Engineering Design with Digital Thread

Victor Singh[*] and Karen E. Willcox[†]

Massachusetts Institute of Technology, Cambridge, MA, 02139

Digital Thread is a data-driven architecture that links together information generated from across the product lifecycle. Though Digital Thread is gaining traction as a digital communication framework to streamline design, manufacturing, and operational processes in order to more efficiently design, build and maintain engineering products, a principled mathematical formulation describing the manner in which Digital Thread can be used for critical design decisions remains absent. The contribution of this paper is to present such a formulation from the context of a data-driven design and decision problem under uncertainty. This formulation accounts for the fact that the design process is highly iterative and not all information is available at once. Output design decisions are made not only on what data to collect but also on the costs and benefits involved in experimentation and sensor instrumentation to collect that data. The mathematical formulation is illustrated through an example design of a structural fiber-steered composite component. In this example, the methodology highlights how different sequencing of small-scale experimentation with manufacturing and deployment lead to different designs and different associated costs.

Nomenclature

A_t^{kl}	Policy parametrization matrix coefficients for ply angle at stage t
$\mathcal{D}_t$	Digital Thread at stage t
$\mathcal{F}_t^m$	Failure index for failure mode m at stage t
I_t^{total}	Total complexity at stage t
$\mathcal{I}_t$	Information space over available resources at stage t
I_t^{fib}	Fiber complexity at stage t
I_t^{thick}	Thickness complexity at stage t
K	Number of spatial basis functions used per stage
L	Number of Digital Thread feature basis functions used per stage
M	Number of structural failure modes
$\mathcal{P}_t$	Probability space for uncertain input variables at stage t
$\mathcal{T}_t$	Space of all tools, methods, and processes at stage t
V_t	Value function at stage t
$\mathcal{V}_t$	Space of design and manufacturing parameters for all products at stage t
Z_t^{kl}	Policy parametrization matrix coefficients for thickness at stage t
$c_t^{\text{fib}}, c_t^{\text{thick}}$	Weighting coefficients for fiber and thickness complexity, respectively
d_t^l, d_t^a, d_t^m	Measurement data for strains, failure stresses, and manufacturing times at stage t, respectively
$\mathbb{E}[\cdot]$	Expectation operator
f_t^{fib}	Complexity feature for fiber angle at stage t
f_t^{thick}	Complexity feature for component thickness at stage t
g_t	Stage constraint at stage t
$p(\cdot)$	Probability distribution
$d_t \in \mathcal{Q}_t$	Measurement data collected from product lifecycle and associated space at time t
r_t	Stage cost at stage t
$u_t \in \mathcal{U}_t$	Decision variable at time index t and associated space
$u_t^d, u_t^a, u_t^z, u_t^s$	Decision variables for high level decisions, fiber-steering angle, component thickness, and sensor placement at stage t, respectively
v_t	Volume associated to $\mathcal{B}_t$
$\mathbf{x}_t \in \mathcal{B}_t$	Spatial coordinate on component body $\mathcal{B}_t$ at stage t
$\delta\mathcal{B}_t$	Boundary on component body $\mathcal{B}_t$ at

[*]Graduate student, Department of Aeronautics and Astronautics, victorsi@mit.edu, Student Member AIAA

[†]Professor of Aeronautics and Astronautics, kwillcox@mit.edu, Associate Fellow AIAA

	stage t
$y_t \in \mathcal{Y}_t$	Input variables and associated space at stage t
y_t^l, y_t^a, y_t^m	Input variables for loads, material allowables, and manufacturing model parameters at stage t, respectively
Φ_t	Digital Thread transition model at stage t
β	Volume penalty parameter
$\boldsymbol{\epsilon}$	Strain tensor (in Voigt notation)
ν_t	Measure associated to $\mathcal{Y}_t$
ϕ_t^l	l^{th} feature basis function for policy parametrization at stage t
μ_t	Policy function at stage t
π_t	Policy at stage t
ψ_t^k	k^{th} spatial basis function for policy parametrization at stage t
$\boldsymbol{\sigma}$	Stress tensor (in Voigt notation)

Subscripts

$t \in \mathbb{N}_0$	Non-dimensional time index or stage

Superscripts

$*$	Designation for optimal quantity or function

I. Introduction

Digital Thread introduces the idea of linking information generated from all stages of the product lifecycle (e.g., early concept, design, manufacturing, operation, post-life, and retirement) through a data-driven architecture of shared resources (e.g., sensor output, computational tools, methods, and processes) for real-time and long-term decision making [1, 2]. Furthermore, Digital Thread is envisioned to be the primary or "authoritative" data and communication platform for a company's products at any instance of time [2, 3]. It is important to distinguish the related concept of *Digital Twin* [2], which is a high-fidelity digital representation to closely mirror the life of a particular product and serial number (e.g., loading history, part replacements, damage, etc.). The Digital Twin can come in the form of a high-fidelity computational model or a combination of models and tools of sufficient fidelity to simulate the life history of the corresponding product. Digital Thread then can be viewed as containing all the information necessary to generate and provide updates to a Digital Twin.

Of particular relevance is the process in which Digital Thread can be used in the design of the next generation of products as illustrated in Figure 1. Here, multiple stages across the product lifecycle feed information into the Digital Thread. Such information can be used to make informed choices on future designs, as well as to reduce uncertainty in design parameters and process costs. Additionally, such information may uncover more efficient strategies for operation. Carrying out design decisions adds new information to the product lifecycle, changing the state of the Digital Thread. This whole process can be mathematically described with a data-driven design approach and decision problem under uncertainty. This paper formulates that mathematical description using the tools of Bayesian inference and decision theory.

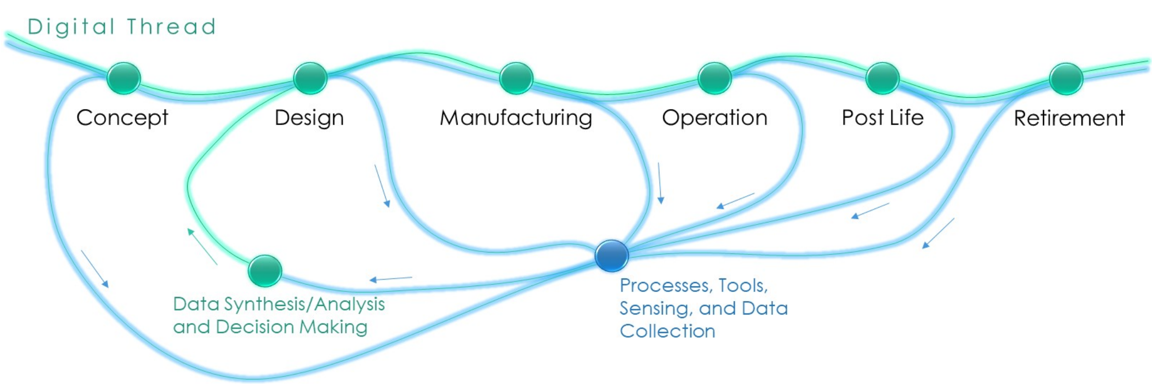

Figure 1. Illustration of engineering design with Digital Thread.

To add context to this discussion, consider the design of a structural component where the operational loads are uncertain. Practical approaches to deal with such uncertainty include over-conservative designs, damage tolerance policies, or redesigns and retrofits if analysis proves insufficient to ensure component integrity. Though the first generation of the component design may necessarily succumb to such practical approaches, data collected during the operational life of the first generation can be fed back and used to better design the next generation. For instance, components in operation can be equipped with strain sensors to infer the loading conditions. Strain data collected from these sensors can then be fed back to design with improved knowledge about the loading conditions for the next generation of component designs. Such a data-driven approach can improve the efficiency of designs over subsequent product generations of the component. To allow for such an approach, however, information and resources from different stages of the product lifecycle must be communicated back to design. This is achieved with Digital Thread.

Despite the range of data-driven technologies that currently exist across the product lifecycle, a unified treatment of data-driven design and decision making under uncertainty using Digital Thread remains absent in the literature. Prior work on Digital Thread has been targeted more so for enterprise level risk/value assessments [4], addressing upcoming challenges and future visions [5,6], and establishing requirements from the vantage point of model-based systems engineering [7,8], product lifecycle management, and additive manufacturing [9,10]. However, such a treatment must account for the fact that the design process is highly iterative and not all information is available at once. Design decisions must be made not only on what data to collect but also on the costs and benefits involved in experimentation and sensor instrumentation to collect that data. Furthermore, since full use of acquired information can become computationally prohibitive, it is critical to evaluate what minimal information is sufficient for design decisions and in what form that information is needed. In particular, the following questions remain unexplored: 1) How can Digital Thread be expressed mathematically and how can it be used for data-driven design? 2) What benefits can future designs gain from learning about previous designs through Digital Thread? 3) How are design decisions influenced by Digital Thread over the short term or long term? 4) Does there exist significant performance improvements in design with Digital Thread compared to traditional design? These are questions this paper aims to answer.

This remainder of this paper is organized as follows. Section II presents relevant background on design approaches currently used in practice and the state-of-the-art work on Digital Thread. Section III introduces the design problem of interest in this paper. Section IV details the underlying decision problem that we will solve. Section V demonstrates our approach on a particular setup of the design and decision problem. Section VI gives concluding remarks.

II. Background

To understand the benefits that design with Digital Thread can achieve, this section provides some background on current engineering design practices and state-of-the-art work on Digital Thread. The systems-level view of a product involves understanding the entire product lifecycle. The *product lifecycle* (in an engineering context) are all the stages from the initial concept, design, manufacturing, deployment, operation and post life services, through to the product's retirement/disposal. *Product lifecycle management* (PLM) [11] is the combination of strategies, methods, tools, and processes to manage and control all aspects of the product lifecycle over multiple products. These aspects might include integrating and communicating processes, data, and systems to various groups across the product lifecycle. A key enabler of efficient PLM has been the development and implementation of *Model-Based Engineering* (MBE) where data models or domain models communicate design intent rather than through document-based exchange of information [12,13], the latter in which can result in lossy transfer of the original source. Examples of MBE data models include use of mechanical/electronic computer aided design tools and modeling languages such as system modeling language (SysML), unified modeling language (UML) and extensible markup language (XML). *Model-Based Systems Engineering* applies the principles of MBE to support systems engineering requirements related to formalization of methods, tools, modeling languages, and best practices used in the design, analysis, communication, verification, and validation of large-scale and complex interdisciplinary systems throughout their lifecycles. [13–16].

PLM provides a means of understanding where uncertainties enter the overall product lifecycle and for incorporating this information back into design. The next step is designing under uncertainty itself. There are many fields relevant to the task of design under uncertainty, including uncertainty quantification, robust optimization, stochastic programming, optimal experimental design, multidisciplinary design optimization. Aspects of each of these fields find relevance in the decision problem associated to Digital Thread. For instance, the impacts of uncertainty from design parameters, modeling errors, and noisy measurements on the performance of specific products requires an understanding of how uncertainty propagates throughout the various stages of their lifecycles. This task can be accomplished with methods from *uncertainty quantification*, which explores identification, characterization, and ultimately reduction of uncertainty of a simulation or physical system [17]. In such systems, uncertainty quantification usually analyzes predicted outputs or specific quantifies of interest [17], often (but not always) representing uncertainty via a probabilistic model.

