem \( \min_{x} f(x) \), \( x \in X \), where \( X = X_1 \times X_2 \times \ldots \times X_m \) and each \( X_j \) is a closed convex set. The optimization vector may be partitioned into “blocks” of coordinates: \( x_j \in X_j \) and \( j = 1, \ldots, m \). An optimization subproblem for each coordinate block (\( j = 1, \ldots, m \)) may then be defined: \( \min_{x_j} f(x) \), \( x \in X_j \). Each subproblem may either be solved simultaneously (Jacobi iteration) or in sequence using the most recently updated values for \( x \) (Gauss-Seidel method), and then iterated. BCD converges to the solution of the original problem if each subproblem has a unique solution. When the iterated case 4 sequential design approach is used, it is a BCD solution to the fully integrated plant and control-design (co-design) optimization problem:

\[
\min_{x_p, u(t)} \phi(\xi(t), u(t), x_p, t) \\
\text{subject to:} \quad g_p(\xi(t), x_p) \leq 0 \\
\xi(t) - f(\xi(t), u(t), x_p, t) = 0
\]

The solution to (10) is the system-optimal design; it accounts for all dynamic system interactions and plant–control-design coupling, resulting in minimal \( \phi(\cdot) \). This solution is our standard of comparison for all active system design methods. Note that the plant-design constraints do not depend directly on \( u(t) \), but they are influenced indirectly by control design through state trajectories. The formulation in (10) is often referred to as the simultaneous co-design method, since plant- and control-design decisions are made simultaneously.

In a BCD solution of (10) using the Gauss-Seidel method (iterated sequential method case 4), the number of coordinate blocks is \( m = 2 \); \( j = 1 \) corresponds to the plant-design coordinate block, and \( j = 2 \) corresponds to the control-design coordinate block. If we assume for generality that vector \( x \) is a discretization [e.g., \( U \) from Eq. (7)] or parameterization (e.g., full-state feedback gain matrix \( K \)) of the control design, the co-design problem becomes a nonlinear program. Satisfaction of the state equations is an important distinction between solution approaches, and it will be discussed in the next section. The BCD algorithm is given in Algorithm 1, where \( k \) is an iteration counter, \( \epsilon \) is a convergence tolerance, and system dynamics are satisfied implicitly via simulation. Note that, in step 1.3, the most recently updated value of \( x_p \) is used and control designs must also satisfy plant constraints.

**Algorithm 1 Co-design Using Block Coordinate Descent**

**Input:** Initial design variables \( x^1 \)

**Output:** Optimal design variables \( x \)

1. Initialize \( x^1 \) and \( \epsilon \), set \( k = 0 \)

repeat

1.1: \( k = k + 1 \)

1.2: \( x_{p}^{k+1} = \arg \min_{x} \phi(\xi(t), x_{p}^{k}, x) \), subject to \( g_p(\xi(t), x_{p}^{k}, x_p) \leq 0 \)

1.3: \( x_{p+1}^{k+1} = \arg \min_{x} \phi(\xi(t), x_K^{k}, x_{p+1}^{k+1}) \), subject to \( g_p(\xi(t), x_p, x_{p+1}^{k+1}) \leq 0 \)

until \( |x_{p}^{k+1} - x_{p}^{k}| \leq \epsilon \)

While BCD is often capable of producing system-optimal solutions, it can be computationally inefficient depending on the problem at hand. This potential inefficiency motivates alternative solution approaches that will be described in the following sections. To illustrate BCD solution efficiency, consider this simple quadratic objective function:

\[
\phi(x) = a_1(x_1 - b_1)^2 + a_2(x_2 - b_2)^2 + a_3x_1x_2
\]

III. Existing Uses of MDO for Dynamic System Design

When applying multidisciplinary design optimization to the design of engineering systems, the aim is to account for interactions between multiple disciplines (such as structural and aerodynamic analysis) or physical subsystems (such as engine and wing). A core objective of MDO is to integrate (previously independent) disciplinary analyses and design activities to yield better system performance and reduced system development time and cost [41,42,116–118].

Dynamic properties are of fundamental importance to the value of many multidisciplinary engineering systems, but application of MDO to dynamic system design has often been done only in a simplified or limited way. For example, rather than using simulation of nonlinear dynamics, simplified dynamic analysis is used, such as frequency domain analysis [119–126], steady-state analysis [127], or

pseudostatic models [95, 105, 128]. Also, in many MDA models, the interactions between disciplinary analyses are treated as static. In co-design studies, a simplified dynamic or static model is often used for physical-system design, while a more complete dynamic model is used for control-system design. This misalignment between physical and system design formulations will prevent identification of a system-optimal solution.

One important factor that complicates the use of MDO for dynamic system design is the static nature of fundamental MDO formulations. With only a few exceptions, MDO formulations have not been developed in a way that addresses system dynamics explicitly. For example, Haftka et al. explained that the analysis associated with MDO generally consists of nonlinear algebraic equations [129]. While dynamic system analysis may indeed be algebraic after discretization of differential equations, the “algebraic equation” mindset prevents deeper treatment of the unique needs of dynamic design interaction directly [9, 46, 110, 114, 123, 127, 147].

Although many MDO frameworks can be challenging to use for dynamic system design, multidisciplinary optimization of dynamic systems has been investigated extensively in specific application areas, most notably in the design of dynamic structures [47, 78, 79, 137–146]. Many studies have addressed control-structure design interaction directly [9, 46, 110, 114, 123, 127, 147–154], while some focus on passive dynamic systems [119, 155–158]. Sensor and actuator placement combined with control design is another extensively studied area of dynamic system design optimization [142–162]. Other important applications include automotive suspension and powertrain design [24, 104, 107, 163–171], robotic system design [5, 7, 8, 20, 21, 95, 108, 109, 113, 147, 172–175], and energy systems [90, 121, 176–178].

