AEOLUS: Advances in Experimental Design, Optimal Control, and Learning for Uncertain Complex Systems
Center Progress Report

Contents

1 Introduction **1**

2 Multiscale modeling for advanced manufacturing and materials **1**
- 2.1 Phase-Field Models of Phase Change in Complex Systems . . . 1
- 2.2 Reduced-order modeling multiscale testbed problems with moving interfaces . . . 3
- 2.3 Nonlocal Cahn-Hilliard models for phase separation with sharp interfaces . . . 3

3 Model inadequacy **4**

4 Learning from data: Low-dimensional modeling and reduced models **6**
- 4.1 Lift & Learn: Learning low-dimensional models for an additive manufacturing solidification process . . . 6
- 4.2 Reduced-order models for multiscale PDE problems with moving interfaces . . . 7
- 4.3 Learning optimal reduced models from scientific data in advanced manufacturing and materials applications . . . 8

5 Transport methods for large-scale Bayesian inference and optimal design **10**
- 5.1 Stein variational Newton methods for BIP and OED . . . 11
- 5.2 Coupling methods for nonlinear ensemble filtering and smoothing . . . 11

6 Intrinsically Bayesian robust theory for optimal operator design, error estimation, experimental design, and model reduction **12**

7 Optimal uncertainty quantification for efficient experimental design of physical systems **13**

8 Optimization under uncertainty **14**

References **16**

A Organizational chart **18**

B Work breakdown **18**

C AEOLUS publications, September 2018 – September 2019 **20**
- C.1 Publications appeared or accepted . . . 20
- C.2 Publications submitted . . . 21

D AEOLUS invited presentations, September 2018 – October 2019 **22**
- D.1 Major keynote and plenary lectures at international conferences and other distinguished lectures . . . 22
- D.2 Other conference and workshop invited talks . . . 23

E Other significant dissemination **26**
E.1 Software 26
E.2 Short courses/summer schools taught and long courses developed 26
E.3 Books published or in manuscript form 26
E.4 Workshops and conferences co-organized 26
E.5 Other dissemination 27

AEOLUS: Advances in Experimental Design, Optimal Control, and Learning for Uncertain Complex Systems
Summary of Center Research

1. Introduction

The AEOLUS Center is dedicated to developing a unified optimization-under-uncertainty framework for (1) learning predictive models from data and (2) optimizing experiments, processes, and designs, all in the context of complex, uncertain energy systems. The AEOLUS center will address the critical need for a principled, rigorous, scalable, and structure-exploiting capability for exploring parameter and decision spaces of complex forward simulation models.

This report summarizes key progress in the first year of AEOLUS research. We first describe our efforts in multiscale modeling for advanced manufacturing and materials, including the development of a rich set of testbed problems (§2) and quantification of their model inadequacy (§3). We present several complementary approaches being pursued for development of low-dimensional and reduced-order models, with a particular focus on learning models from data (§4). We then highlight our progress in developing new methods to achieve optimal experimental design (§5,§6), uncertainty quantification in support of experimental design §7), and optimal design under uncertainty (§8).

2. Multiscale modeling for advanced manufacturing and materials

AEOLUS research is developing multiscale modeling capabilities and testbed problems in additive manufacturing and materials self-assembly. These testbed problems involve a complex interplay between multiple physical phenomena at multiple scales, and they embody the mathematical challenges that are the driving focus of the AEOLUS center: learning models from data and optimal control under uncertainty.

2.1. Phase-Field Models of Phase Change in Complex Systems

We are studying and developing models of phase change in polymeric materials, a subject at the center of modern methods of nanomanufacturing, and of related phenomena arising in the evolution of microenvironments in biological systems, particularly tumor growth in mammals. The Direct-Self-Assembly (DSA) of block copolymers describes the spontaneous phase change in multiphase materials, such as block copolymers, that can result in distinct patterns of monomer species with periodic structures of nanoscale dimension, some as small as 5 nm. These patterns guide lithography processes that produce templates for the fabrication of nanoscale devices that are very important in semiconductor technology. Similarly, phase-field models of vascular tumor growth share many features of the DSA models of block copolymers, although their overall complexity can be much greater. Among key challenges in these studies are the development and rigorous analysis of mathematical and computational models of phase change in polymers, the resolution of inverse problems whereby model parameters and initial conditions are determined that fit observational data or data stemming from higher-fidelity "ground truth" models, optimal control/design strategies to manipulate boundary conditions and environments to yield desirable morphologies, i.e. patterns in equilibrium states, and the development of computer codes and methods that produce accurate computer simulations based on these models.

We have developed implementations of the numerical solution of finite element approximations of continuum (phenomenological) models of phase change in copolymers, both Cahn-Hilliard (CH) models as well as nonlocal CH models following the Ohta-Kawasaki density functional theory. Such energy functional enables capturing typical possible morphologies of copolymers, as shown in Fig. 1. Rather extensive numerical experiments have been performed on local and nonlocal CH models, treating the model as one of gradient flow or, alternatively, treating the problem of determining equilibrium states as a nonconvex optimization problem, directly minimizing the free energy. The nonlocal component of the

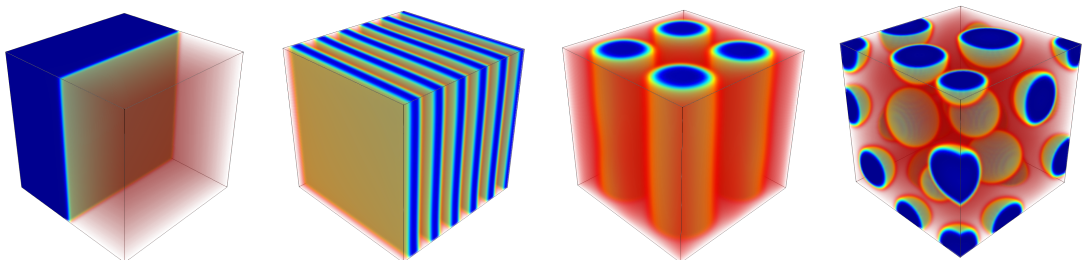


Figure 1: Morphologies of block-copolymers from the steady state finite element solutions of the nonlocal Cahn-Hilliard model: (a) macro-phase seperation; (b) lamellar; (c) cylinders; (d) spheres.

energy is introduced via a PDE constraint, along with a global mass conservation constraint, and the constrained optimization (non-convex due to the double well potential) is solved via Newton iteration on the KKT conditions, with Gauss-Newton globalization. This direct energy minimization method rapidly determines the final morphology for the nonlocal CH problem. Fig. 2 shows 2D simulations of block copolymer final morphology using the two solution strategies, indicating a 40X speedup using the new Newton method.