Design decisions associated to Digital Thread will be determined through minimizing some cost metric subject to constraints. Given the uncertainty just discussed, optimization methods to solve this problem will be inherently stochastic based. Additionally, there are different ways to treat the stochastic optimization problem. One way is through *robust optimization* where a stochastic optimization problem is cast into a deterministic one through determining the maximum/minimum bounds of the sources of uncertainty and performing an optimization over the range of these bounds [18]. Alternatively, *stochastic programming* treats the uncertainty with probabilistic models and optimization is performed on an objective statement (and possibly constraints) involving some mean, variance, or other probabilistic criteria [19]. *Uncertainty-based multidisciplinary design optimization* methods seek to include considerations of reliability and robustness in a system-level design formulation [20–22].

To be able to design over multiple product generations, design decisions must be made not just on current knowledge but also on potential future information. For instance, learning from operational data will require deciding if the current generation of products should be designed to help improve future collection of measurements (e.g., through optimally placed sensors or tailored structural architecture) or be designed only to satisfy immediate metrics of performance. These design decisions will have to be guided through some metric of assessing benefits and costs. A field that explores this problem is *optimal experimental design* (OED), where the objective is to determine experimental designs that are optimal with respect to some statistical criteria or utility function [23]. In our context, experimental designs refer to actual design decisions. Additionally, design decisions in the Digital Thread setting will be sequential in nature, where decisions of one generation will impact that of the next. This problem is explored in a sub-field of OED known as *sequential optimal experimental design* where experiments are conducted in sequence, and the results of one experiment may affect the design of subsequent experiments [24].

In order to create a digital realization of more complicated components and systems, however, the Digital Thread must be able to account for information from data sources across multiple processes and disciplines. In recent efforts, methodologies such as Single Digital Thread Approach to Detailed Design (STAnDD) [25] work towards integrating multidisciplinary processes and analysis disciplines to more effectively evaluate product costs and performances. Additive manufacturing provides an additional opportunity to use Digital Thread for integration of data-driven tools and physics-based simulations. In particular, use of the principles of Digital Thread in additive manufacturing has led to improvement in supply chain and production operations, reduction in time to design and manufacture parts, and decrease in manufactured part variability [9]. Alternatively, improvements in manufacturing technology have allowed for exploring other manufacturing strategies such as employing maintenance as part of a product's standard lifecycle to minimize overall costs [26].

A relatively recent opportunity that motivates Digital Thread is the concept of an "attritable" vehicle, where the prospect of manufacturing and deploying small unmanned aerial vehicles at much lower costs opens up an opportunity for design to incorporate information gained through operational data (structural strains, air measurements, flight paths, energy consumption, payload history, and other related quantities) collected from sensors on-board the vehicles. Some use cases of attritable vehicles have been explored with the Revolutionary Affordable Architecture Generation and Evaluation (RAAGE) [27] framework where design, operations, manufacturing, and costs are integrated to enable rapid exploration of different cost-reducing designs and strategies.

III. Design Problem

This section describes the design problem considered in this paper and defines the elements of the design problem along with their mathematical models.

III.A. Design Problem Description

We develop and illustrate our methodology in the context of a specific design problem, although the formulation is general and applies broadly across design problems. We consider a composite tow-steered (fiber-steered) planar component, where our objectives are to find the optimal fiber-steering and optimal component thickness subject to the design and constraint metrics. This component may be, for example, a small systems bracket on a larger assembly, a structural panel, or an aerodynamic control surface. In this paper, we explore the specific example of the design of a chord-wise rib within a wing box section.

A challenge to our design task is the presence of uncertain input variables that will directly influence the design of the product. In this problem specifically, the uncertain inputs are the loading the component will experience in operation, the material properties of the component, and the specific manufacturing timestamps. Situations where these variables have most relevance is during the early stage of design, when testing and experimentation have yet not taken place or when a brand new product is brought to market for which only partial design information can be used from other sources due to its novelty.

Large uncertainties in these inputs can lead to conservative designs that can be costly both to manufacture and operate. Thus, our goal is to collect data to reduce these uncertainties and thus to minimize costs. We can collect data through three different paths: Material properties can be learned through collection of measurements from coupon level experiments; manufacturing timestamps can be learned from a combination of a bill of materials, timestamps of individual processes, and other related documentation when a prototype or product is manufactured; and operating loads can be learned from strain sensors placed on the product in operation.

Although the task of learning the uncertain input variables through measurements can be addressed with methods from machine learning, this task in the context of the overall design problem is made complicated by the fact that collecting data comes at a cost. To see this, we illustrate the Digital Thread for this design problem in Figure 2. Here we see that collecting necessary data requires both time and financial resources. Though material data can be obtained fairly readily and quickly during the design phase, manufacturing data can only be obtained once a prototype is built. Additionally, operational data can only be obtained once a prototype or product is built, equipped with sensors, and put into operation. Depending on the scale of the component, the manufacturing process can take weeks or months and putting a product into operation with proper functionality of all its parts and sensors can take much longer.

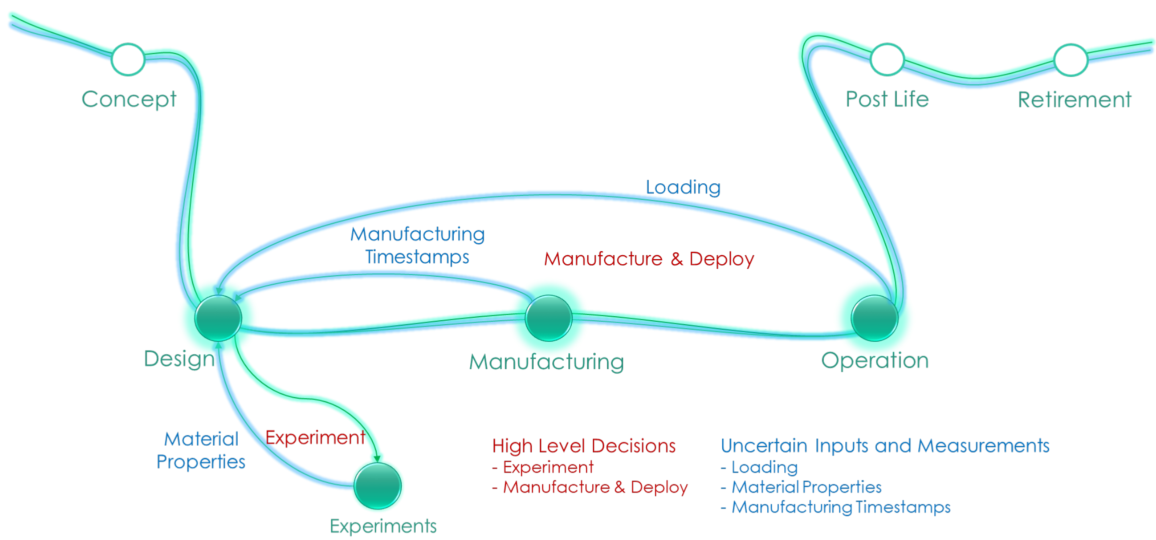

Figure 2. Illustration of Digital Thread for the design problem. The stages of the product lifecycle of focus here are the stages between design and operation.

With these considerations, we see that the overall design problem is underpinned by a decision problem under uncertainty, over potentially multiple product generations. Thus, our ultimate goal is to reduce input uncertainty to the degree necessary to minimize total expected costs over multiple product generations. We accomplish this goal using Digital Thread. In the following subsections, we set up the necessary elements of our design problem in Section III.B and the necessary modeling using Digital Thread in Section III.C.

III.B. Design Problem Elements

The key elements of the design problem can be broken down into four pieces: the uncertain input variables (what we would like to learn), the measurement data (what we learn from), the Digital Thread itself (how to represent what we know), and the decision variables (the decisions and design choices we can make). Figure 3 provides an illustration of these elements in the context of the example wingbox rib. The four elements are described in detail as follows:

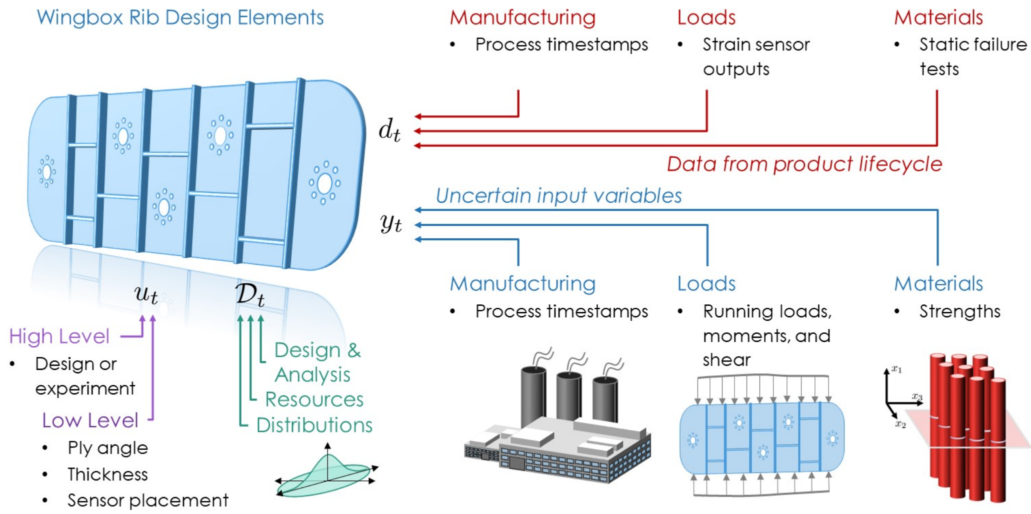

Figure 3. Illustration of the four key elements of the design problem for the wingbox rib example.

1. **Input Variables:** The input variables are the loads experienced in operation, material properties of the component, and the timestamps of each manufacturing process. Mathematically, we denote the input variables as $y_t \in \mathcal{Y}_t$ where $t \in \mathbb{N}_0$ is a non-dimensional time index and $\mathcal{Y}_t$ denotes the associated space. The time index t indicates the number of sequential design decisions made from some initial starting point where $t = 0$. The physical time duration between stage t and stage $t + 1$ is allowed to vary. We also give a time index on the space $\mathcal{Y}_t$ to allow for flexibility in scenarios where the vector of input variables can grow in size due to addition of new variables that were not considered previously. A detailed list of all input variables used for the design problem are provided in Table 1.

   We decompose further the input variables as $y_t = [y_t^l, y_t^a, y_t^m]^T$, where:

   (a) **Loads**: $y_t^l(\mathbf{x}_t) : \mathcal{B}_t \to \mathbb{R}^5$ are the input loads as a function of spatial coordinates $\mathbf{x}_t \in \delta\mathcal{B}_t$ on the component body $\mathcal{B}_t \subset \mathbb{R}^3$ with boundary $\delta\mathcal{B}_t \subset \mathcal{B}_t$ (running loads, running moments, and transverse shear).