While substantial depth exists in select application areas for dynamic system design optimization, fundamental MDO formulations specifically designed for dynamic systems are largely not available. Many of the aforementioned studies employed either the basic multidisciplinary feasible (MDF) formulation [57, 179], where all analysis tasks are performed nested within a single optimization algorithm, or some form of the sequential design processes discussed previously. Often, even when multidisciplinary analysis is performed, design is only conducted within one discipline (e.g., aeroelasticity). When multidisciplinary design is performed that incorporates physical-system design, some aspect of dynamic analysis is usually simplified (e.g., co-design with static plant analysis), whereas extensive work in the area of optimal control [44] fully embraces the complexities of system dynamics but addresses only one design discipline: control.

All the components required for fully integrated dynamic system design (multidisciplinary analysis, physical-system design methodologies, optimal control, etc.) exist, but they are usually in fragmented form, except for specific case studies. Development of unifying MDO frameworks that can integrate these components would support the more general application of integrated dynamic system design, broadening the impact of MDO. This is an important opportunity for transformational progress in the design of dynamic engineering systems and an exciting direction for new fundamental work in MDO.

A. Current MDO Formulations and Dynamic System Design

We will now explore approaches for using existing MDO formulations for dynamic system design and highlight some of the difficulties that can arise. First, we will look at how the MDF formulation can be applied to dynamic system design problems, and then we will explore distributed MDO methods.

When using MDF formulation, all analysis tasks are performed within the optimization algorithm loop. MDF formulation may be used to solve the fully integrated problem defined in Eq. (10) [9, 166, 180] or parts of a sequential design problem [114]. Starting with the simplest case, consider the (potentially multidisciplinary) plant-design optimization problem introduced in Fig. 1. If we are using a case 4 objective function that incorporates active control, the MDF formulation for the plant-design portion of the sequential approach is:

\[
\min_{\xi_p} \phi(\xi(t), x_c, x_p) \\
\text{subject to: } g_p(\xi(t), x_p) \leq 0 \\
\text{where: } \dot{\xi}(t) - f(\xi(t), x_c, x_p, t) = 0
\]  

(12)

Here, the control design \( x_c \) is fixed, and the state equations are solved for the state trajectories \( \xi(t) \) using a forward simulation algorithm (such as a Runge–Kutta algorithm [52]) for every plant design \( x_p \) proposed by the optimization algorithm. In other words, system analysis is nested within the optimization problem. Solution of the state equations requires time discretization \( t_1, \ldots, t_n \), where \( t_0 = 0, t_n = t_f, h_i = t_{i+1} - t_i \), and \( n_i \) is the number of time steps). As noted previously, the state trajectory solution may be represented in matrix form \( \Xi \), where the \( i \)th row of \( \Xi \) corresponds to \( \xi_i(t) \). Also note that the state equations in Eq. (12) may span multiple engineering disciplines [cf. Eqs. (1–3)].

MDF formulation is a discretize-then-optimize approach since discretization is performed before optimization. When applied to dynamic system design problems, MDF formulation is also known as the single-shooting method [52].

Once the MDF solution to the plant-design problem is obtained \( (x_p) \), the optimal control problem may be solved either via conventional optimal control methods (e.g., optimize-then-discretize methods based on PMP) or discretize-then-optimize methods such as direct transcription or MDF formulation. These are good alternatives when the system Hamiltonian derivatives needed for PMP-based solutions are not obtained easily. The MDF formulation of the optimal control problem is:

\[
\min_{x_c, x_p} \phi(\xi(t), x_c, x_p) \\
\text{subject to: } g_p(\xi(t), x_p) \leq 0 \\
\text{where: } \dot{\xi}(t) - f(\xi(t), x_c, x_p, t) = 0
\]

(13)

Sequential design processes will not produce system-optimal solutions unless a case 4 formulation is iterated and BCD convergence conditions are satisfied. A more efficient approach is to use MDF formulation to solve the simultaneous problem defined in Eq. (10):

\[
\min_{x_c, x_p} \phi(\xi(t), x_c, x_p) \\
\text{subject to: } g_p(\xi(t), x_p) \leq 0 \\
\text{where: } \dot{\xi}(t) - f(\xi(t), x_c, x_p, t) = 0
\]

(14)

Here, the analysis is completely integrated. Often, this can be done within a single software environment, but if this is not possible, disparate analysis tools may be integrated using techniques such as cosimulation to coordinate multiple simulation environments [182].

Nested co-design is a method that may be viewed as a special case of MDF formulation [106, 123]. An outer optimization loop optimizes the plant design, and an inner optimization loop identifies the optimal control for each plant design tested by the outer loop. Note that this inner optimization loop may have a simulation nested within it if a closed-form optimal control method (such as LQR [53]) or an AAO optimal control approach (such as direct transcription) is not employed, resulting in a double nesting. One advantage of nested
co-design is the ability to use existing optimal control algorithms (e.g., LQR, DT) to solve the inner-loop problem efficiently without the complication of managing plant-design variables. The outer-loop formulation is:

$$\min_{x_p} \phi_1(x_p), \quad \text{subject to: } g_p(x_p) \leq 0 \quad (15)$$

where for every objective function evaluation, the following inner-loop problem is solved to obtain $\phi_1(\cdot)$:

$$\min_{x_t} \phi(x(t), x_c, x_p)$$

subject to: \hspace{1cm} $g_p(x(t), x_p) \leq 0$

where: \hspace{1cm} $x(t) - f(x(t), x_c, x_p, t) = 0 \quad (16)$

The function $\phi_1(\cdot)$ in the outer-loop problem is an optimal-value function that is calculated by solving the inner-loop problem. Plant design $x_p$ is held fixed during the inner-loop solution. Note that the plant-design constraints are retained (at least those influenced by state trajectories); $g_p(\cdot)$ must be retained in any implementation where plant- and control-design problems are solved separately (e.g., nested methods or the sequential methods described previously). Otherwise, design feasibility cannot be assured. Several have proposed using LQR to solve the inner-loop problem for linear systems \[46, 106, 183\]. If a detailed plant-design formulation with substantial plant constraints is used, LQR may not be practical, as it cannot manage inequality plant constraints. In this case, a discretize-then-optimize approach would be more appropriate for the inner loop. Also, by the nature of the nested co-design method, the plant and control objectives are aligned since the outer-loop objective is also, by the nature of the nested co-design method, the plant and control objectives are aligned since the outer-loop objective is aligned with control objectives are temporarily independent, enabling coarse-grained parallel computing. One of the key points here is that, in MDF formulation, the coupling variables are optimization variables, whereas in MDF formulation, they are not. We might consider using MDF formulation if the dimension of $y$ is small, while MDF formulation usually is more appropriate for densely coupled problems.

IDF formulation might be applied to dynamic system design problems in one of several ways, distinguished by how the separate analysis functions are defined. To illustrate the first of three IDF formulation approaches discussed in this paper, suppose the system analysis is in the form of a Simulink® model. We could group model components into clusters; a simulation of each cluster would comprise an analysis function, and the signals connecting the clusters would become the coupling variables. Each cluster corresponds to a portion of the derivative function, similar to the partitioned state equations in Eqs. (1) and (2). For example, the model in Fig. 3, based on the vane airflow (VAF) sensor problem in \[179\], can be partitioned into blocks used to compute torque due to air resistance and blocks used to quantify the dynamic response of the sensor vane. This partition cuts across torque $\tau$ and position $\theta$ signals (time histories are shown in Fig. 3).

The challenge with this decomposition approach is that each signal corresponds to a time history (a function-valued quantity) and not just a scalar or small vector. Signals that cut across partitions are coupling variables, adding significantly to the number of optimization variables in the IDF formulation. In the VAF example, IDF formulation requires the optimization algorithm to specify the complete torque and position trajectories so that the two subsystems may be simulated independently for each optimization function call.

If we assume that partitioned signals in this first IDF approach correspond to states, we can define $\xi(t)$ as the subset of state trajectories that cut across system partitions. Solving this IDF formulation problem using nonlinear programming requires that we discretize $\xi(t)$. The matrix $\Xi$ is a subset of discretized state trajectories, where the $i$th row corresponds to the value of states that cross partitions at time $t_i$. $\Xi$ are coupling variables, so they are included in the set of optimization variables in this IDF formulation:

$$\min_{x, \xi, \Xi} \phi(\Xi, x_p, x_c)$$

subject to: \hspace{1cm} $g_p(\Xi, x_p) \leq 0$

$$\Xi - a(\Xi, x_p, x_c) = 0 \quad (18)$$

Here, $\Xi$ is determined by simulating each of the subsystems. Each simulation requires a priori specification of the state trajectories that are inputs to the corresponding subsystem (i.e., the corresponding elements of $\Xi$). The components of $\Xi$ are local copies of state trajectories that correspond to trajectories computed in other subsystem simulations. Having the optimization algorithm specify these local copies enables independent simulation of each subsystem. The relationship between these quantities is made more clear by the last constraint in Eq. (18). The analysis functions $a(\cdot)$ here output the state trajectories that cross subsystem boundaries as computed by the simulations. The analysis function outputs must match the local state trajectory copies $\Xi$. In other words, if we specify $\Xi$ and use it in evaluating $a(\cdot)$, the resulting state trajectories must match the input $\Xi$. If the last constraint is satisfied, $\Xi$ is a fixed point, and the solution...
to this decomposed problem will match the solution to the undecomposed MDF formulation.

An accurate solution using this specific IDF approach requires fine discretization of state trajectories \( \Xi \). This fine discretization, however, increases IDF formulation problem dimension since \( \Xi \) is an optimization variable, often resulting in computationally expensive solutions. In other high-dimension formulations (such as direct transcription), problem structure and easily obtained analytical derivatives can be exploited for efficient solution. Analytical derivatives are difficult to obtain for Eq. (18) due to the subsystem simulations. While IDF partitioning enables coarse-grained parallelism, its problem structure does not readily allow for further efficiency improvements.

High-dimension coupling variables (such as \( \Xi \)), also known as vector-valued coupling variables (VVCVs), can be approximated with low-dimension representations to aid efficient computation [186–188]. We can extend the usefulness of IDF and other MDO formulations for the design of dynamic systems using this and other workarounds, but “force-fit” approaches like this are fundamentally limited. There is a bound on how far trajectory representation dimension can be reduced before solution accuracy suffers. We need to explore fundamentally different solution methods that fit the properties of dynamic system design problems more naturally.