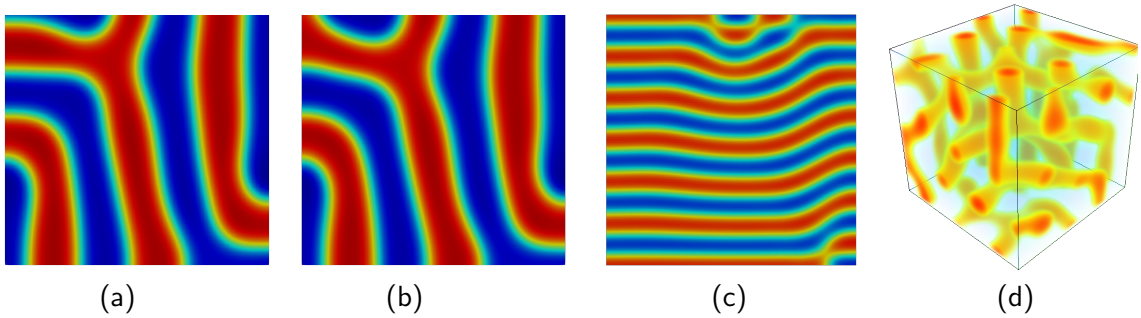


Figure 2: Equilibrium state of nonlocal Cahn-Hilliard equation: (a) gradient flow solution taking 3115 steps to reach equilibrium with a final energy of 1.874. (b) new Newton method based on direct minimization of nonconvex free energy taking 80 iterations to reach equilibrium with a final energy of 1.873. Morphology of a block copolymer simulated using models based on SCFT in (c) 2D and (d) 3D.

We have also just completed a detailed study of Self-Consistent Field Theory (SCFT) of block copolymers. It is known that the nonlocal Ohta-Kawasaki model can be derived from SCFT under suitable assumptions; SCFT models can be used to provide training data to calibrate the phenomenological parameters in the nonlocal CH model. We have developed preliminary finite element codes based on SCFT and have obtained numerical solutions of representative examples of 2D and 3D morphologies of block copolymers; see Fig. 2. At this juncture, the computation time required for numerical solutions based on the SCFT analysis is very lengthy compared to that of CH models and may generally be too expensive computationally to be competitive. Further work will be done on ways to make these simulations run more efficiently.

In future work, we intend to implement Bayesian approaches for parameter estimation and model validation developed under the learning components of AEOLUS for the nonlocal CH models using synthetic data furnished by higher fidelity SCFT approximations and imaging data provided by Brookhaven National Laboratory. The resulting inferred CH phase-field models, with quantified uncertainties in parameters, will then provide a basis for optimization of directed self-assembly processes by optimal control of substrate chemical boundary conditions—a challenging optimization under uncertainty problem.

2.2. Reduced-order modeling multiscale testbed problems with moving interfaces

Our driving applications comprise physical and chemical processes that result in rapidly evolving interfaces. The time evolution of these interfaces is strongly coupled to bulk properties of several fields. This strong coupling manifests mathematically as stiffness and ill-conditioning, which in turn dictate the use of high-resolution both in space and time. At each time step, nonlocal interactions need to be resolved, typically through the solution of elliptic PDEs in evolving domains with variable coefficients. As a result, simulations for such problems are expensive and present a major computational barrier for higher level tasks like parameter estimation, design, and uncertainty quantification. Our overarching goal is to construct reduced order models that can be used to dramatically accelerate the forward problem.

The first driving application we consider for constructing reduced order models is **additive manufacturing**. In this problem, the interface between different microstructures (composition, phase) evolves as a function of the temperature, composition, geometry, and interfacial energies. Solidification and crystal growth result in a microstructure that controls the bulk properties and the functionality of the material. For our initial investigations, we have chosen a three-species, phase-field model that is a good approximation of a more complex multispecies model for the microstructure evolution of a nickel-base superalloy during laser powder bed fusion [19]. Our model includes temperature, chemical potential, and phase. The phase-field tracks the evolution of the solid-liquid interface, and can reproduce classical features of dendritic evolution, for example the formation of primary and secondary arms.

Its mathematical formulation involves a set of nonlinear parabolic equations for the three species evolution, with elliptic operators involving nonlinear, inhomogeneous, and anisotropic coefficients. The formulation couples two time-dependent diffusion equations (temperature and chemical potential) to a time-dependent reaction-diffusion equation (phase field). Although there have been several studies that consider a single dendrite, we consider a multiscale setup in which we can simultaneously consider 100s to thousands of dendrites and simulate the properties of the bulk structure as depicted in Figure 3.

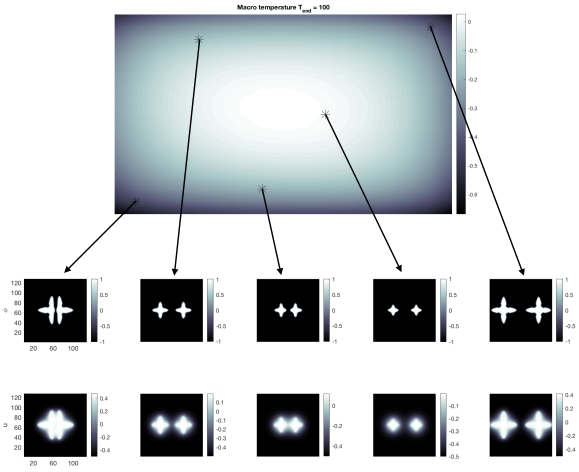
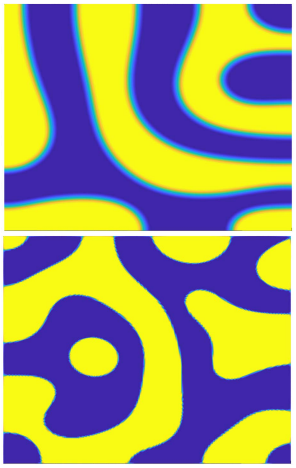


Figure 3: Illustration of a multiscale simulation to track the solidification of a material with microstructure. The top geometry indicates the macroscale model in which we solve a variable coefficient heat equation with source term and heat conductivity that depend on the microstructure. In the bottom two rows of figures, the top row depicts the phase field function (white is the solid phase and black is the liquid phase) at different points of the macro domain. The bottom figure depicts the chemical potential.

Our second application for reduced-order models concerns the **rheology of complex fluids**, that is, fluid suspensions with a large number of deformable particles. The suspension dynamics is governed by long-range hydrodynamic interactions between the deformable particles and the elastic energy of the particle-fluid interface, e.g., due to bending and tension. We have chosen this application because we have an existing integral equation solver that can simulate the forward problem with spectral accuracy. The suspension evolution is governed by quasi-static Stokes equations with jump interface conditions, where the jump is related to the derivatives of the elastic energy of the interface. Due to highly nonlinear fluid-structure interaction dynamics, the moving interfaces, and the multiple scales, numerical simulations of such flows are challenging and expensive.