   (b) **Material Properties**: $y_t^a \in \mathbb{R}^{12}$ are the properties of the material (elastic moduli, Poisson ratios, and material strength properties).

   (c) **Manufacturing Parameters**: $y_t^m \in \mathbb{R}^p$ is a p-dimensional parameter input of a model to calculate timestamps of a manufacturing process consisting of n steps. We use the manufacturing process and associated parameters detailed in the Advanced Composite Cost Estimating Manual (ACCEM) cost model [28, 29] for the design problem where $p = 57$ and $n = 63$.

   In our formulation, the input variables can be random and are represented probabilistically. Thus we require a random variable counterpart to model each uncertain input variable. These random variables also require definitions of the appropriate probability space and maps; for clarity of presentation to the engineering audience, we omit those details here.

2. **Measurement Data:** Measurement data are material properties from coupon level experiments, timestamps for each manufacturing process, and output from strain sensor measurements of the product in operation. These measurements are performed at different stages of the product lifecycle as illustrated in Figure 4. Mathematically, we denote the measurement data as $d_t \in \mathcal{Q}_t$ where $\mathcal{Q}_t$ denotes the associated space at stage t.

   As done for the input variables, we further decompose the measurements as $d_t = [d_t^l, d_t^a, d_t^m]^T$, where:

Table 1. Table of input variables and their descriptions.

y_t^l	Variable Description
1	Normal running load
2	Tangential running load
3	Normal running moment
4	Tangential running moment
5	Transverse shear load

y_t^a	Variable Description
1	Fiber compressive modulus
2	Fiber tensile modulus
3	In-plane normal compressive modulus
4	In-plane normal tensile modulus
5	Poisson ratio $1 \rightarrow 2$
6	Poisson ratio $2 \rightarrow 1$
7	Shear modulus
8	Fiber compressive strength
9	Fiber tensile strength
10	In-plane normal compressive strength
11	In-plane normal tensile strength
12	Shear strength

y_t^m	Variable Description
1	Clean lay-up tool surface
2	Apply release agent to surface
3	Position template and tape down
4	Ply deposition - automatic
5	Tape layer
6	Transfer from plate to stack
7	Transfer from stack to tool
8	Clean curing tool
9	Apply release agent to curing tool
10	Transfer lay-up to curing tool
11	Debulking (disposable bag)
12	Sharp male bend
13	Sharp female bend
14	Male radial
15	Female radial
16	Set up
17	Details, prefit, disassemble, clean
18	Apply adhesive

y_t^m	Variable Description
19	Assemble detail parts
20	Trim part
21	Apply porous separator film
22	Apply bleeder plies
23	Apply non-porous separator film
24	Apply vent cloth
25	Install vacuum fittings
26	Install thermocouples
27	Apply seal strips
28	Apply disposable bag
29	Seal edges
30	Connect vacuum lines, apply vacuum
31	Smooth down
32	Check seals
33	Disconnect vacuum lines
34	Check autoclave interior
35	Load lay-up tray
36	Roll tray in
37	Connect thermocouple
38	Connect vacuum lines, apply vacuum
39	Check bag, seal, and fittings
40	Close autoclave
41	Set recorders
42	Start cure cycle and check
43	Remove charts, open autoclave
44	Disconnect thermocouple leads
45	Disconnect vacuum lines
46	Roll tray out of autoclave
47	Remove lay-up from tray
48	Remove disposable bags
49	Remove thermocouples
50	Remove vacuum fittings
51	Remove vent cloth
52	Remove non-porous separator film
53	Remove bleeder plies
54	Remove porous separator film
55	Put used material aside
56	Remove layup
57	Clean tool

(a) **Strains**: $d_t^l \in \mathbb{R}^p$ with $p = 3q$ are the three strain components (in-plane strain components) for $q = 72$ sensor locations.

(b) **Materials**: $d_t^a \in \mathbb{R}^{12}$ are the material properties obtained from coupon level experiments. Coupon level experiments here involve the static failure of composite test specimens of appropriate loading and geometry in order to acquire data about material properties and failure strengths (y_t^a in Table 1) used for structural analysis.

(c) **Manufacturing Timestamps**: $d_t^m \in \mathbb{R}^n$ are the timestamps for each of the $n = 63$ steps of the manufacturing process described earlier.

Since collected data can be noisy and there exists uncertainty in methods, tools, and processes, the measurement data can be random and thus also require random variable counterparts as is done for the uncertain input variables.

3. **Digital Thread:** In the context of our design problem, the definition of Digital Thread should address the following key questions: 1) What is our current knowledge of the uncertain input variables of interest? 2) What resources do we have and how do we use them? 3) What products do we have and what information is available about them? 4) What is the least amount of information (no redundancy) we need to make efficient decisions? With this in mind, we can view Digital Thread as being a minimal digital realization of our state of knowledge and its associated uncertainty. We target a minimal realization through the use of sufficient statistics [30], where current knowledge of the uncertain input variables given past information will be represented by conditional probabilities. With this representation, we formalize the definition of Digital Thread through composition of three different spaces:

(a) **Statistics**: The first space, designated as $\mathcal{P}_t$ at stage t, defines the probability functions that characterize the uncertain input variables. An element of this space gives the probability of the uncertain input variables given past history (data and decisions) and the current decision yet to be taken. Note, we could have also defined this space in terms of a finite number of parameters if the distributions and their evolutions remain in some family of distributions (e.g., mean and variances for Gaussian distributions) or in terms of a collection of samples for more arbitrary distributions if necessary.

(b) **Products**: The second space, designated as $\mathcal{V}_t$ at stage t, is the space associated to design, manufacturing, and operating specifications for each product as well as their current conditions in the lifecycle. Specifically, this includes design geometry, manufacturing plans, operating protocols, measurement instructions, and operation/maintenance/repair history. An element of this space has a combination of numerical, categorical, and textual information.

(c) **Resources**: The final space, designated as $\mathcal{T}_t$ at stage t, formalizes our knowledge of the methods, tools, and processes available across the product lifecycle. This space contains information and protocols of methods, tools, processes, and algorithms for all stages across the product lifecycle. An element of this space also has a combination of numerical, categorical, and textual information.

With the three spaces defined, we formalize Digital Thread as follows. We define Digital Thread at stage t as $\mathcal{D}_t \in \mathcal{I}_t$ where $\mathcal{I}_t \subseteq \mathcal{P}_t \times \mathcal{T}_t \times \mathcal{V}_t$. This definition of Digital Thread, in addition to containing distributions on the uncertain input variables through $\mathcal{P}_t$, also provides information about existing products through $\mathcal{V}_t$ as well as information about all underlying tools, methods, and processes used for those products through $\mathcal{T}_t$. We also note that through the use of the spaces $\mathcal{P}_t$ and $\mathcal{V}_t$, we have all the necessary resources to build a Digital Twin corresponding to a particular product in operation.

4. **Decision Variables:** The decision variables (or control actions) permit us to make decisions about whether to perform experiments, or to manufacture and deploy a new design. Associated with manufacturing and deployment is additional specification of design parameters such as the fiber-steering angle, component thickness, and sensor location placement. We designate $u_t \in \mathcal{U}_t$ where $\mathcal{U}_t$ is the space of available decisions at stage t. More specifically, we decompose the decision variable as $u_t = [u_t^d, u_t^a, u_t^z, u_t^s]^T$, where:

(a) **High-Level Decision**: $u_t^d \in \{0, 1\}$ designates a binary decision between performing experiments ($u_t^d = 0$) and manufacturing and deploying a new design ($u_t^d = 1$).

(b) **Fiber Angle**: $u_t^a(\mathbf{x}_t) : \mathcal{B}_t \to [-\pi, \pi]$ is the fiber angle within the component as a function of spatial coordinates $\mathbf{x}_t$.

(c) **Thickness**: $u_t^z(\mathbf{x}_t) : \mathcal{B}_t \to \mathbb{R}_+$ is the component through thickness as a function of spatial coordinates $\mathbf{x}_t$.

(d) **Sensor Placement**: $u_t^s \in \mathbb{R}^{3q}$ are the sensor spatial coordinates for $q = 72$ sensors.

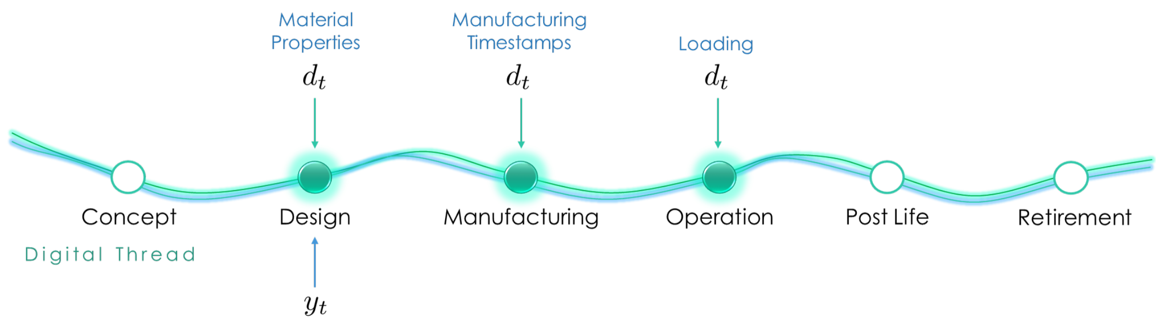

Figure 4. Illustration of flow of uncertain inputs and measurement data in the overall product lifecycle highlighted for the stages between design and operation.

III.C. Design Problem Modeling

With the four key elements defined, we can formulate the manner in which these variables influence each other and evolve over time. Specifically, we are interested in modeling the distribution associated to the uncertain input variables y_t and their evolution as parameters are learned from measurement data, the likelihood of the measurement data d_t, and the change in the state of the Digital Thread $\mathcal{D}_t$ when new data is collected and decisions are performed. These models will then ultimately layout a data assimilation approach in the context of Bayesian filtering and give the underlying mechanics for the decision problem. The models of interest are described as follows:

1. **Distribution of Uncertain Input Variables:** We model the statistics of y_t using a conditional probability model of the form $p(y_t|\mathcal{D}_t, u_t)$. Here the probability of the uncertain inputs depends on both the current state of the Digital Thread and the action taken.

2. **Likelihood of Measurement Data:** We model the statistics of d_t using a likelihood probability model of the form $p(d_t|y_t, \mathcal{D}_t, u_t)$. Here the likelihood of the measurement data depend on the uncertain input variables, the current state of the Digital Thread, and action taken.

3. **Digital Thread Evolution:** The Digital Thread changes from stage t to $t+1$ as new decisions are made and data is collected. We model this process as $\mathcal{D}_{t+1} = \Phi_t(\mathcal{D}_t, u_t, d_t)$ where $\Phi_t : \mathcal{I}_t \times \mathcal{U}_t \times \mathcal{Q}_t \to \mathcal{I}_{t+1}$ is the transition model at stage t.