Cutting across signals is not the only decomposition available for dynamic systems. Problems may be partitioned temporally by splitting the simulation into \( n_T \) time segments instead of partitioning system model elements. The state trajectories across each one of these time segments is obtained via simulation. We can consider each of these independent simulations to be an analysis function, and the state of the system between time segments makes up the set of coupling variables. The coupling variables in this case are normally lower dimension than the coupling variables in the first IDF variant given in Eq. (18). This second IDF decomposition approach, known as multiple shooting [24,52], has practical motivations. It helps ameliorate numerical instabilities for highly nonlinear systems, and it enables coarse-grained parallel computing. The IDF formulation for this approach is:

\[
\begin{align*}
\min_{x_p, x_c, Y} \quad & \phi(\Xi, x_p, x_c) \\
\text{subject to:} \quad & g_p(\Xi, x_p) \leq 0 \\
& \xi_i(\Xi, Y) = 0 \quad i = 1, 2, \ldots, n_T \quad (19)
\end{align*}
\]

Here, \( Y \) is the matrix of coupling variables; each row corresponds to the state value at the beginning of a time segment. The defect constraints \( \xi_i(\cdot) \) ensure the initial state values for each time segment (corresponding to rows of \( Y \)) match the final state value from the previous time-segment simulation (corresponding to the appropriate rows of \( \Xi \)). These quantities are illustrated in Fig. 4. The optimization algorithm chooses initial state values for the simulations in time segments 1 and 2. These initial values correspond to rows in \( Y \). In time segment 1, the state trajectory is obtained by simulating through \( t_1 \). The defect constraint quantifies the difference between \( \xi_i \) (the
state value at $t_1$ obtained via simulation in time segment 1) and the initial state value used in time segment 2. At IDF convergence, these two quantities should match. While not widely used for co-design at present, this multiple-shooting approach is a well-known optimal control technique.

The third IDF approach involves a popular model for co-design problems that addresses the link between plant analysis and control design. Suppose the objective function takes the form $\phi(\xi(t), x_c, x_p)$ but the dependence of $\phi(\cdot)$ on $x_p$ is through intermediate variables $y$. For example, Allison et al. modeled a passive–active automotive suspension where the system dynamics model depends on spring stiffness and damping coefficients $k_s$ and $c$, respectively [24]. These coefficients, however, are not independent design variables. They are intermediate (coupling) variables that depend on other quantities that designers have direct control over, such as geometric dimensions. If $y = \{k_s, c\}$, we can represent this dependence using the analysis function notation, $y = a(x_p)$, and the IDF formulation becomes:

$$
\min_{x_p, x_c, y} \phi(a(\Xi, x_p), x_c)
$$
subject to:

$$
g_p(\Xi, x_p) \leq 0
$$
$$
y - a(x_p) = 0
$$

As with the other IDF approaches, $\Xi$ is obtained via simulation, but in this case, a single un-decomposed simulation is used. The coupling variables here are low dimension, and they offer yet another opportunity for problem decomposition. In this way, we can look at co-design as a two-discipline (plant design and control design) MDO problem [114, 166, 189].

Other decompositions are possible. Engineers may define subsystems that correspond to observers, distributed control systems, or multiple disciplines associated with plant design. Hybrids of the three IDF formulation variants discussed previously may also be implemented. For example, the second and third variants could be combined to create a formulation with separate plant and control analyses, and multiple shooting is used to perform the simulation in a distributed manner for each subsystem.

The IDF formulations involve a single optimization problem. Alternative “multilevel” MDO formulations employ multiple distributed optimization problems that each solve a piece of the system design problem [190, 191]. Multilevel formulations are especially useful when desirable sparsity patterns exist in both analysis couplings and design variable dependence structures [133]. One important multilevel formulation is augmented Lagrangian coordination (ALC) [130, 192], which is a nonhierarchical generalization of analytical target cascading [131].

The three dynamic system decompositions described previously for IDF formulation also apply to ALC. For example, the third IDF formulation variation has been demonstrated in several ALC studies where analysis functions compute parameters that are used in a dynamic simulation. Alexander et al. demonstrated how to use this approach to solve an electric vehicle design problem that involves function-valued coupling variables [186–189, 193]. An ALC subproblem is defined for electric motor design, and a second subproblem is defined for powetrain or vehicle system design. The vehicle design subproblem computes the system objective function based on dynamic simulation. This vehicle-level simulation depends on motor properties, such as the torque-speed curve, that are computed by the motor subproblem. These properties are coupling variables that link motor design to the system objective function. While these coupling variables are not time histories, they are function valued and need to be represented using VVCVs. Reduced-dimension representations for this specific problem have been investigated, some of which render solution via ALC practical.

Another way to extend the third IDF decomposition approach to ALC involves defining ALC subproblems for plant and control design but then capitalizing on existing optimal control theory to develop an optimize-then-discretize solution for the controls subproblem. Allison and Nazari demonstrated this approach using an electric circuit design problem [189]. As with the electric vehicle studies, the objective function in this case is linked to the plant-design via plant analysis functions and coupling variables. The suspension co-design problem discussed previously is another example of this decomposition approach, since the objective function is linked to plant-design variables via coupling variables: $y = \{k_s, c\} = a(x_p)$.