2.3. Nonlocal Cahn-Hilliard models for phase separation with sharp interfaces



Cahn-Hilliard partial differential equation models are in common use for modeling problems where different substances are separated by regularized interfaces, that is, diffuse interfaces that approximate sharp physical interfaces. The thickness of the diffuse interface is proportional to a regularizing parameter so it cannot be reduced via grid-size reduction; see the top figure. On the other hand, the use of nonlocal Cahn-Hilliard models can result in sharp interfaces so that if, for example, continuous finite element discretizations are used, the thickness of the interface is of order of the grid size (see the bottom figure) and if discontinuous FEMs are used, the interface can be totally sharp. Thus, we propose a nonlocal variant of the Cahn-Hilliard interface model with a non-smooth potential of double-well obstacle type. We have considered different variants of the nonlocal contributions in the model such as nonlocal operators and nonlocal boundary conditions. In contrast to the local setting, the proposed model allows for sharp interfaces. Here, the choice of the obstacle potential plays an important role in our analysis, because it guarantees the strict separation of the substance into pure phases for nontrivial interactions. Mathematically, this introduces additional inequality constraints that, in a weak form, lead to a coupled system of variational inequalities, which at each time instance can be restated as a constrained optimization problem. We analyze a discretization of the problem in space and time based on finite elements and implicit-explicit time stepping methods that can be realized efficiently. We provide numerical experiments to support our theoretical findings in one and two spatial dimensions.

3. Model inadequacy

Work on physics-based model inadequacy representations for complex mathematical models has focused on two areas during the past year. First, in continuation of previous work, an inadequacy formulation is being developed for a model of scalar transport through a heterogeneous porous medium. Second, these ideas are being extended to develop an inadequacy representation for the non-local Cahn-Hilliard model based on self-consistent field theory (SCFT) in the context of the block co-polymer application. In both problems, we aim to exploit the existence of a hierarchy of physical modelss to build realistic inadequacy representations for low fidelity physical models.

Porous Media Transport through heterogeneous porous media depends on dynamics across a wide range of scales, the smallest of which can neither be resolved nor observed. For this reason homogenized models of mean transport neglect the smallest scales, though they are known to affect macroscale dynamics. Similar issues arise in multiscale material models and many other physical modeling domains, due to limited computational resources.

Thus, work has focused on the process by which an inadequacy representation can be formulated and informed for problems with unresolved microscale dynamics. To represent the model hierarchy in the porous media flow application, two simplified models were developed: a higher-fidelity model that resolves all relevant scales, and a lower-fidelity model that does not resolve the smallest scales. Both models are much simpler than would be required for a practical problem, but the nature of the inadequacy in the low-fidelity model is the same as those used in practice. The inadequacy representation is then intended to account for the uncertainty introduced in the low-fidelity model by modeling the effects of the small scales. Specifically, a stochastic operator is added to the low-fidelity model evolution equations at the point where the effects of the microscale dynamics are modeled. An important part of inadequacy model development is identifying constraints that prior knowledge places on the representation. In this case, linearity and spatial shift invariance of the governing equations lead to requirements that the operator be linear, enabling its characterization by its eigendecomposition, and requiring its eigenfunctions to be the Fourier modes, respectively. Since the eigenfunctions are fixed, its eigenvalues

must be the source of stochasticity in the representation.

In previous inadequacy treatments, the form of the distribution of the eigenvalues of the operator would be assumed, and the hyperparameters of assumed form would then be inferred. This year work has instead focused on a using direct simulation of the high-fidelity model to compute samples from the eigenvalues' distribution. Preliminary results of this sampling indicate high levels of correlation between the eigenvalues, especially at higher frequencies, as shown in Figure 4.

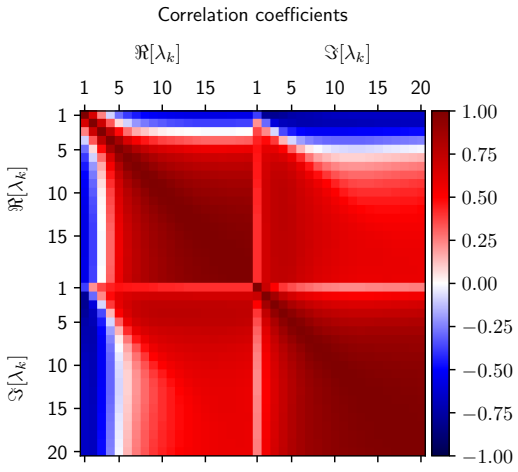


Figure 4: Correlations between the real and imaginary parts, as well as correlations across wavenumber are significant.

Sensitivity analysis studies were performed to discover which aspects of the eigenvalues' distributions are most important for prediction. Using summary statistics computed from eigenvalue samples, different distributions were constructed that, e.g. included information about correlations without perfectly reproducing marginal distributions, or reproduced marginal distributions without accounting for correlations. The different distributions were then forward propagated to generate sample evolutions of the transported concentration field. As shown in Figure 5, accounting for correlations leads to evolutions that better maintain positivity of concentrations, as is physically required.

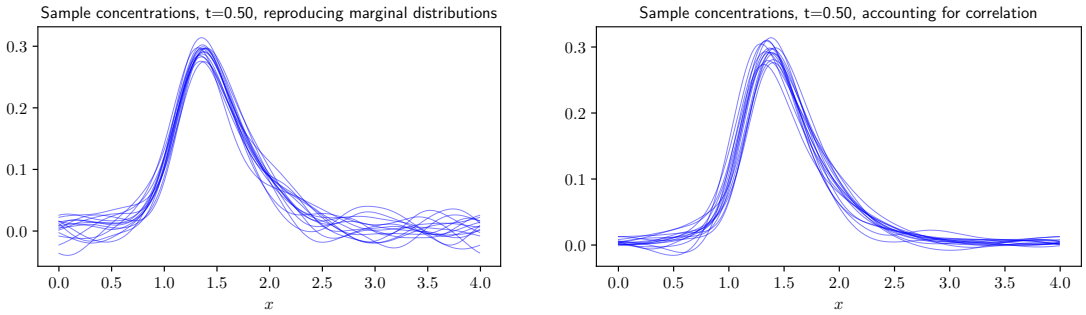


Figure 5: The distribution focused on characterizing the marginal distributions of the eigenvalues (left) produces nonphysical concentration fields with strong oscillations in their tails. The distribution that focused on characterizing the correlation structure produced much more physical evolutions.