With these models in place, the process of data assimilation can be described in the following way. For data assimilation, we are interested in the posterior distribution $p(y_{t+1}|\mathcal{D}_{t+1}, u_{t+1})$. Using the law of total probability and recognizing that y_t is not dependent on the future control u_{t+1} given $\mathcal{D}_{t+1}$:

$$p(y_{t+1}|\mathcal{D}_{t+1}, u_{t+1}) = \int_{\mathcal{Y}_t} p(y_{t+1}|y_t, \mathcal{D}_{t+1}, u_{t+1}) p(y_t|\mathcal{D}_{t+1}) \, \mathrm{d}\nu_t \tag{1}$$

This integral is over a product of two distributions. The first distribution represents any additional changes to the uncertain input variables in the next stage. This process occurs whenever information is inherited from a previous design and is modified for the new design. A relevant example is when loads on a new design, where the new design may be larger or more elongated than the previous design, are derived using information from loads of the previous design. The second distribution in the integral in (1) involves the actual data assimilation process. To see this, we use the fact that $\mathcal{D}_{t+1}$ is by design a sufficient statistic for past information (initial condition and measurement/control history) and Bayes rule to rewrite as:

$$\begin{aligned} p(y_t|\mathcal{D}_{t+1}) &= p(y_t|\mathcal{D}_0, u_0, ..., u_t, d_0, ..., d_t) \\ &= p(y_t|\mathcal{D}_t, u_t, d_t) \\ &= \frac{p(d_t|y_t, \mathcal{D}_t, u_t)p(y_t|\mathcal{D}_t, u_t)}{p(d_t|\mathcal{D}_t, u_t)} \end{aligned} \tag{2}$$

which is just a product of the prior and likelihood models we presented earlier (with the appropriate normalization). Note, correlation of the inputs is handled automatically through an appropriate prior distribution on the inputs and conditional distributions $p(y_t|\mathcal{D}_t, u_t)$ and $p(d_t|y_t, \mathcal{D}_t, u_t)$ used for the updates in (1).

We have now established a mathematical definition of Digital Thread and the underlying mechanics for updating the Digital Thread as new data are made available. Moreover, the data assimilation formulation shows a sequential nature in which the Digital Thread is updated where first a decision is made, followed by a realized input variable (the actual loads are experienced in operation, the actual manufacturing timestamps are available, or the coupons for testing are built with the actual material), followed by newly received data, and finally an update of the Digital Thread itself. Figure 5 illustrates this flow graphically. In Section IV, we develop the overall decision problem that incorporates these mechanics.

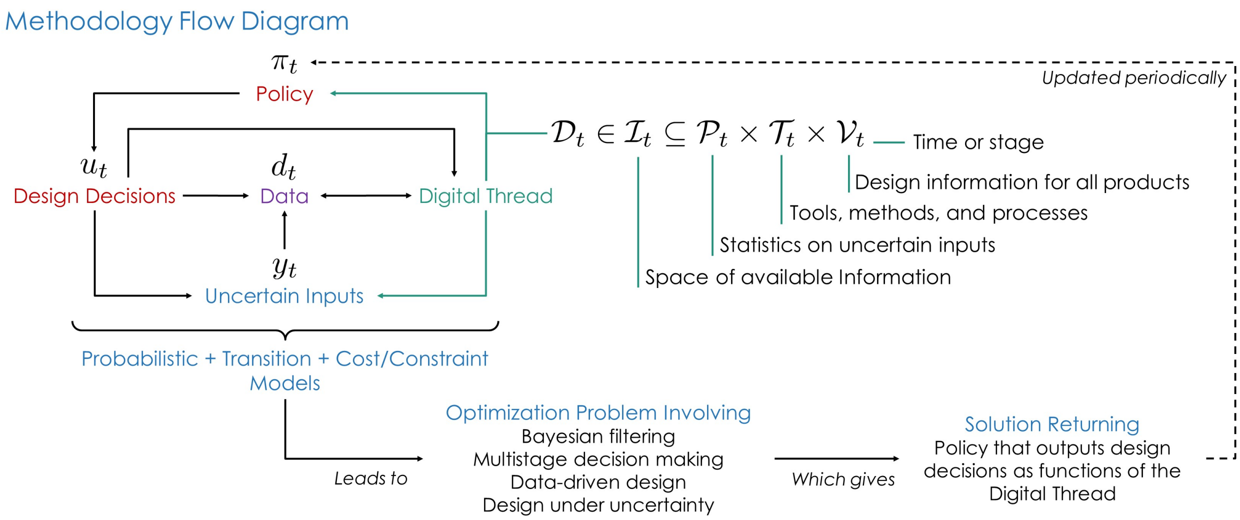

Figure 5. Flow diagram of complete methodology.

IV. Decision Problem

In the following subsections, we set up the necessary elements for the decision problem in Section IV.A and present the decision problem in Section IV.B.

IV.A. Decision Problem Elements

For a decision problem to be properly formulated, we need a means of steering the evolution of $\{\mathcal{D}_t, y_t\}$ using the decision variables u_t to satisfy desired metrics. This is accomplished through use of stage-wise cost and constraint functions, which are described below:

1. **Cost Function:** The cost function at stage t is given by a function $r_t(\mathcal{D}_t, u_t, y_t) : \mathcal{I}_t \times \mathcal{U}_t \times \mathcal{Y}_t \to \mathbb{R}$. For this problem setup, we approach cost from a geometrical point of view [31], which can serve as an alternative for monetary based cost metrics without explicit dependence on manufacturing processes. In particular, we view the cost of a composite component composed of key complexity features. For the

design problem we explore $N = 2$ features related to fiber angle variation and component thickness. The two features have the following forms:

(a) **Fiber Angle:** We express complexity of the fiber angle as the norm squared of the fiber angle gradient divided by the thickness. In particular, let a fiber direction (angle) at some spatial coordinate $\mathbf{x}_t$ with through thickness u_t^z be described by u_t^a. Then the feature associated with that fiber angle is given by

$$f_t^{\text{fib}}(\mathbf{x}_t, u_t) = \frac{1}{u_t^z(\mathbf{x}_t)} ||\nabla u_t^a(\mathbf{x}_t)||_2^2 \tag{3}$$

where $f_t^{\text{fib}}(\mathbf{x}_t, u_t) : \mathcal{B}_t \times \mathcal{U}_t \to \mathbb{R}_+$ is the feature associated to the fiber angle variation. The metric here penalizes small thicknesses coupled with aggressive fiber angle changes and favors larger thicknesses with benign fiber angle changes. In addition, the metric favors straight in-plane fiber paths. This metric is a small modification of the metric presented in Ref. 32.

(b) **Component Thickness:** We penalize two aspects of the component thickness: thickness variation across the component body and overall volume. The feature associated with component thickness u_t^z is given by

$$f_t^{\text{thick}}(\mathbf{x}_t, u_t) = ||\nabla u_t^z(\mathbf{x}_t)||_2^2 + \beta \tag{4}$$

where $f_t^{\text{thick}}(\mathbf{x}_t, u_t) : \mathcal{B}_t \times \mathcal{U}_t \to \mathbb{R}_+$ is the feature associated to the thickness variation and $\beta \in \mathbb{R}_+$ scales the penalty associated with total volume.

The complexity associated to each feature is given by an integration of that feature over the component body $\mathcal{B}_t$:

$$\begin{aligned} I_t^{\text{fib}}(u_t) &= \int_{\mathcal{B}_t} f_t^{\text{fib}}(\mathbf{x}_t, u_t)\, \mathrm{d}v_t \\ I_t^{\text{thick}}(u_t) &= \int_{\mathcal{B}_t} f_t^{\text{thick}}(\mathbf{x}_t, u_t)\, \mathrm{d}v_t \end{aligned} \tag{5}$$

The total complexity is given by an additive model of the form:

$$I_t^{\text{total}}(u_t) = c_t^{\text{fib}} I_t^{\text{fib}}(u_t) + c_t^{\text{thick}} I_t^{\text{thick}}(u_t) \tag{6}$$

for some weighting coefficients $c_t^{\text{fib}}, c_t^{\text{thick}} \in \mathbb{R}_+$. The cost function is then given as $r_t(\mathcal{D}_t, u_t, y_t) = I_t^{\text{total}}(u_t)$.

2. **Constraint Function:** The constraint function at stage t is given by a function $g_t(\mathcal{D}_t, u_t, y_t) : \mathcal{I}_t \times \mathcal{U}_t \times \mathcal{Y}_t \to \mathbb{R}$. The constraint function we use is based on structural failure criteria. An arbitrary failure criteria for structural analysis can be put in the form $\mathcal{F}_t^m(\boldsymbol{\sigma}, \boldsymbol{\epsilon}, y_t^a)$ where (in Voigt notation) $\boldsymbol{\sigma} \in \mathbb{R}^6$ denotes a stress tensor, $\boldsymbol{\epsilon} \in \mathbb{R}^6$ denotes a strain tensor, and m denotes the m^{th} mode of M total failure modes. This expression is local and denotes the failure index at spatial coordinate $\mathbf{x}_t$. Failure occurs whenever $\mathcal{F}_t^m(\cdot) > 1$. The constraint function is then given by a maximization over spatial coordinates on the component body and failure mode:

$$g_t(\mathcal{D}_t, u_t, y_t) = \max_{\mathbf{x}_t \in \mathcal{B}_t, m \in [1,\ldots,M]} \mathcal{F}_t^m(\boldsymbol{\sigma}(\mathbf{x}_t, u_t, y_t), \boldsymbol{\epsilon}(\mathbf{x}_t, u_t, y_t), y_t^a) - 1 \tag{7}$$

We investigate the Tsai-Wu criteria [33] as the failure metric. For the composite structural model, we use the small displacement Mindlin-Reissner plate formulation [34, 35] specialized for composites.

IV.B. Decision Problem Formulation

In setting up the decision problem, we would like to obtain answers to the following set of questions: Does one invest early in small scale experiments to reduce the uncertainties for a subset of the uncertain input variables? Or does one proceed to manufacturing and deployment to gain revenue through sales and gather other sources of data including specific manufacturing timestamps and operating conditions? When is one decision favored over the other and under what conditions?