The details of the two ALC approaches discussed previously are available in the literature [186–189,193]. Another possible ALC formulation based on the second IDF decomposition (multiple shooting) is introduced here. Suppose the objective function is of the form presented in Eq. (5), but in discretized form:

$$
\phi(\Xi, x_c) = \psi(\xi_0, T_F) + \int_0^T L(\Xi, x_p, x_c) \, dt
$$

Here, we assume that numerical integration is performed using discretized state and control trajectories to compute the Lagrange cost. The final state value at $T_F$ is $\xi_n$. Observe that this function is additively separable if the problem is partitioned temporally into $n_T$ smaller time segments. If $T_j$ is the time at the end of time segment $j$ ($T_0 = 0$ and $T_{n_T} = T_F = T_j$), and if $\Xi^{(j)}$ and $x_c^{(j)}$ are the discretized state and control trajectories over time segment $j$, the objective function may be rewritten as:

$$
\phi(\Xi, x_c) = \psi(\xi_0, T_F) + \sum_{j=1}^{n_T} \int_{T_{j-1}}^{T_j} L(\Xi^{(j)}, x_p, x_c^{(j)}) \, dt
$$

The optimization problem now can be divided into $n_T$ subsystems, where the $j$th objective function is the $j$th term of the sum in Eq. (22), and the Mayer term is included in the objective function for subsystem $n_T$. If coordinated using ALC, the decomposed problem is equivalent to the original undecomposed problem. In this ALC approach, the states at time-segment interfaces are coupling variables, and defect constraints ensure that the states at these interfaces are consistent. More specifically, if $x_c^{(j)}$ is the state at the beginning of time segment $j$ (the coupling variable), it must match $\xi_c^{(j-1)}$, which is the state at the end of time segment $j - 1$ obtained via simulation. The coupling variable $x_c^{(j)}$ is an independent optimization variable in the $j$th ALC optimization subproblem, whereas $\xi_c^{(j-1)}$ is computed in subproblem $j - 1$ and held fixed in subproblem $j$. The formulation for ALC subproblem $j$ ($j \neq n_T$) is:

$$
\min_{x_p^{(j)}, x_c^{(j)}, \xi_c^{(j)}} \int_{T_{j-1}}^{T_j} L(\Xi^{(j)}, x_p^{(j)}, x_c^{(j)}) \, dt + \int_{T_{j-1}}^{T_j} \psi(\xi_c^{(j)}, x_p^{(j)}, x_c^{(j)}) \, dt
$$
subject to:

$$
g_p(\Xi^{(j)}, x_p^{(j)}) - \xi_c^{(j)} \leq 0
$$

Instead of posing the defect equations as equality constraints as in the multiple-shooting formulation of IDF given in Eq. (19), the defect
constraints are enforced using an augmented Lagrangian penalty function $\pi()$. A coordination algorithm guides all of the subproblems toward agreement so that, at ALC convergence, the defect constraints are satisfied within a given tolerance. Another primary difference between this ALC formulation and the corresponding IDF formulation is that optimization tasks are distributed in addition to analysis tasks. Also, $x^{(l)}$ is a local copy of the plant-design vector, and at ALC convergence, the copies from all subproblems must match [enforced with $\pi()$]. Here, plant and control variables are solved for simultaneously in the optimization subproblem. A variant of this ALC formulation is to adopt a nested approach similar to Eqs. (15) and (16). The outer loop of the ALC subproblem would solve for $x^{(l)}$ and $y^{(l)}$, whereas $x^{(l)}$ would be obtained in an inner optimization loop.

To summarize, MDO has been used extensively for solving dynamic system design problems. In most cases, however, a simple MDF formulation has been used to solve the simultaneous or nested co-design problems (or parts of a sequential design process), or the problem has been limited to MDA with single-discipline design (e.g., optimal control). Solving dynamic system design problems with MDO formulations that are more sophisticated than MDF formulation has proven to be challenging. Some success has been realized via distributed MDO methods, particularly when specialized optimization algorithms are employed or the unique structure of dynamic systems is exploited (e.g., IDF variant two: multiple shooting). However, some approaches for decomposing tightly integrated dynamic system models produce VVVCVs (e.g., IDF variant one) or are otherwise poorly suited for solving dynamic system design problems. The solution approach in these cases is not a good fit. Instead of attempting to “force-fit” a particular solution method onto a given problem, new MDO strategies should be developed and explored that are compatible with the unique demands of dynamic system design. Multidisciplinary design problems that involve complex system dynamics are fundamentally different from the static, pseudostatic, simplified dynamic problems that much of MDO development has been based on. Returning to the foundations of MDO and developing formulations specifically for dynamic systems will advance both MDO research and efforts to design increasingly complex dynamic systems. The remainder of this paper outlines promising directions for building up a more general theory for multidisciplinary dynamic system design optimization.

### IV. Intrinsically Dynamic MDO Formulations

The need is clear for MDO methodologies that are deeply compatible with the nature of dynamic system design problems, but how do we move forward? Optimization has been used very successfully for a number of specific dynamic system design applications, and in some cases, MDO has been applied; but how do we move toward a more general theory for MDSO? MDSO must extend to a wide array of dynamic engineering systems, address dynamic issues directly, and be used more comprehensively throughout the product development process. We envision MDSO as a vital branch of MDO and believe that MDSO should embody the following characteristics:

1) MDSO should be intrinsically dynamic. Most real engineering systems are dynamic. Many are nonlinear. System dynamics must be a core component of MDSO formulations, comparable to its importance in optimal control.

2) MDSO should be multidisciplinary and integrated. MDSO should incorporate both multidisciplinary analysis and multidisciplinary design. Integration should be central to MDSO, spanning analysis domains, time scales, and length scales. Decomposition should be strategic and congruent with dynamic system characteristics.

3) MDSO should be systems oriented. Legacy design mindsets should be replaced with a balanced approach to dynamic system design, including avoiding unnecessary multiobjective co-design formulations and bias toward control design. Additionally, larger systems of systems views should be incorporated into MDSO.

4) MDSO should use passive dynamics. A deeper treatment of physical-system dynamics in an integrated MDSO approach enables greater use of passive dynamics, reducing control-system demands and advancing system capabilities.