Block Co-Polymer As in the porous media problem, in the block co-polymer application, a hierarchy of physical models is available. Of these models, one of the most computationally tractable is the non-local Cahn-Hilliard-like model, which describes the macroscopic monomer density in terms of a scalar PDE. As shown by Choksi and Ren [5], this model can be derived from a higher-fidelity model—a self-consistent mean field theory [14]—by invoking a number of simplifying assumptions. These assumptions include linearization about a high temperature condition and use of long- and short-wave approximations to analytically compute an inverse Fourier transform. The result of these assumptions is that the relationship between the density and the potential is greatly simplified, such that the potential can be eliminated analytically. To derive an appropriate model inadequacy representation, we are working to develop random perturbations of the assumptions that are consistent with SCFT while maintaining the tractability of the non-local Cahn-Hilliard-like model.

4. Learning from data: Low-dimensional modeling and reduced models

AEOLUS is pursuing several different directions to tackle the challenging task of deriving low-dimensional models and reduced-order models for our target multiscale problems in additive manufacturing and materials self-assembly.

4.1. Lift & Learn: Learning low-dimensional models for an additive manufacturing solidification process

One of the applications under development as part of the Exascale Computing Project (ECP) is ExaAM, the Exascale Additive Manufacturing project, which is developing a set of simulation capabilities for modeling additive manufacturing processes, with emphasis on laser and electron beam powder bed fusion. One of the key physical phenomena is solidification, for which phase field provides a mature, well-posed formulation. However, it is extremely expensive computationally because of the stringent spatial resolution requirements associated with the use of a diffuse interface especially in the presence of high curvatures and/or in the presence of solute diffusion. In addition, there are several model parameters whose determination in the context of the phase field formulation are tied to the mathematical description of the diffuse interface and therefore pose restrictions on the temporal and spatial resolution required for model convergence and relevance to the behavior of real material systems. Therefore, developing reduced-order models (ROMs) for phase field simulations of solidification is an attractive test bed for AEOLUS.

The mechanical properties of additively manufactured parts are influenced by the solidification microstructure, and in order to be able to build parts with desired properties one must be able to control the microstructure (e.g., size and orientation of the dendrites, primary and secondary dendritic arm spacing, solute concentration, etc.). Forward simulation models that have the required spatial resolution to accurately capture the solidification structure of an additively manufactured part have computational costs that are prohibitively high for achieving control. In AEOLUS, we are pursuing reduced-order modeling as a critical enabler for achieving models that are sufficiently accurate and at the same time sufficiently efficient to achieve control. Our target problems pose several challenges for existing model reduction methods. For example, our initial testbed model couples phase field simulation of the solidification with the heat equation. For this mdoel, the character of the temperature and order parameter state solutions changes drastically throughout a given simulation (as the solidification front progresses) and the forward simulation models have highly nonlinear dependence on system parameters. Classical methods such as the proper orthogonal decomposition (POD) that seek straightforward approximations of the state in a linear subspace require many modes to achieve accuracy in a problem such as this one, rendering the resulting POD reduced models inefficient.

We have developed a projection-based ROM for a one-dimensional model problem of the testbed described in §2.2. This testbed problem solves a phase field simulation of the solidification coupled with the heat equation. Written in their native form, the governing equations solve for the spatial and time evolution of the temperature field $T(x,t)$ and the order parameter $\phi(x,t)$. The order parameter ϕ denotes the phase field state, where $\phi = 0$ and $\phi = 1$ correspond to the liquid and solid states, respectively. These two coupled nonlinear equations can be transformed to a set of cubic differential algebraic equations (DAEs), with the introduction of six auxiliary states. The resulting lifted DAE system then has eight unknown variables (compared to two for the original system), but the cubic structure of the lifted formulation is desirable because it admits an efficient ROM that does not require additional approximations of the nonlinear terms (i.e., it avoids so-called "hyper-reduction" which is used in existing ROM methods but often fails for multiscale problems). We call this reduced-order modeling approach "Lift & Learn."

We do *not* discretize the lifted PDE system, but rather collect snapshots of the original system

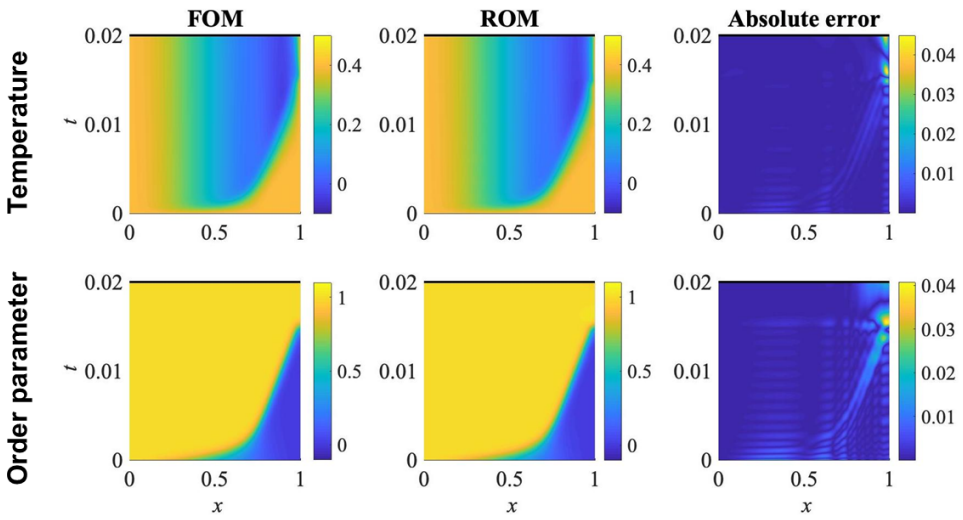


Figure 6: A comparison of full-order model (FOM) and ROM predictions of temperature and order parameter. The FOM and ROM have 2002 and 23 degrees of freedom, respectively.

simulation and then apply the lifting transformations to the snapshots. This gives us a set of snapshots representing temperature T, order parameter ϕ, and the six auxiliary variables. From these snapshots, we compute the POD basis and then learn the ROM operator coefficients (linear, quadratic and cubic terms) from projected snapshot data, using a least-squares operator inference method. Figure 6 compares ROM predictions with full model predictions for one dimensional solidification with initial temperature field $T(x, 0) = 0.4$ and initial phase field profile defined by $\phi(x, 0) = 0.5 \cos(\pi x) + 0.5$. The dimension of the POD basis is selected by retaining 99.9% of the cumulative energy of the POD singular values. This results in a ROM dimension of 23: seven modes for differential equations and 16 modes for the algebraic equations.