With all the ingredients in hand and motivations discussed, we formulate the decision problem in the following way. We aim to minimize the expected total cost over a finite horizon (multiple product generations) subject to the immediate design constraints at stage t. Mathematically, this takes the form:

$$\begin{aligned} V_t^*(\mathcal{D}_t') = \min_{\pi_t} \mathbb{E}\Big[\sum_{k=t}^{T} \gamma^{k-t} r_k(\mathcal{D}_k, \mu_k, y_k) | \mathcal{D}_t = \mathcal{D}_t', \pi_t\Big] \\ \text{s.t. } \mathbb{E}\big[g_t(\mathcal{D}_t, \mu_t, y_t) | \mathcal{D}_t = \mathcal{D}_t', \pi_t\big] \leq 0, \quad t \in \{0, ..., T\} \end{aligned} \tag{8}$$

where a policy $\pi_t = \{\mu_t, ..., \mu_T\}$ defines a sequence of functions over the horizon T from time t. Each function μ_t returns the control action u_t, i.e. $\mu_t(\cdot) = u_t$. Here, $V_t^*(\mathcal{D}_t) : \mathcal{I}_t \to \mathbb{R}$ is the optimal value function or cost-to-go for Digital Thread $\mathcal{D}_t$ at stage t. The parameter $\gamma \in [0, 1)$ is a discount factor. The expectation is taken over the uncertain inputs $\{y_t, ..., y_T\}$ and measurement data $\{d_t, ..., d_T\}$ with respect to the joint probability distribution $p(\cdot|\mathcal{D}_t = \mathcal{D}_t', \pi_t)$. Given the recursive structure of the decision statement, the optimal cost can be expressed in terms of the following Bellman Equation using Bellman's principle of optimality [30]:

$$\begin{aligned} V_t^*(\mathcal{D}_t) = \min_{u_t \in \mathcal{U}_t} \mathbb{E}\big[r_t(\mathcal{D}_t, u_t, y_t) + \gamma V_{t+1}^*(\Phi_t(\mathcal{D}_t, u_t, d_t)) | \mathcal{D}_t, u_t\big] \\ \text{s.t. } \mathbb{E}\big[g_t(\mathcal{D}_t, u_t, y_t) | \mathcal{D}_t, u_t\big] \leq 0, \quad t \in \{0, ..., T\}, \quad V_{T+1}^* = 0 \end{aligned} \tag{9}$$

The solution to this dynamic program yields an optimal policy π_t^* that specifies new designs and changes to the Digital Thread for each step t up to the horizon T. Each function μ_t^* of the optimal policy is a function of $\mathcal{D}_t$ (a feedback policy), i.e. $\pi_t^* = \{\mu_t^*(\mathcal{D}_t), ..., \mu_T^*(\mathcal{D}_T)\}$.

A summary flow diagram of the complete methodology is given in Figure 5.

V. Demonstration

This section demonstrates our methodology on a specific setup of the decision problem introduced in Section IV.B.

V.A. Design Geometry

The component we analyze for design is shown in Figure 6. This component is a chord-wise rib from a wing box section for a small fixed wing aircraft with wingspan around 50 ft. The overall geometry has five holes of various radii with non-straight top and bottom edges.

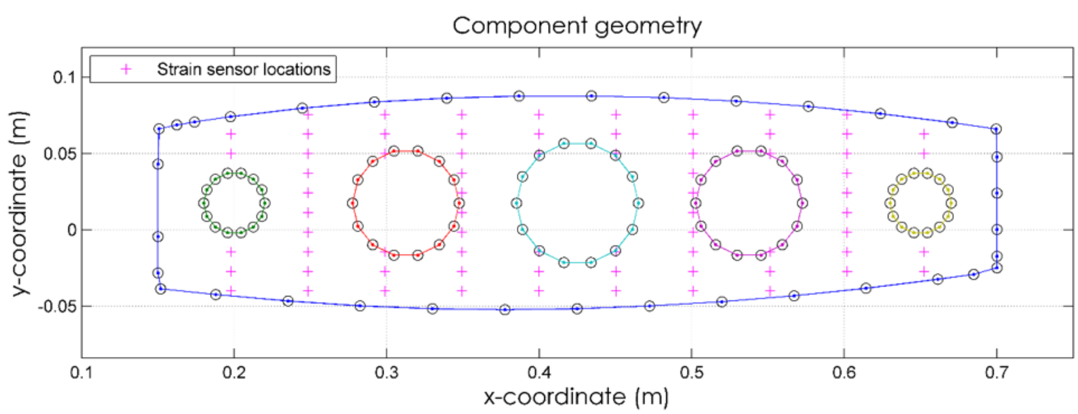

Figure 6. Illustration of component geometry. Sensor locations used for strain sensor measurements are displayed in magenta "+" marks

V.B. Decision Problem of Interest

For the demonstration, we focus on the greedy version of the decision problem, where the discount factor is set to $\gamma = 0$. In this case the decision problem statement takes the simpler form:

$$\begin{aligned} V_t^*(\mathcal{D}_t) = \min_{u_t \in \mathcal{U}_t} \, & \mathbb{E}\big[r_t(\mathcal{D}_t, u_t, y_t) | \mathcal{D}_t, u_t\big] \\ \text{s.t. } & \mathbb{E}\big[g_t(\mathcal{D}_t, u_t, y_t) | \mathcal{D}_t, u_t\big] \leq 0, \quad t \in \{0, ..., T\}, \quad V_{T+1}^* = 0 \end{aligned} \tag{10}$$

Note, that in this particular setup the strategy of sensor placement is not present due to the absence of V_{t+1}^*, and hence d_t, in the Bellman equation. That is, the policies that are generated are reactive to measurements as opposed to having selection of where to best place sensors for the next design. This is because having $\gamma = 0$ cancels out any contribution of future collected data in the optimization at each stage t. Non-zero discount factors will require solving the more general optimization problem given by (9).

To explore the behavior of design decisions on costs, we evaluate our formulation over three stages $t \in \{0, 1, 2\}$ where we are allowed one set of coupon level experiments labeled with E ($u_t^d = 0$), and two manufacturing and deployments of a new design labeled with D ($u_t^d = 1$). Enumerating these possibilities leads to two high-level decisions $\{EDD, DED\}$ that we can take (the option DDE is excluded because it does not make sense to experiment after the products are already manufactured and deployed). For example, the sequence DED means to manufacture and deploy a new design first, followed by performing coupon level experiments second, and the manufacturing and deploying another new design. The second design benefits from data collected from both coupon experiments and operational measurements of the previous design. The two high-level decisions $\{EDD, DED\}$ are distinguished by whether coupon experiments should be performed first or second. Letting each high-level decision correspond to a specific policy, we have two policies to compare. Our methodology will assess which policy generates lower costs for the example problem in addition to returning specific designs for each policy.

Note, enumerating and assessing high level decisions as separate policies (e.g., EDD and DED) will scale combinatorially based on the number of stages and high-level decisions available for each stage. However, in practice the number of high-level decisions of interest tend to be few for design-oriented objectives. In addition, the volatile nature of engineering design can bring upon new scenarios, objectives, and problem restructuring that can not all be accounted for practically in an optimization of a past stage. Thus, the number of stages to be evaluated will also be few. Reduction in computational time for a large number of polices can be achieved by updating and evaluating policies in parallel if necessary.

V.C. Setup of Design Problem Elements

For the demonstration, data and statistical parameters are generated synthetically using computational models and baseline loads information. The specific representation of the four elements are presented in the following:

1. **Uncertain Input Variables:** We model the uncertain input variables using Gaussian random variables with specified initial means and covariances. The small displacement Mindlin-Reissner plate model gives a linear mapping between loads and strain measurements, the map from material properties to measured material properties using coupon experiments is linear, and the map from manufacturing parameters to manufacturing timestamps is setup to be linear for this particular demonstration; thus, the uncertain input variables remains Gaussian over subsequent stages t.

2. **Measurement Data:** We model the measurement data using the aforementioned linear models that map input variables to measurement data with additive Gaussian noise. Strain sensor placement is assumed fixed for this demonstration.

3. **Digital Thread:** In this case, the Digital Thread comprises the distributions on the uncertain input variables as well as parameters and labels for the particular finite element solver and failure criteria, the noise parameters of the measurements, design geometry for components in operation, and cost/constraint parameters. The distributions on the uncertain input variables are calculated using Kalman Filters with the prediction and analyze steps reversed, as prescribed by the Bayesian filter in (1).

4. **Decision Variables:** The decision problem is solved using a policy parametrization technique. Since the ply angle and thickness are spatial in nature while the Digital Thread lives in information space, the policy parametrization will involve functions over both $\mathcal{B}_t$ and $\mathcal{I}_t$. In particular, the low-level control is determined using the following parametrization of the policy:

$$\begin{aligned}\mu_t^a(\mathbf{x}_t, \mathcal{D}_t) &= \sum_{k=1}^{K}\sum_{l=1}^{L}\psi_t^k(\mathbf{x}_t)A_t^{kl}\phi_t^l(\mathcal{D}_t)\\ \mu_t^z(\mathbf{x}_t, \mathcal{D}_t) &= \sum_{k=1}^{K}\sum_{l=1}^{L}\psi_t^k(\mathbf{x}_t)Z_t^{kl}\phi_t^l(\mathcal{D}_t)\\ \mu_t^s &= [\mathbf{x}_t^1, ..., \mathbf{x}_t^q]^T\end{aligned} \tag{11}$$

where $A_t^{kl}, Z_t^{kl} \in \mathbb{R}$ are the policy parametrization matrix coefficients for ply angle and thickness, respectively. Here, $\psi_t^k(\mathbf{x}_t) : \mathcal{B}_t \to \mathbb{R}$ is the k^{th} basis function over spatial points on the component body $\mathcal{B}_t$ and $\phi_t^l(\mathcal{D}_t) : \mathcal{I}_t \to \mathbb{R}$ is the l^{th} basis function over features of the Digital Thread $\mathcal{D}_t$ in information space $\mathcal{I}_t$. The features of the Digital Thread used here are the means and variances of the uncertain input variables extracted from the Digital Thread. The basis functions over spatial points on the body are expressed using Gaussian radial basis functions where the arguments are pre-scaled to lie on the unit hypercube. We use K basis functions centered at K points over spatial coordinates, where the centers of the radial basis functions correspond to the nodes of the finite element model. The basis functions over the features of the Digital Thread are expressed using L monomials.

   The parametrizations in (11) are then optimized using a policy gradient method [36, 37] applied to the decision statement (10). Expectations in (10) are approximated using Monte Carlo simulations. Monte Carlo errors are estimated numerically. We determined that approximately 100 samples were sufficient to ensure errors in the Monte Carlo approximations of the expectations in (10) were within 2%, based on numerically estimated statistics of $r_t(\cdot)$ for feasible designs.

Other modeling choices for the setup: At stage 0, the Digital Thread starts with large uncertainties in both the input loadings and material strength properties. The input loadings known at stage 0 have a mean that is 1.5 times larger than the mean of the input loadings the component will actually see in operation. The material strength properties known at stage 0 have means 0.9 times that of the actual material strength properties. The material moduli and Poisson ratios are assumed fixed and known. We set the operating costs to be 0.05 times the costs of manufacturing and deploying of the component. Data can be collected from manufacturing to operation only after a design has been manufactured and deployed from a previous time step. Costs to perform coupon level experiments are assumed negligible in comparison to other costs for this particular setup. The scalars in the cost model specified in (4) and (6) are set to $\beta = 0.001$, $c_t^{\text{fib}} = 0.2$, and $c_t^{\text{thick}} = 10^5$.

V.D. Results

In Figure 7, we show the typical effect of updating our knowledge of the uncertain loads by learning from the data collected from strain sensor measurements after deployment of a design. Here, the loads used for the initial design are compared to the loads estimated from operational data of the previously deployed design. These estimated loads are then used for the second design. The mean and two standard deviations of the variance for the initial estimate of loads (before any data assimilation) are shown with the red dashed-dotted line and red shading, respectively. Similarly, the mean and two standard deviations of the variance for the estimate of the loads after a design is deployed are shown with the blue dashed line and blue shading, respectively. The actual loads to be learned are shown with the thick magenta line. The large shifts in the mean and variance reduction of the running loads and normal running moment after data assimilation indicate that the design of the next generation can be built lighter (and therefore at a lower cost) than the previous generation. This is because the loads for this example setup are learned to be of lesser magnitude than what was used for the design of the previous generation. However, in order to have obtained this knowledge, we had to deploy a design in the first place—and that has a cost.