5) MDSO should be parallel. Computational resources increasingly rely on more processors to enhance performance rather than processor speed. It is imperative that algorithms used with MDSO are parallel in nature to exploit this trend.

In the remainder of this paper, we discuss four important fronts for advancing the state of MDSO and satisfying these five requirements.

#### A. Balanced Co-design

Most co-design studies have been performed with a strong emphasis on control design, and they tend to deemphasize physical-system design [92,107]. For example, while controls engineers often recognize the importance of integrating control design with plant design, they often construct co-design implementations with fairly simplified plant-design formulations. Dependent quantities often are treated as independent plant-design variables. For example, Fathy et al. treat spring and damper coefficients as design variables [107,163], when in reality, these quantities depend on detailed geometric design variables, such as spring wire and helix diameters. In other words, spring and damper coefficients are coupling variables y that link plant- and control-design analyses. Using coupling variables in place of design variables results in an incomplete problem formulation; its solution produces plant requirements instead of a plant design, and it usually neglects plant-design constraints $g_p()$ (e.g., suspension packaging, fatigue, damper temperature, etc. [24]). Plant-design simplification is particularly problematic when working with nested co-design formulations. The plant constraints, at least those that depend on state trajectories, should be included in both the inner and outer loops [Eqs. (15) and (16)], but simplified plant design obscures this requirement. Fathy et al. presented the nested co-design formulation sans $g_p()$ where the inner loop is solved using LQR (for linear systems) [106]. LQR, however, cannot incorporate plant constraints, so a more general inner-loop solution method, such as MDF or direct transcription [24,52], must be used when more realistic plant-design models are used in co-design.

Generations of engineers have developed mechanical systems without active control. Design paradigms appropriate for passive systems have evolved and matured, and they now permeate the collective engineering consciousness. These mindsets often are taken as given, and it is hard to imagine any other design perspective, even for engineers seeking multidisciplinary design solutions. These legacy design approaches are evident in many existing co-design formulations in two ways. First, many studies posit that co-design is by nature multiojective [110–112], i.e., the plant-design objective is distinct from the control-design objective (case 2, 3, or 5 plant design). Often the objective used for the plant is passive or static, a clear artifact of legacy design paradigms. These co-design formulations overlook to some degree the integrated nature of actively controlled systems. Active systems are single, unified systems, not two systems each with a distinct design objective. Plant and control systems work together toward a systemwide goal [194]. Co-design formulations therefore should include a single systemwide objective that reflects the primary purpose of the overall system (case 4 plant design) [149]. As noted earlier, co-design problems may indeed be multiojective due to inherent system tradeoffs (e.g., cost vs performance), but a problem is not automatically multiojective if it is a co-design problem. If fundamental tradeoffs exist, the set of multiple system objectives should be used consistently across both plant and control designs, qualifying as a case 4 plant design.

Second, co-design studies often assume unidirectional coupling between plant and control designs, i.e., control performance depends on plant design: $\phi(a(x_p), x_p, q(t))$, but not vice versa. This premise is understandable if a control system is viewed as an “add-on” to the physical-system due to legacy design mindsets. The properties of active dynamic systems, however, typically depend simultaneously on plant and control designs. If this is the case, unidirectional
formulations are incomplete. Many properly modeled plant-design constraints depend on dynamic response $\xi(t)$, which depends on both plant and control design. For example, material fatigue constraints depend on stress oscillation properties, which are a function of state trajectories (see the active suspension example in [24]). While bidirectional coupling is challenging to model, it is required for co-design formulations to accurately represent the system design problem. Any of the co-design formulations discussed previously that use a case 4 objective and include plant-design constraint dependence on $\xi(t)$ are bidirectional.

A balanced approach to co-design, such as the approach demonstrated in [24], enables engineers to construct a formulation based on what is best for the overall system, rather than retaining elements of legacy design formulations from plant or control design. These formulations appropriately balance plant- and control-design depth, have single system objective functions (or consistently applied sets of multiple objectives), and account for bidirectional coupling.

B. Passive Dynamics

Passive dynamics refers to the dynamic behavior of a system without active control. Many systems are designed to operate passively (e.g., most automotive suspension systems [168,195], vibration absorbers [158], and passive walkers [196]), and often this design approach is desirable to reduce system complexity and improve stability and reliability. Passive dynamics, however, play a critical role in active systems as well. The physical elements of a system should be designed so that their passive dynamic properties combine synergistically with active control to enhance performance [110,153,197,198]. Doing so can have a profound impact on dynamic performance and energy consumption. For example, extensive research in building design has resulted in numerous passive technologies for reducing energy consumption [199], such as night ventilation [200,201], passive cooling [202], solar walls [203], and passive systems combined with advanced control systems [204].

Most passive dynamics studies to date have employed a sequential design approach; the plant is designed first as a passive system (case 1 or 2 plant design), followed by control-system design. While this strategy may be effective, it cannot fully exploit the synergy between physical and control systems, and it cannot achieve system optimality. Co-design can be an effective approach for tailoring passive dynamics to enhance active system performance, but only if plant design is treated with sufficient depth and if the plant-design objective matches the system objective (case 4 plant design). Allison successfully demonstrated the co-design of a robotic manipulator where the control effort and energy required to perform a task were reduced dramatically by capitalizing on synergy between passive dynamics and active control [20,21]. Future MDSDO development should enhance our ability to leverage passive dynamics. In some cases, this may even eliminate the need for active control, or if not, significantly reduce control-system complexity and energy requirements.

C. Direct Transcription

Direct Transcription is a class of discretize-then-optimize optimal control methods that was introduced in Sec. II. Here, we explore additional details, discuss its extension to co-design, and examine its role in emerging MDSDO developments.