The UT and ORNL teams are working closely together to develop these reduced-order modeling capabilities in the context of ExaAM application targets. In order to make the initial demonstration more computationally tractable, we have worked to reduce the ExaAM solidification component (MEUMAPPS-SL) for simpler materials (e.g. pure metals). In FY20 we will demonstrate the full "lift and learn" process for the initial problem and begin designing more realistic challenge problems as next steps. Ultimately we plan to leverage the ROMs developed within AEOLUS in ExaAM within design optimization loops.

4.2. Reduced-order models for multiscale PDE problems with moving interfaces

We have combined novel discretization schemes with machine-learning in order to construct an effective reduced model, and we have extended this reduced model construction methodology to numerical homogenization algorithms. Both methods are related to microstructure evolution of systems governed by multispecies nonlinear PDEs.

Integration of operator-splitting and machine learning. The frequent solution of large linear systems (typically discretizations that attempt to capture nonlocal interactions) creates a major computational bottleneck. Using our complex fluid simulation framework, we have proposed a machine learning augmented reduced model for such problems. The basic idea is to devise an operator-splitting scheme that allows the construction a high-dimensional approximation for expensive interface operators. We have reported results this method in [16]. Our model replaces expensive nonlinear operators with regression functions. Given the physical parameters of the interface, our model generalizes to arbi-

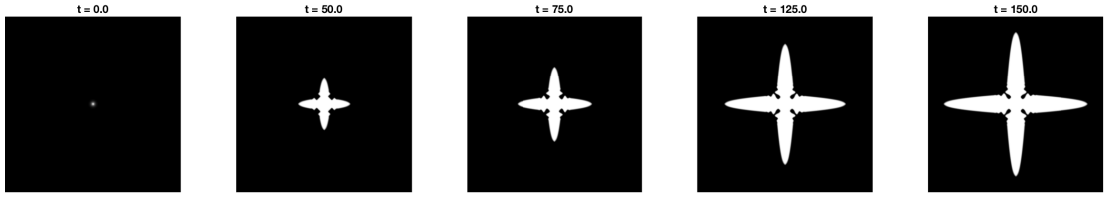


Figure 7: Time evolution of the phase field function of a single dendrite using a three-species PDE model (temperature, chemical potential, and phase field).

trary geometries and boundary conditions without retraining the regression approximation. Our reduced model simulations are qualitatively as accurate as and approximately an order of magnitude faster than the simulations performed with the numerical scheme using the same number of degrees of freedom as the reduced model. The important features of our scheme are that (1) it generalizes to dynamics that are completely different than the ones used in the training phase; (2) it is tightly integrated with the forward problem. Our method is not a black box approximation and the reduced model still uses several components of the PDE forward solver.

Multiscale phase-field models for crystal growth. We have completed the implementation of the homogenization scheme described in [13] in one and two dimensions. That scheme significantly accelerates dendrite growth simulations since it doesn't require full revolution of the forward problem. However, this scheme two-way couples the macroscale and microscale and ends up requiring hundreds of single-dendrite simulations to capture the growth dynamics (see Figure 3 for the full model dynamics and Figure 7 for the single dendrite dynamics). As a result it is still very expensive, even in 2D. To address this cost, we are considering an adaptive scheme for the macro-micro model in which we don't track the evolution of all the dendrites but only the ones for which the dynamics are sufficiently different some default representative dynamics. This requires methods for defining similarity between dendrites and their state, methods to interpolate between different dendrite shapes, and robust error criteria when we do such regressions. Preliminary results using a metric that tracks the temperature history of a dendrite give less than one per cent error when the latent heat and initial shapes don't vary significantly in the macro domain. But if there are significant coarse-scale variations in properties that control growth (e.g., latent heat) this approach can lead to 50% or higher errors. We are currently examining algorithms to better control this error while speeding up the overall computations. In a different research thread (that uses the same model), we have also start working on a optimization problem in which we control the heat deposition in order to optimize the final microstructure.

4.3. Learning optimal reduced models from scientific data in advanced manufacturing and materials applications

In this section we summarize our recent accomplishments towards learning optimal reduced models from scientific data in advanced manufacturing and materials applications, described in our recent efforts [8, 9, 20]. In addition, we have also developed several other related efforts whose details are described in [6, 7, 11, 12, 15, 18, 23, 24, 25].

The Natural Greedy Algorithm (NGA) for reduced bases in Banach spaces. We recently introduced and analyzed a novel reduced basis approach, used to construct an approximating subspace for a given set of data. Our technique, which we call the Natural Greedy Algorithm (NGA), is based on a recursive approach for iteratively constructing such subspaces, and coincides with the standard, and the extensively studied, Orthogonal Greedy Algorithm (OGA) in a Hilbert space. However, for a given set of data, the NGA is straightforward to implement and overcomes the

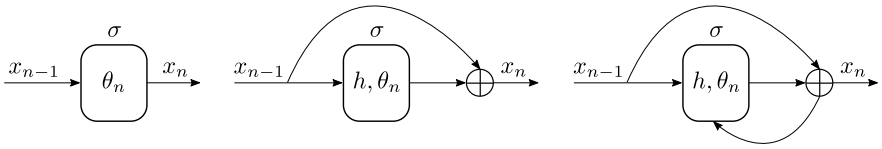


Figure 9: From left to right: feed forward layer, residual layer, proposed implicit residual layer.

explosion in computational effort introduced by the OGA when applied in a general Banach space, as we utilize an entirely new technique for projecting onto the appropriate subspaces. We rigorously analyzed the NGA, and demonstrate that it's theoretical performance is similar to the OGA, while the realization of the former results in significant computational savings through a substantially improved numerical procedure. Furthermore, we show that the empirical interpolation method (EIM) can be viewed as a special case of the NGA. Finally, several numerical examples are used to illustrate the advantages of our NGA compared with other greedy algorithms and additional popular reduced bases methods, including EIM and proper orthogonal decomposition (POD). For example, consider the following parametric family given by

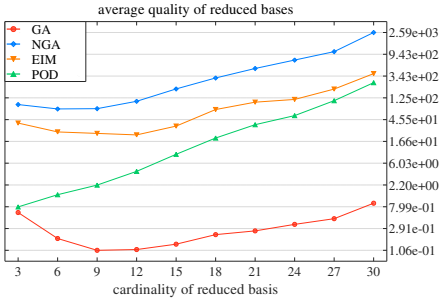


Figure 8: The average quality of various reduction techniques ranging from $m = 3$ to $m = 30$ reduced bases.

$$\mathcal{F}(x, \mu_1, \mu_2) = \exp(x + 2\mu_1 + 3\mu_2)\Big(\exp\Big(-\pi \Big| x - \frac{\mu_1}{2} \Big| \Big) \arcsin\Big(\sin\big(2\pi \exp(\mu_1)\, x\big)\Big) + \exp\Big(-\pi \Big| x + \frac{\mu_2}{2} \Big| \Big) \arcsin\Big(\sin\big(\exp(\pi - \mu_2)\, x\big)\Big)\Big),$$

where $(\mu_1, \mu_2) \in \Omega = [-2, 2]$, and $x \in \mathcal{D} = [0, 2] \times [0, 2]$. Figure 8 show the quality and the convergence respectively, of various reduced basis techniques for the parametric family in $\mathcal{X} = L_1[-2, 2]$. We note that in this case all algorithms perform similarly in terms of approximation accuracy, however the computational complexity of NGA appears to be the smallest (even smaller than that of EIM) due to the especially simple formula for norming functionals in case $p = 1$.