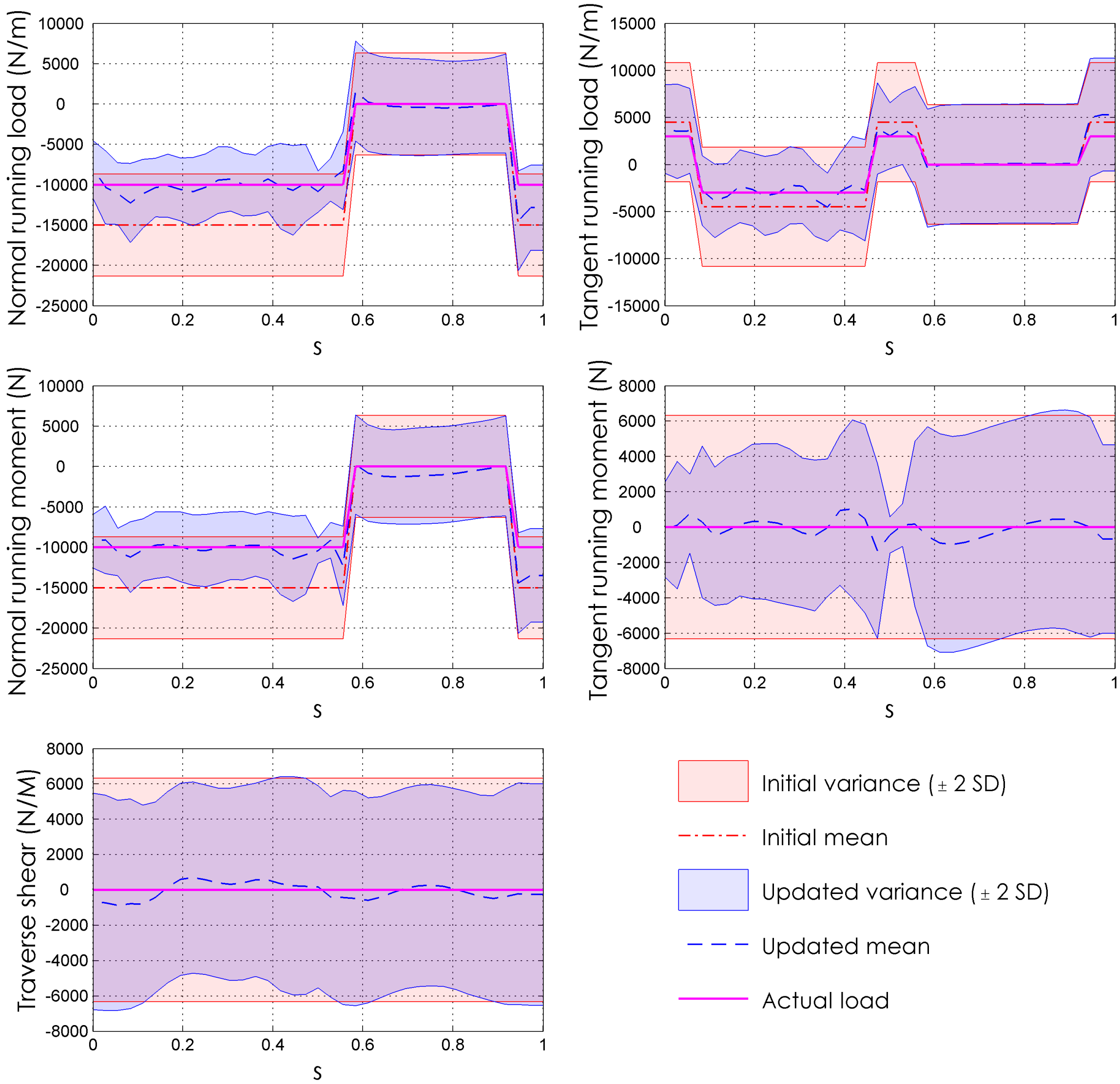

Figure 7. Demonstration of uncertain loads learned through strain measurements after manufacturing and deploying a design. Loads are presented here as functions of a non-dimensional variable $s \in [0, 1]$ that runs counter-clockwise around the outer perimeter of the component starting at spatial coordinates $(x, y) \approx (0.700, 0.025)$ m. SD - standard deviation.

In Figure 8, we show the typical effect of updating our knowledge of the uncertain material strength properties by learning from performed coupon level experiments. Here the estimates of material strength properties known initially are compared to the estimates after learning from coupon level experiments. The probability density function for the initial estimate of strength properties is given by the red shaded curve. Similarly, the probability density function for the strength properties after performing coupon level experiments is given by the blue shaded curve. The actual strength properties to be learned are shown with the thick magenta vertical line. The large shifts in the mean and variance reduction of the strength properties after performing the coupon experiments indicate that the next design to be deployed will benefit from higher and more precise material strength property estimates and therefore be lighter and of lower cost. Of course, to have obtained this knowledge, we had to forgo deploying a design earlier and the potential benefits it could have provided.

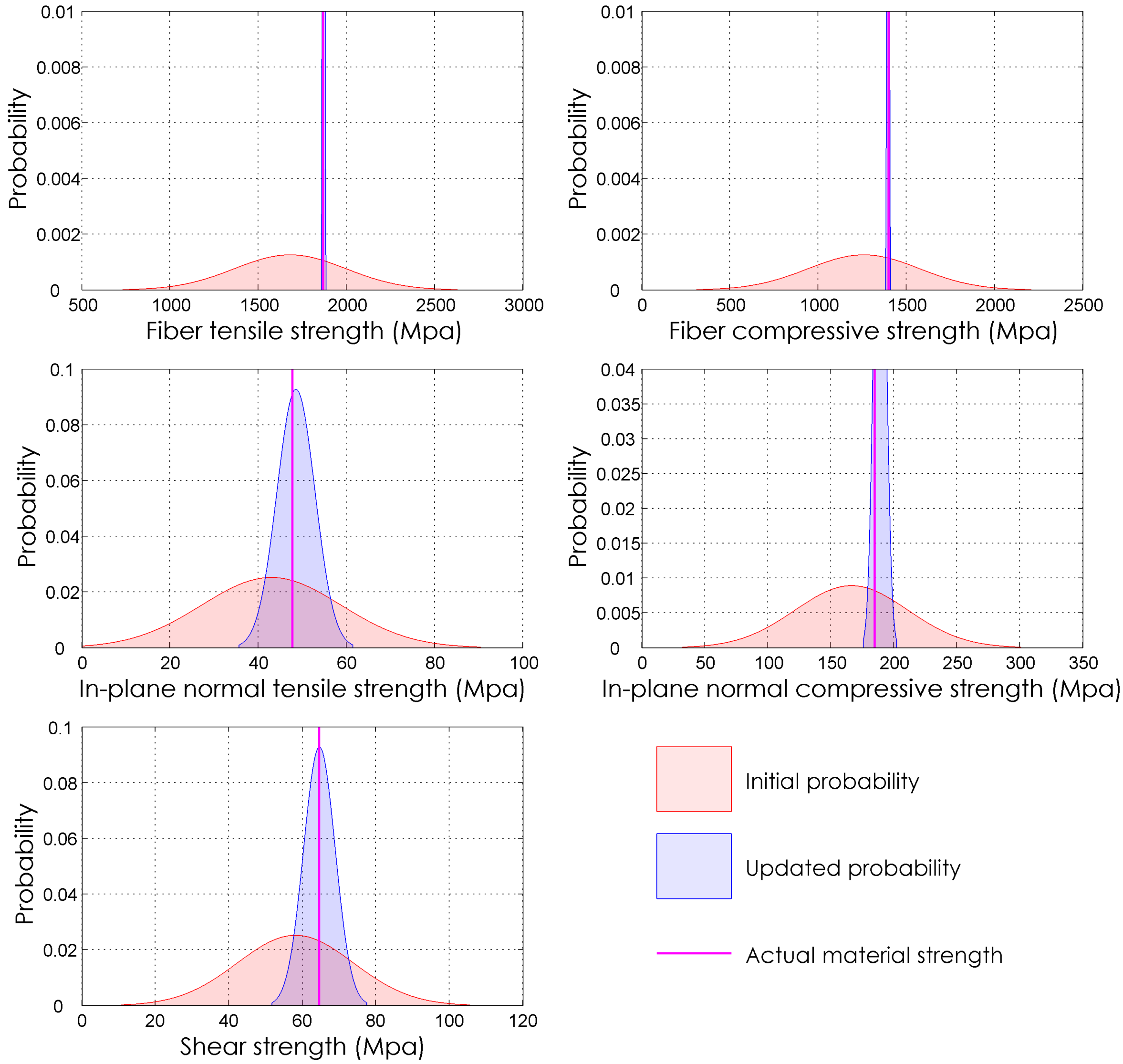

Figure 8. Demonstration of material strength allowables learned through coupon level experiments.

Typical first and second designs generated using policy EDD are shown in Figure 9 and Figure 10, respectively. Similarly, the first and second designs generated using policy DED are shown in Figure 11 and Figure 12, respectively. What can be observed is that both policies tend to favor fiber-steering for the first design (most noticeably near the top surface of the component) and thickness reduction for the second design. However, the first design produced by policy EDD benefits from the improved estimates of the material strength properties collected from the coupon level experiments while the first design for policy DED can only rely on initial estimates of those properties, which are conservative for this setup. As a result, the first design generated by policy EDD is thinner overall compared to the design generated by policy DED.

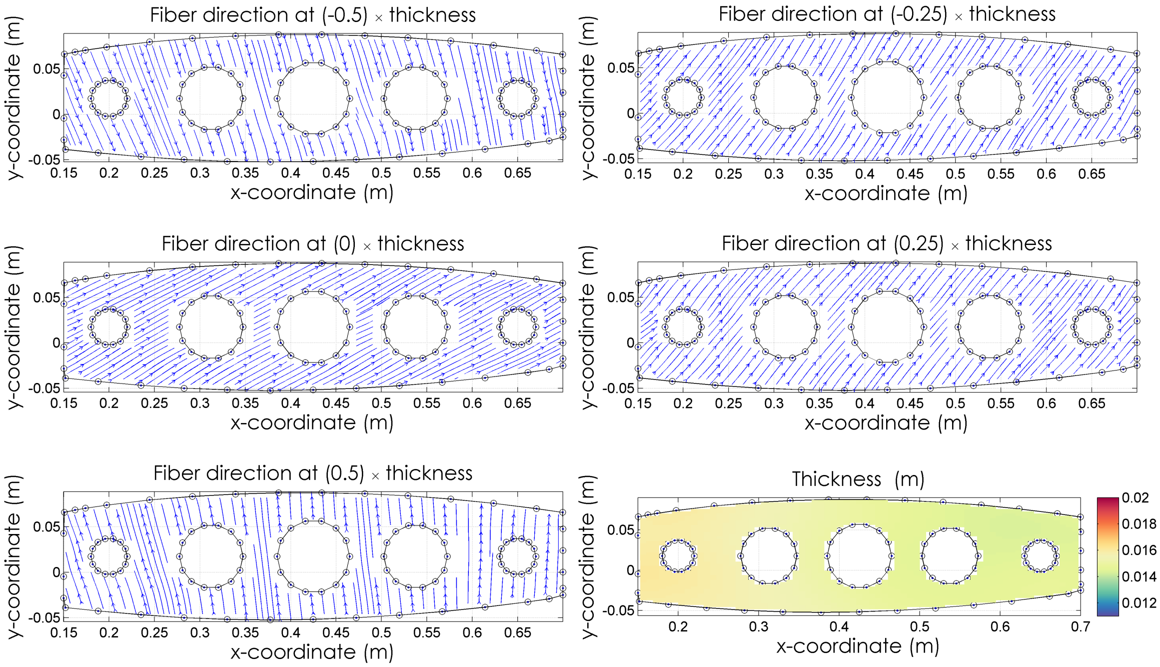

Figure 9. First design generated from policy EDD. Ply direction is shown as function of scaled depth through the thickness, where -0.5 corresponds to the bottom surface, 0 to the mid-plane, and 0.5 for the top surface.