Exploring the relationship between DT (an AAO method) and IDF formulations for dynamic system design provides some useful insights. When using IDF formulation, some of the system analysis burden is shifted to the optimization algorithm via consistency constraints, whereas in AAO, the optimization algorithm performs all system analysis directly. In the case of IDF for dynamic systems (variant 2, or multiple shooting), simulations of subdivided time segments comprise the analysis functions, and consistency (defect) constraints ensure continuity between time segments. Now, consider what happens if the time-segment size is reduced to that of a single time step. The number of coupling variables would increase to $n_t \times n_t$ (the number states times the number of time steps), and a consistency constraint would be required for each time step. The coupling variables would then be the complete set of discretized states $x$, and the consistency constraints would be the discretized state equations.

When DT is applied to optimal control, the defect constraint Jacobian typically has a sparse diagonal structure that supports efficient problem solution. Wang and Arora have explored this property for a variety of DT formulations including trapezoidal, compressed Hermite–Simpson, separated Hermite–Simpson, Newmark’s method, central difference, piecewise Hermite interpolation, and cubic B-spline interpolation [65]. They calculated the number of nonzero elements in the Jacobian for each formulation with and without sparsity. Sparse formulations were only linear with respect to $n_t$, while full formulations were quadratic. It was also shown that increasing $n_t$ results in all formulations converging to the same solution. Finally, other qualitative advantages and disadvantages for the various formulations, such as smoothness and implementation difficulty, were discussed. Although their formulations did not include an active control element, the extension of the sparsity pattern for the discretized control was straightforward; the resulting pattern exhibits a sparse diagonal structure that is similar to the state variables [24].

The optimal control formulation for DT (without inequality constraints) was presented in Eq. (7). Allison et al. [24], Deshmukh and Allison [66], and Herber and Allison [67] demonstrated an extension of DT for the solution of co-design problems, and Tava and Suzuki employed a similar technique for launch vehicle co-design [205]. Using DT for co-design produces a problem that requires satisfaction of optimality conditions for both plant and control designs, in addition to satisfying defect constraints. The following is a simultaneous DT co-design formulation:

$$\min_{x_p, U} \sum_{i=1}^{n_t} L(x_p, u, \xi) h_i,$$

subject to: $\xi(x_p, U, \Xi) = 0$, $g_p(x_p, \Xi) \leq 0$ (24)

In the DT co-design extension, defect constraints are dependent on $x_p$, increasing constraint Jacobian density. In most practical co-design problems, the coupling between plant and co-design is bidirectional, i.e., plant constraints depend on state trajectories, which in turn depend on control design. The dependence of plant constraints on both state and plant-design variables increases the constraint Jacobian density further. Allison et al. has shown the Jacobian sparsity pattern for a realistic co-design problem [24]. While initial studies have addressed these challenges for specific co-design problems, many open questions remain regarding the extension of DT for co-design. One of these problems is a more computationally expensive time derivative function due to the plant-design models required. Recent work in derivative function surrogate modeling can help reduce DT expense in this case [206]. Note that other DT co-design formulations are possible, such as nesting a DT optimal control implementation within a plant optimization outer loop or using DT to solve an ALC controls subproblem as described in [189].

Computational efficiency, parallelism, and numerical stability are desirable properties of DT, but other qualities motivate on a more fundamental level the investigation of formulations like DT for dynamic system design. For example, Eq. (24) imposes no assumptions on control structure, which is especially helpful during early-stage design when control architecture is undefined. DT solutions provide insights into upper system performance limits without the restrictions imposed by specific control-system designs. Open-loop solutions also often provide insights into complex system dynamics and possible directions for physical-system design [90], and they can also serve as a basis for developing implementable feedback control systems. Most importantly, DT addresses system dynamics directly; dynamics are an integral part of the MDO formulation in Eq. (24). The DT co-design extension is a fully integrated
approach for dynamic system design that manages control, plant design, and state variables simultaneously.

While DT is promising for co-design, it can be challenging to implement at present. Commercial and open-source software is available for using DT to solve optimal control problems, but using DT for co-design changes the underlying problem structure and currently requires custom software development. Also, as a fully integrated AAO approach, DT usually does not mesh well with popular modeling environments. Sophisticated multidisciplinary analysis is difficult to incorporate, requiring integration at the equation level, a decidedly “advanced maneuver.” Progress must be made in theoretical and algorithmic development, the development of design tools, and awareness among design engineers before DT can become a practical solution for co-design.

D. Dynamic System Topology Optimization

Design optimization with respect to continuous variables, such as geometric dimensions or control parameters, is fairly mature and often can be performed with efficient gradient-based algorithms. In continuous optimization, however, system configuration is defined a priori. The design space is restricted, and in essence, we are “tuning” existing designs rather than generating completely new designs. Fundamentally new designs require configuration or topology modifications, i.e., changes in the existence of or interaction between system elements [207]. Topology design is traditionally the domain of engineering creativity and intuition, but often we lack the intuition required to make decisions regarding large-scale complex system that deviate very far from established design configurations (particularly if dynamics are important). The development of efficient methodologies for topology design is particularly important because new configurations can precipitate significant improvements in system performance, often much more so than continuous optimization alone. Success in topology optimization will help MDO break free of its “gilded cage” [208] and transition from design improvement to design synthesis.