Robust learning with implicit residual networks. In the work [20] we proposed an entirely novel deep architecture utilizing residual blocks inspired by implicit discretization schemes. As opposed to the standard feed-forward networks, the outputs of the proposed implicit residual blocks are defined as the fixed points of the appropriately chosen nonlinear transformations. We show that this choice leads to improved stability of both forward and backward propagations, has a favorable impact on the generalization power of the network and allows for higher learning rates. In addition, we consider a reformulation of ResNet which does not introduce new parameters and can potentially lead to a reduction in the number of required layers due to improved forward stability and robustness. For the second example, we consider another small test problem illustrated in Figure 10. It consists of 513 points organized in two differently labeled spirals. Here we use the implicit network architecture with 6 hidden nodes on each of the 25 layers and $tanh$ activation insted of $ReLU$. Figure 10 illustrates convergence of the networks and the classification results. One can see that the proposed implicit scheme is more accurate and robust than the classical ResNet. Finally, we also derive the memory efficient reversible training algorithm and provide numerical results in support of our findings.

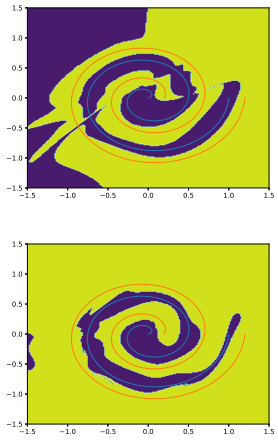
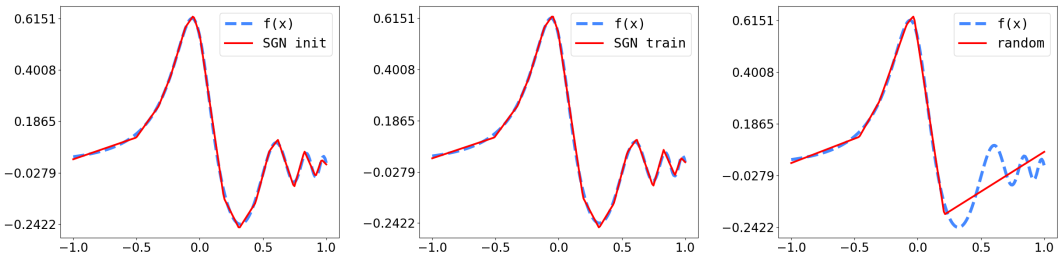


Figure 10: Classification results of the classical ResNet (L) and implicit networks (R).

Greedy shallow networks (GSN): an innovative approach for constructing and training neural networks. In [8] we developed a novel greedy approach to obtain a single layer neural network approximation to a target function with the use of a ReLU activation function. In our approach we construct a shallow network by utilizing a greedy algorithm where the set of possible inner weights acts as a parametrization of the prescribed dictionary. To facilitate the greedy selection we employ an integral representation of the network, based on the ridgelet transform, that significantly reduces the cardinality of the dictionary and hence promotes feasibility of the proposed method. Our approach allows for the construction of efficient architectures which can be treated either as improved initializations to be used in place of random-based alternatives, or as fully-trained networks, thus potentially nullifying the need for training and/or calibrating based on backpropagation. Numerical experiments demonstrate the tenability of the proposed concept and its advantages compared to the classical techniques for training and constructing neural networks. See, e.g., Figure 11 where we compare the new deep network architecture designed with our GSN approach with an architecture designed by classical random initialization.



(a) 3.27e-02 **(b)** 3.09e-02 **(c)** 6.96e-02

Figure 11: Approximation comparisons via: (a) GSN initialization, (b) neural network with GSN initialization, (c) neural network with random initialization. Values under the images indicate the ℓ_2-approximation error on the test set.

5. Transport methods for large-scale Bayesian inference and optimal design

Learning from data, with quantified uncertainty, is a cross-cutting problem in the AEOLUS effort. Moreover, our broader goal is to "close the loop" between inference and observation—using existing models (accounting for their parametric uncertainties and inadequacies) to drive the optimal collection of new data, where optimality is defined according to particular prediction or design goals. These problems induce enormous computational challenges: Bayesian inference in high or infinite parameter dimensions, and in generically non-Gaussian settings; inference for dynamical models with sequentially arriving data; and an outer loop of optimization for optimal experimental design (OED), where inference must be performed for multiple realizations of the data at each iteration.

To address these challenges, we are focusing on optimization-driven inference methodologies that employ *transportation of measure* as a unifying concept. Compared to more traditional approaches such as Markov chain Monte Carlo (MCMC) or sequential Monte Carlo (SMC), transport offers potentially significant gains in efficiency; a natural ability to harness modern massively parallel and heterogeneous computing hardware; multiple "hooks" for identifying and exploiting low-dimensional structure; the ability to harness intrusive (e.g., gradient and Hessian) model information when available, and to proceed efficiently when it is not; and rich tradeoffs between computational effort and accuracy. Many transport approaches can be viewed as variational Bayesian alternatives to asymptotically exact sampling methods such as MCMC, but with a different *representation* of the posterior distribution (or the joint parameter-data distribution) that is adaptable and infinitely refinable.

We are building on some of our past work in transport-based inference to create: (i) new nonparametric variational inference methodologies; (ii) fast estimators of posterior normalizing constants for

OED; and (iii) ensemble-based sequential inference methods. We detail these developments below.

5.1. Stein variational Newton methods for BIP and OED

To address the critical curse-of-dimensionality challenge in Bayesian inverse problems, we developed a projected Stein variational Newton (pSVN) method [4] based on SVN [10], which constructs a composition of transport maps represented by radial basis functions that push forward the samples drawn from prior distribution of the parameter to its posterior distribution, by minimizing a Kullback–Leibler (KL) divergence between the pushforward prior and the posterior We exploited the essential property that the posterior often departs from the prior in only an intrinsic low-dimensional subspace, and projected the high-dimensional parameters to this subspace, informed by the Hessian of the parameter-to-observable map. We demonstrated the advantages of pSVN in (1) effectively alleviating the curse of dimensionality, (2) preserving high accuracy in computing statistics of some QoI w.r.t. the posterior, (3) converging rapidly with convergence rate independent of the number of samples, (4) achieving good strong scalability w.r.t. the number of cores. More details are provided in Figure 12.