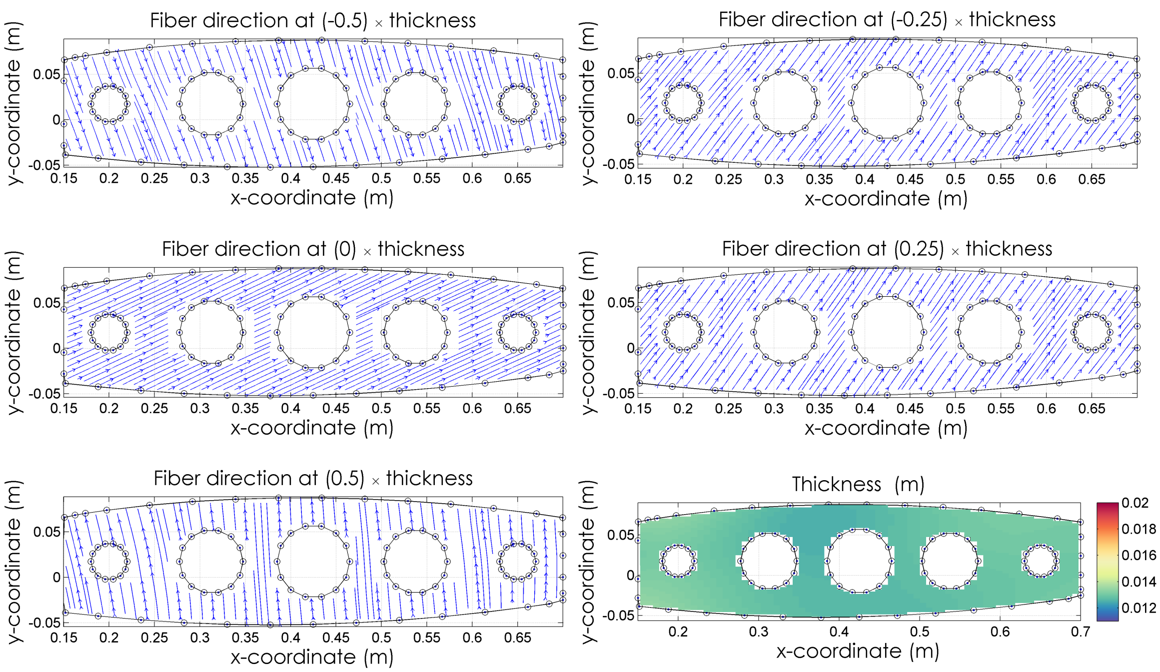

Figure 10. Second design generated from policy EDD. Ply direction is shown as function of scaled depth through the thickness, where -0.5 corresponds to the bottom surface, 0 to the mid-plane, and 0.5 for the top surface.

Though the second design produced from both policies benefit from both operational and coupon level

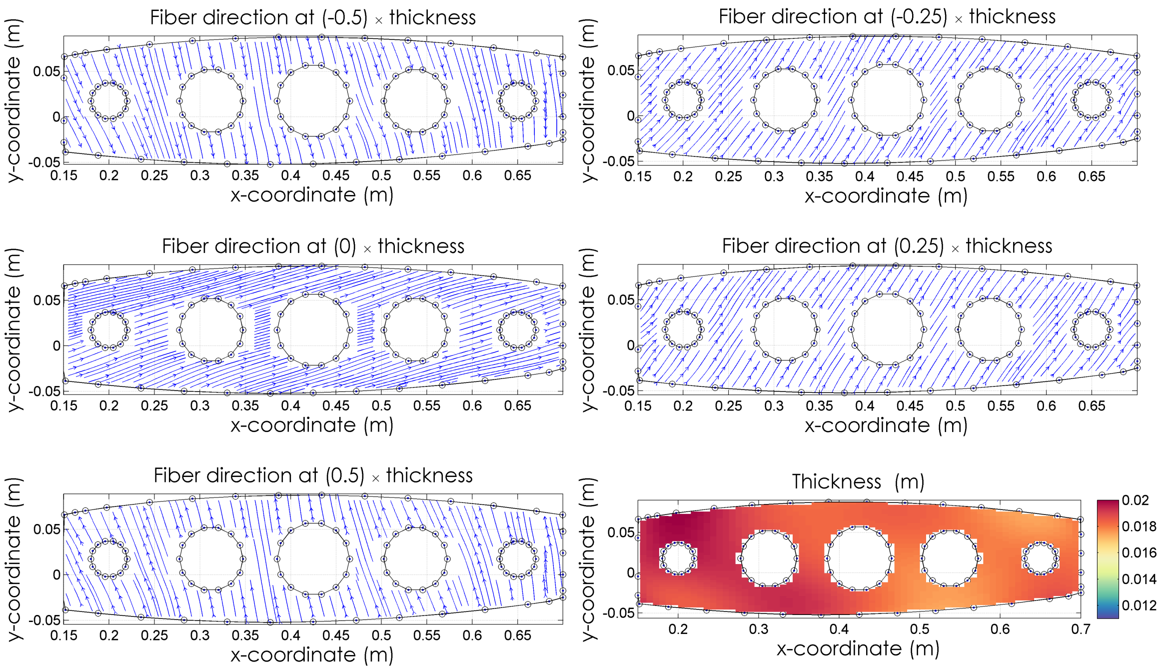

Figure 11. First design generated from policy DED. Ply direction is shown as function of scaled depth through the thickness, where -0.5 corresponds to the bottom surface, 0 to the mid-plane, and 0.5 for the top surface.

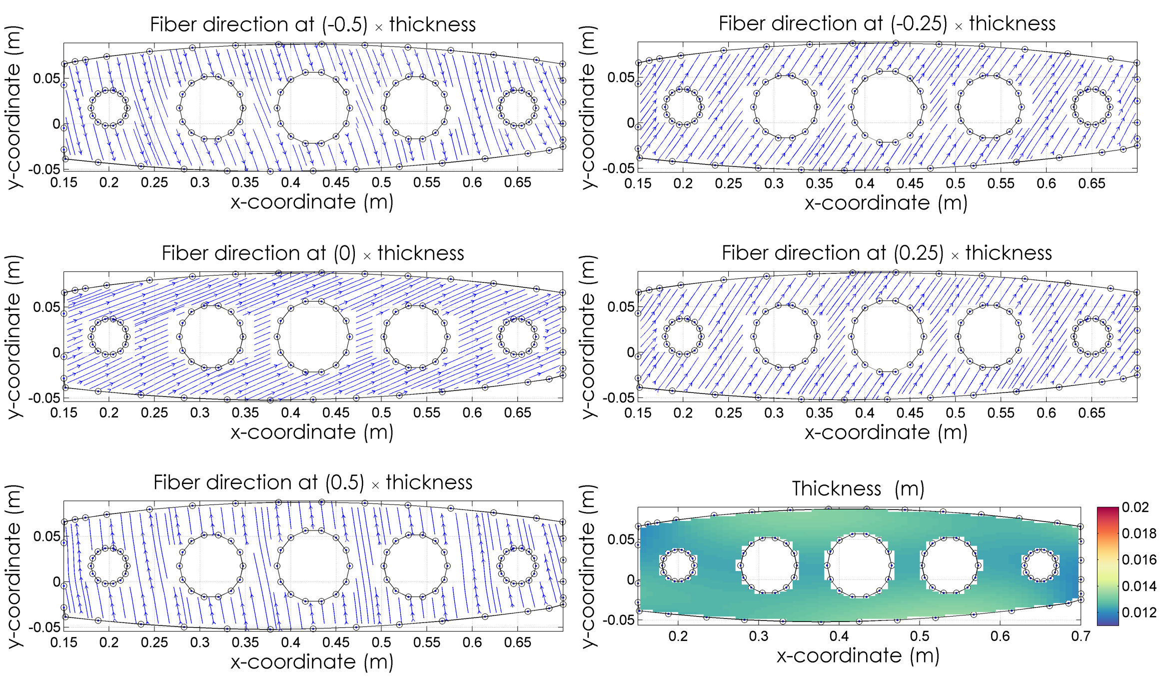

Figure 12. Second design generated from policy DED. Ply direction is shown as function of scaled depth through the thickness, where -0.5 corresponds to the bottom surface, 0 to the mid-plane, and 0.5 for the top surface.

experimental data, the costs for each policy are accumulated differently as can be seen in Figure 13. Here, we see that the best policy for this setup is EDD—that is, the best strategy from the given initial state of the Digital Thread is to first perform experiments to drive down the uncertainty of the material strength properties, and second to manufacture and deploy that design to learn about the uncertain loading conditions from data collected through operation. Interestingly, we see that manufacturing and deploying first leads to higher overall costs as a result of designing heavier and more conservative designs from the lack of data about the material strength properties earlier. Additionally, the corresponding operational costs are higher and accrued over a longer timeframe. These differences are more significant when the operational costs are higher than the 5% used here. The results overall show that material strength properties have a larger impact on the overall costs than do the input loads. This observation is made more surprising by the fact that the means of the material strength properties were only 10% away from the true values compared to 50% for the input loadings. It is important to note that the designs generated by both policies take into account the variances of the uncertain input variables. For this setup, these variances are relatively high with respect to the mean, so the designs generated by both policies reflect robustness to a wide range of possible uncertain inputs. Reducing input variance can lead to more specialized designs (more specific tailoring of fiber angles and thickness) through cost savings obtained by excluding uncertain inputs that are less likely to occur. Nevertheless, the best policy shows the importance of both the sources of uncertainty and the sequence in which one attempts to reduce them.