Topology optimization methods for continuum systems, such as homogenization methods for structural topology optimization [172,209,210], are well established. These methods, however, do not apply to dynamic systems with discrete components with unique properties or functions (as opposed to a homogeneous continuum). For example, hybrid electric vehicle (HEV) powertrains combine multiple power sources to provide forward motion, and we have numerous options in how to specify the number, type, and connectivity of these power sources. Traditionally, engineers have explored design configurations via engineering intuition or by enumerating possible configurations [211] (sometimes aided by automated modeling [25]) and comparing the optimal designs of each. In either case, we are limited to investigating systems of only very small size, or limited to only partial enumerations of larger systems. For example, design of genetic regulatory circuits, a critical element of synthetic biology, has plateaued at a maximum circuit size of six nodes using scientific intuition or exhaustive enumeration [97,212,213]. Improved methods are required to develop circuits of practical size.

Traditional topology design problems are too large to use exhaustive enumeration as a solution approach. We can solve larger non-continuous problems using heuristic methods, although these approaches lead to improved instead of optimal solutions. For example, heuristic filters based on engineering knowledge can help reduce the number of design configurations that we need to compare. Liu [214] and Liu and Peng [215] demonstrated the use of heuristic filters in HEV powertrain design. Rule-based techniques can also be used to generate feasible system topologies and reduce the number of designs that must be compared [216]. Genetic algorithms and other heuristic algorithms have also been applied successfully for topology design [217], but they are largely limited to continuum design problems or small heterogeneous systems.

Many current topology optimization approaches use centralized decision-making, i.e., all configuration decisions are micromanaged by an optimization algorithm, and the available decisions are usually determined a priori. Centralization limits the scale of systems that can be considered. We need methods that can scale up to larger systems and that can address the specific needs of dynamic system configuration design problems. Theory and methods from complex systems research offer some promising directions for efficient large-scale system topology design. In complex systems, global behavior emerges from distributed local decisions (e.g., market systems, ant colonies, and immune systems). Keeping decisions local makes scaling up to very large systems possible. If engineers adopt design methods that involve automated local decision-making, they can concentrate their efforts on the rules that guide local decisions instead of on managing a large number of low-level decisions. In this way, a complex systems approach can help us abstract topology design problems and operate at a higher level to “transcend the overwhelming details of individual systems” ([218] p. 172). Novel topology design approaches based on complex systems will allow us to explore new design configurations without requiring decisions at the lowest level. Initial work in applying cellular automata to structural topology optimization [219,220] and in applying cellular division algorithms to dynamic system topology design problems has produced promising results [221–224], and it is an example of the type of complex systems technology that may enable topology optimization of large-scale dynamic systems.

Individual mechatronic systems often are part of a larger system of systems (SOS), where many mechatronic and other systems are coordinated to perform a larger task [194,225]. Transportation systems [226], space construction [227], military operations [194], and farming [1] are all examples of SOSs. The interface between physical and control systems is important to investigate, but the additional interfaces between individual mechatronic systems and issues surrounding distributed control design [228] lead to a particularly challenging topology design problem. SOS design should address simultaneously the interfaces at the individual system and SOS level.

SOS design is sometimes referred to as “site-level design,” where we are not only interested in the design of individual mechatronic systems but also in how they communicate and interact with each other, the environment, and humans. SOS design is too involved to perform from scratch whenever new needs evolve [226], often due to investment in existing infrastructure, complexity of associated social or economic systems, or due to sheer complexity [19]. When complete system design is impractical, strategic redesign of limited portions of the larger system is performed instead [18,20,21]; design methods that can accommodate uncertain future changes in system requirements are vital in SOS design.

V. Conclusions

The design of multidisciplinary dynamic systems presents unique challenges to engineers, and it is becoming an increasingly important technical issue as the number and complexity of smart and autonomous systems rises, and as their role in society becomes more crucial (e.g., energy and transportation systems). A phenomenal amount of work has been performed in the area of dynamic system design, but it has primarily addressed control-system design. Physical-system design is integral to the dynamic system design problem and must be addressed as well. Unfortunately, dynamic properties are often simplified or neglected when performing plant design. While control-design efforts more fully embrace system dynamics, if the plant-design problem is addressed in conjunction with control design, the plant-design problem is usually simplified (e.g., treating dependent variables as design variables). Co-design methods have been developed to design dynamic systems in a more integrated way, but they often exhibit a strong control-systems emphasis, and they sometimes retain many elements of siloed design methodologies (i.e., separate plant and control design). For example, distinct objectives for plant and control designs are often kept instead of adopting a systemwide objective. A more balanced system design approach is needed. While many have recognized the need for a more balanced and integrated approach to engineering system design,
many of the methods and tools required to put these concepts into practice are lacking.

A fresh systems perspective is needed to advance MDSO. It must be more than merging plant and control designs and constructing interface mechanisms between existing design frameworks. The underlying design philosophies for plant and control systems need fundamental changes; each needs to move from a disciplinary design approach toward a completely integrated approach focused on the system design problem. MDO is the right framework for this transition.

The problem of dynamic system design optimization with balanced consideration of physical- and control-system designs falls squarely into the domain of MDO. While MDO has been applied to the design of active dynamic systems using basic formulations such as MDF formulation, or within the limited scope of specific applications, the established MDO formulations largely do not explicitly address the unique characteristics of dynamic systems. A concerted effort is advocated in the MDO research community to develop more general theory and methodologies for MDSO. These efforts should result in MDO approaches with intrinsically dynamic formulations that provide a balanced approach to co-design, more fully use passive dynamics, embrace more sophisticated dynamic plant-design models, aid early-stage design efforts, and ultimately go beyond individual mechatronic systems to support the design of dynamic systems of systems.

References


doi:10.1080/0305215X.2011.646264


R. Haftka
Associate Editor