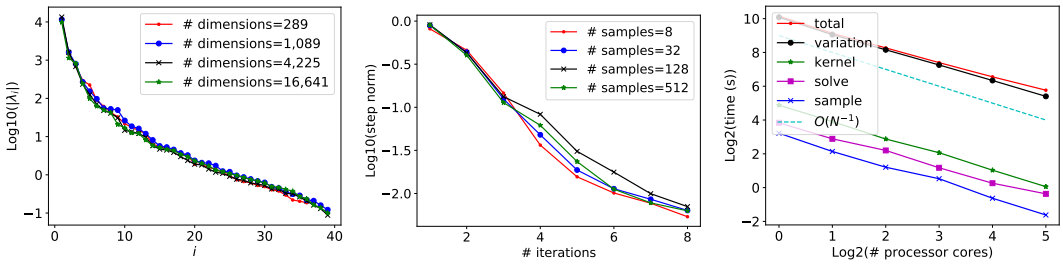
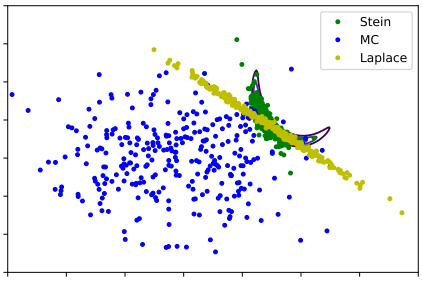


Figure 12: Left: Decay of eigenvalues $\log_{10}(|\lambda_i|)$ with increasing dimension d. Middle: Decay of stopping criterion—the averaged norm of the update $w^l - w^{l-1}$ w.r.t. the iteration number l, with increasing number of samples. Right: Decay of the wall clock time (seconds) of different computational components w.r.t. increasing number of processor cores.

We extended SVN based methods to compute the expected information gain (EIG) [22], a key objective in Bayesian OED, which consists of (1) efficient and accurate sampling from the posterior, (2) an effective decoupling technique (construction and evaluation) to compute the often intractable model evidence appearing in EIG. This method enjoys the combined merits of superior efficiency over MCMC sampling and greater accuracy than Laplace approximation, thus mitigating the cost of the inner Bayesian inverse problem (Figure 13).



d = 2	Reference	Prior	Stein	Laplace
EIG	**14.74**	21.08	**14.71**	15.35
d = 4	Reference	Prior	Stein	Laplace
EIG	**13.02**	29.47	**12.98**	14.40

Reference computed with 10^6 samples from prior. The number of samples from prior, Stein, and Laplace are 500. For Stein, we use 300 samples for map construction and 200 samples for evaluation.

Figure 13: Left: contours of banana-shaped posterior density and samples obtained by Monte Carlo, Stein, and Laplace sampling, indicating superiority of Stein. Right: numerical results for EIG evaluation in dimensions $d = 2, 4$.

5.2. Coupling methods for nonlinear ensemble filtering and smoothing

A grand challenge in Bayesian computation involves the development of online inference algorithms for dynamical models, such as the coupled nonlinear PDE models proposed for additive manufacturing processes in Section 2. Here we wish to estimate both the parameters and state of the system, from

sequentially arriving data, with a cost that does not grow in time and an error that remains uniformly bounded. If the dynamics are linear and all distributions are Gaussian, the classical *Kalman filtering* and smoothing recursions provide a complete solution. In the nonlinear and non-Gaussian application problems at hand, however, inference becomes far more challenging. Existing methods are typically either *consistent* but restricted in dimensionality and scale (e.g., particle filters, smoothers, and more general SMC algorithms for state space models), or relatively scalable but inconsistent with the Bayesian solution and unable to provide any guarantees for meaningful UQ (e.g., the ensemble Kalman filter (EnKF)).

To address these gaps, we seek a consistent, recursive Bayesian inference methodology that preserves many of the desirable features of ensemble methods—e.g., practical performance in high dimensional systems, tractable and non-intrusive computations—while addressing the limitations just described.

Our current work has proposed a new mathematical framework for generalizing the EnKF through the lens of measure transport [21]. We seek *nonlinear* transport maps that couple the prior to the posterior, and methods for estimating these maps efficiently from limited ensembles, using convex optimization. This approach is also much less reliant on tuning parameters to achieve optimal performance.

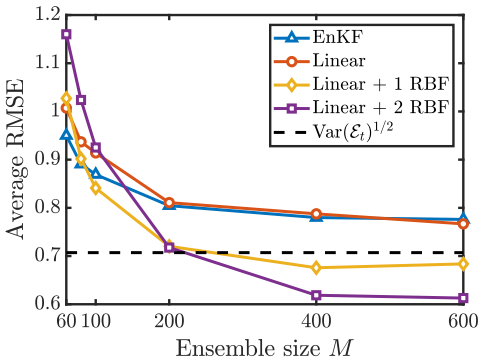


Figure 14: Average RMSE (over 2000 assimilation cycles) for the "hard" Lorenz-96 configuration of [17]. Dashed line is the standard deviation of the observational noise.

The underlying idea of this framework is to build a mapping that pushes forward samples from the joint probability distribution of state and data to the posterior distribution over the state, which is a particular conditional of this distribution. This map is built by first estimating a Knothe–Rosenblatt (KR) map from the joint state-data distribution to a reference or "bridging" measure (e.g., a standard Gaussian), and then inverting part of this map to build the prior-to-posterior transformation. Since only sampling from the likelihood is required, this inference approach can be seen as an instance of approximate Bayesian computation (ABC). The key KR estimation step involves only convex optimization, and the framework offers many opportunities for introducing structure (e.g., sparsity, localization, and low rank) that can improve stability and reduce computational cost. Figure 14 demonstrates filtering performance on a challenging configuration of the benchmark Lorenz-96 system; here, introducing very simple and limited nonlinearities in the map yields significant improvements in state estimation performance over the EnKF, at minimal computational cost [21].

6. Intrinsically Bayesian robust theory for optimal operator design, error estimation, experimental design, and model reduction

An *intrinsically Bayesian robust (IBR) operator* provides optimal expected performance relative to a prior distribution over an uncertainty class of models. It is termed an *optimal Bayesian operator* if new data are incorporated to produce a posterior distribution. The *mean objective cost of uncertainty (MOCU)* is the expected increase in cost from applying an IBR operator instead of an optimal model-specific operator in each model. *Optimal experimental design (OED)* selects an experiment that maximally reduces expected MOCU, which is evaluated relative to the distribution conditioned on the experiment.