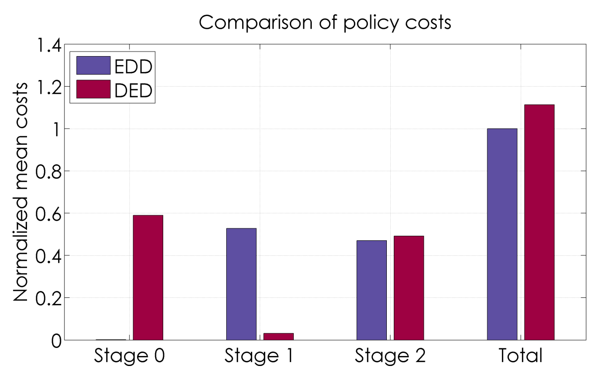

Figure 13. Stage cost comparison of the two policies. Costs are normalized with respect to the best policy.

VI. Conclusions

This paper presented a general mathematical methodology for modeling Digital Thread, for updating it with new information, and for using it to drive design decisions over multiple product generations. Furthermore, this paper presented terminology on how to view the key aspects of the design problem and introduced models that provide the mechanics of the associated data assimilation and decision problem. The associated decision problem simultaneously addresses multistage decision making for data-driven iterative design and design under uncertainty within its formulation. An illustrative composite component design example highlighted the key properties of the approach. Demonstrations show the intricate nature of the underlying data assimilation and associated decision problem, where tradeoffs in designing the current generation of products have to be made based not just on immediate costs but with consideration of their impacts on future designs of the product. In addition, the benefit of acquiring data must be balanced with the cost of or the opportunities lost in obtaining that data. This tradeoff assessment is performed automatically using the decision problem methodology and is shown to result in different designs and costs that reflect the different sequence in which information is acquired. In addition, this methodology can provide quantitative guidelines on which design inputs to target to efficiently reduce overall system uncertainties. Future work will involve demonstration of the method for decision problem setups that go beyond the greedy case and include additional decision complexity such as sensor placement or design tailoring for future data collection. In addition, the methodology will be further evaluated with sensor measurements from physical experimental tests of a structural component.

VII. Acknowledgements

This work is supported in part by AFOSR grant FA9550-16-1-0108 under the Dynamic Data Driven Application System Program (Program Manager Dr. E. Blasch), the MIT-SUTD International Design Center, and the United States Department of Energy Office of Advanced Scientific Computing Research (ASCR) grants DE-FG02-08ER2585 and DE-SC0009297, as part of the DiaMonD Multifaceted Mathematics Integrated Capability Center (program manager Dr. S. Lee).

References

[1] US Airforce, *Global Horizons Final Report: United States Air Force Global Science and Technology Vision - AF/ST TR 13-01*, United States Air Force, 2013.

[2] Kraft, E., "HPCMP CREATE-AV and the Air Force Digital Thread," *in AIAA SciTech 2015, 53th AIAA Aerospace Sciences Meeting, Kissimmee, Florida*, 2015, pp. 1–13, doi:10.2514/6.2015-0042.

[3] West, T. and Pyster, A., "Untangling the Digital Thread: The Challenge and Promise of Model-Based Engineering in Defense Acquisition," *INSIGHT*, Vol. 18, No. 2, 2015, pp. 45–55, doi:10.1002/inst.12022.

[4] Burgaud, F., Durand, J., and Mavris, D., "A Decision-Support Methodology to Make Enterprise-Level Risk/Value Trade-Offs," *in AIAA SciTech 2017, 19th AIAA Non-Deterministic Approaches Conference, Grapevine, Texas*, 2017, pp. 1–14, doi:10.2514/6.2017-0589.

[5] Kobryn, P., Tuegel, E., Zweber, J., and Kolonay, R., "Digital Thread and Twin for Systems Engeering: EMD to Disposal," *in AIAA SciTech 2017, 55th AIAA Aerospace Sciences Meeting, Grapevine, Texas*, 2017, pp. 1–13, doi:10.2514/6.2017-0876.

[6] Zweber, J., Kolonay, R., Kobryn, P., and Tuegel, E., "Digital Thread and Twin for Systems Engeering: Pre-MDD through TMRR," *in AIAA SciTech 2017, 55th AIAA Aerospace Sciences Meeting, Grapevine, Texas*, 2017, pp. 1–26, doi:10.2514/6.2017-0875.

[7] Hedberg, T., Lubell, J., Fischer, L., Maggiano, L., and Feeney, A., "Testing the Digital Thread in Support of Model-Based Manufacturing and Inspection," *Journal of Computing and Information Science in Engineering*, Vol. 16, No. 2, 2016, doi:10.1115/1.4032697.

[8] Sundaram, V. and Brownlow, L., "MBSE Digital System Model for AF DCGS," *in AIAA SciTech 2018, 2018 AIAA Aerospace Sciences Meeting, Kissimmee, Florida*, 2018, pp. 1–17, doi:10.2514/6.2018-1217.

[9] Mies, D., Marsden, W., and Warde, S., "Overview of Additive Manufacturing Informatics: "A Digital Thread"," *Integrating Materials and Manufacturing Innovation*, Vol. 5, No. 1, 2016, pp. 1–6, doi:10.1186/s40192-016-0050-7.

[10] Nassar, A. and Reutzel, E., "A proposed digital thread for additive manufacturing," *in International Solid Freeform Fabrication Symposium, Austin, Texas*, 2013.

[11] Stark, J., *Product Lifecycle Management*, Springer International, 2015, doi:10.1007/978-3-319-17440-2.

[12] Wymore, A., *Model-Based Systems Engineering*, CRC press, 1993.

[13] Ramos, A., Ferreira, J., and Barcelo, J., "Model-Based Systems Engineering: An Emerging Approach for Modern Systems," *IEEE Transactions on Systems,Man, and Cybernetics - Part C: Applications and Reviews*, 2012, pp. 101–111, doi:10.1109/TSMCC.2011.2106495.

[14] Estefan, J., "MBSE Methodology Survey," *INSIGHT-INCOSE J.*, Vol. 12, No. 4, 2009, pp. 16–18, doi:10.1002/inst.200912416.

[15] Cloutier, R., "Introduction to this Special Edition on Model-Based Systems Engineering," *INSIGHT-INCOSE J.*, Vol. 12, No. 4, 2009, pp. 7–8, doi:10.1002/inst.20091247.

[16] Holland, T., *Modeling and Simulation in the Systems Engineering Life Cycle: Core Concepts and Accompanying Lectures*, Springer-Verlag London, 2015, doi:10.1007/978-1-4471-5634-5.

[17] Smith, R., *Uncertainty Quantification: Theory, Implementation, and Applications*, SIAM Computational Science & Engineering, 2014.

[18] Bertsimas, D., Brown, D., and Caramanis, C., "Theory and Applications or Robust Optimization," *SIAM Review*, Vol. 53, No. 3, 2011, pp. 464–501, doi:10.1137/080734510.

[19] Kall, P., Wallace, S., and Kall, P., *Stochastic Programming*, John Wiley & Sons, Chichester, 1994.

[20] Yao, W., Chen, X., Luo, W., van Tooren, M., and Guo, J., "Review of uncertainty-based multidisciplinary design optimization methods for aerospace vehicles," *Progress in Aerospace Science*, Vol. 47, No. 6, 2011, pp. 450–479, doi:10.1016/j.paerosci.2011.05.001.

[21] Du, X. and Chen, W., "Efficient Uncertainty Analysis Methods for Multidisciplinary Robust Design," *AIAA Journal*, Vol. 40, No. 3, 2002, pp. 545–552, doi:10.2514/2.1681.

[22] Kokkolaras, M., Mourelatos, Z., and Papalambros, P., "Design Optimization of Hierarchically Decomposed Multilevel Systems Under Uncertainty," *ASME International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, Volume 1: 30th Design Automation Conference*, 2004, pp. 613–624, doi:10.1115/DETC2004-57357.

[23] Atkinson, A., Donev, A., and Tobias, R., *Optimum Experimental Designs, with SAS (Oxford Statistical Science Series)*, Oxford University Press, 2007.

[24] Huan, X. and Marzouk, Y., "Simulation-based Optimal Bayesian Experimental Design for Nonlinear Systems," *Journal of Computational Physics*, Vol. 232, No. 1, 2013, pp. 288–317, doi:10.1016/j.jcp.2012.08.013.

[25] Gharbi, A., Sarojini, D., Kallou, E., Harper, D., Petitgenet, V., Rancourt, D., Briceno, S., and Mavris, D., "STAnDD: A Single Digital Thread Approach to Detailed Design," *in AIAA SciTech 2017, 55th AIAA Aerospace Sciences Meeting, Grapevine, Texas*, 2017, pp. 1–13, doi:10.2514/6.2017-0693.

[26] Thomsen, B., Kokkolaras, M., Mnsson, T., and Isaksson, O., "Quantitative Assessment of the Impact of Alternative Manufacturing Methods on Aeroengine Component Lifing Decisions," *ASME Journal of Mechanical Design*, Vol. 139, No. 2, 2016, pp. 021401–1–021401–10, doi:10.1115/1.4034883.

[27] Pagan, J., Huynh, D., Schafer, S., Pinon, O., and Mavris, D., "Revolutionary Affordable Architecture Generation & Evaluation - Application to a System of Attritable UAVs," *in AIAA SciTech 2018, 2018 AIAA Aerospace Sciences Meeting, Kissimmee, Florida*, 2018, pp. 1–25, doi:10.2514/6.2018-1743.

[28] Gutowski, T., Hoult, D., Dillon, G., Neoh, E., Muter, S., E.Kim, and Tse, M., "Development of a Theoretical Cost Model for Advanced Composite Fabrication," *Composites Manufacturing*, Vol. 5, No. 4, 1994, pp. 231–239, doi:10.1016/0956-7143(94)90138-4.

[29] Corporation, N., "Advanced Composites Cost Estimating Manual (ACCEM)," *AFFDL-TR-76-87*, 1976, pp. 1–80.

[30] Bertsekas, D., *Dynamic Programming and Optimal Control*, Vol. I, Athena Scientific, 3rd ed., 2005.

[31] Tse, M., *Design Cost Model for Advanced Composite Structures*, PhD Thesis, MIT, 1992.

[32] Brooks, T. and Martins, J., "High-fidelity Aerostructural Optimization of a High Aspect Ratio Tow-steered Composite Wing," *in AIAA SciTech 2016, 57th AIAA/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference, San Diego, CA*, 2016, pp. 1–16, doi:10.2514/6.2016-1179.

[33] Tsai, S. and Wu, E., "A General Theory of Strength for Anisotropic Materials," *Journal of Composite Materials*, 1971, pp. 58–80, doi:10.1177/002199837100500106.

[34] Mindlin, R., "Influence of Rotatory Inertia and Shear on Flexural Motions of Isotropic, Elastic Plates," *ASME Journal of Applied Mechanics*, 1951, pp. 31–38.

[35] Reissner, E., "The Effect of Transverse Shear Deformation on the Bending of Elastic Plates," *ASME Journal of Applied Mechanics*, 1945, pp. A69–A77.

[36] Sutton, R., McAllester, D., Singh, S., and Mansour, Y., "Policy Gradient Methods for Reinforcement Learning with Function Approximation," *In Advances in Neural Information Processing Systems*, 2000, pp. 1057–1063.

[37] Kakade, S., "A Natural Policy Gradient," *In Advances in Neural Information Processing Systems*, 2002, pp. 1531–1538.