Optimal Bayesian classifier. [27] We construct an optimal Bayesian classifier (OBC) with serially dependent training observations. The training observations are generated from a multidimensional vector autoregressive process, and there exists uncertainty about parameters governing the model. We find an OBC under the assumption of known covariance matrices of white-noise processes.

Robust filtering and experimental design. [26] We derive robust linear filtering and experimental design for systems governed by a stochastic differential equation (SDE) under model uncertainty, assuming that the physical process is modeled via a SDE system with unknown parameters, and the

signals are degraded by blurring and additive noise. Since the system is nonstationary, the Wiener-Hopf equation is not solved in closed form. Hence, we discretize the problem to obtain a matrix system.

MOCU-based active learning for block polymer phase identification. We have a computational model (for example, PDE or its surrogates) which connects the features (x) of the block polymer of interest with its phases (y). We wish to identify the phase by characterizing $\Pr((y|x,\theta)$, where θ parameterizes uncertainty. The uncertainty can be reduced by sequential data collection (active learning/sampling). We solve the problem to sequentially choose samples from the feature (design) space for phase identification. The utility function is to reduce the cost of the uncertainty based on the residual error by the IBR classifier, leading to MOCU-based active learning.

Multi-Objective MOCU. Given two cost functions that measure different aspects of operational performance, we define a weighted cost function, an IBR operator relative to this cost function, and MOCU for each specific value of the weight λ. The multi-objective MOCU as the average over λ relative to a distribution of λ. The concept extends to a finite number of cost functions.

Optimal-Bayesian-transfer-learning for classifier error estimation. Transfer learning simultaneously learns from source domains that have lots of labeled data, and transfers the relevant knowledge to a target domain with limited labeled data to improve prediction performance. Optimal Bayesian Transfer Learning (OBTL) has source and target domains related through a joint prior density of the model parameters. We extend the theory developed for learning OBTL classifiers to error estimation on the target domain utilizing data from both the target and source domains.

MOCU-based reduced order modeling. When modeling complex systems, physical (spatial-temporal) fields of interest are often governed by certain physics principles. PDE and large-scale ODE systems after discretization are often the computation tools to simulate the systems behavior so that design, control, and optimization can be studied. However, for these high-consequence decision making problems, often we have to solve large-scale PDE/ODEs many times. Reduced-order models (ROMs) aim to achieve the best approximation to the original full-order model (FOM). For operational objectives such as design or control, we are formulating ROMs that have minimal performance reduction with respect to the objectives.

7. Optimal uncertainty quantification for efficient experimental design of physical systems

Autonomous REsearch systems (ARESs), colloquially called robot scientists, offer the advantage of high throughput scientific research. AERSs are characterized by their ability to autonomously execute experiments, perform on-board characterization of experimental responses, learn from these observations and dynamically plan subsequent experiments, often with a particular experimental objective in mind. Examples of such ARES include ADAM, EVE and ARES. Our work this past year deals with the planning component of an ARES, which uses an aggregate of prior knowledge (accounting for uncertainty) and collected data to select the experiments to perform. Because planning on an ARES is to be done autonomously, it is important to be systematic and explicit in the planning logic. More specifically, it is important to properly encode 1) the tunable knobs or choices the ARES has control over 2) the overall goal or objectives of running the experiment, 3) what data the ARES has as observations, 4) some definition of state, from state of belief (what the ARES) knows about the system to be studied, to some physical state (how the ARES is currently set, for example) and finally 5) an understanding of how that state evolves after each experiment. In our work this year we have applied Bayesian experimental design (BED) techniques to the design of experiments under uncertainty. We have focused on three key ingredients. First is how to model state in an ARES. Second is a consideration of the overall learning objectives, or modalities – how specific modalities impact the appropriateness of the decision-making rules, or policies that govern which experiments an ARES performs. As a demonstration we have applied BED to a problem that is ubiquitous in physics, chemistry and materials sciences: phase diagrams. We shall see how BED techniques can be applied to the problem of efficiently learning key features of phase

diagrams, or the phase diagram in its entirety.

A common goal in areas of science and engineering that rely on making accurate assessments of performance and risk (e.g. aerospace engineering, finance, geophysics, operations research) for complex systems is to be able to guarantee the quality of the assessments being made. Very often, the knowledge of the system is incomplete or contains some form of uncertainty. There can be uncertainty in the form of the governing equations, in information about the parameters, in the collected data and measurements, and in the values of the input variables and their bounds. One of the most common cases is that initial conditions and/or boundary conditions are known only to a certain level of accuracy. Even in the case where the dynamics of a system is known exactly at a fine-grained level, computationally more tractable coarse-grained models of the system often have to be derived under approximation, and thus contain uncertainty. One way to refine a model of a system under uncertainty is to perform experiments to help solidify what is known about the parameters and/or the form of the governing equations. Sampling methods (such as Monte Carlo) can be used to determine the predictive capacity of the models under the resulting uncertainty. Unfortunately, this determination can be computationally costly, inaccurate, and in many cases impractical.

One promising approach to dealing with this challenge is Optimal Uncertainty Quantification (OUQ) developed by Owhadi et al. OUQ integrates the knowledge available for both mathematical models and any knowledge that constrains outcomes of the system, and then casts the problem as a constrained global optimization problem in a space of probability measures; this optimization is made tractable by reducing the problem to a finite-dimensional effective search space of discrete probability distributions, parameterized by positions and weights. Much of the early work with OUQ has been to provide rigorous certification for the behavior of engineered systems such as structures under applied stresses or metal targets under impact by projectiles.

In this past year we have demonstrated the utility of the OUQ approach to understanding the behavior of a system that is governed by a partial differential equation (and more specifically, by Burgers' equation). In particular, we solved the problem of predicting shock location when we only know bounds on viscosity and on the initial conditions. We have calculated the bounds on the probablity that the shock location occurs at a distance greater than some selected target distance, given there is uncertainty in the location of the left boundary wall. Through this example, we demonstrated the potential to apply OUQ to complex physical systems, such as systems governed by coupled partial differential equations. We compare our results to those obtained using a standard Monte Carlo approach, and show that OUQ provides more accurate bounds at a lower computational cost. As OUQ can take advantage of solution-constraining information that Monte Carlo cannot, and requires fewer assumptions on the form of the inputs, the predicted bounds from OUQ are more rigorous than those obtained with Monte Carlo.

8. Optimization under uncertainty

We have developed and further extended our computational framework for optimal control and design governed by PDEs under uncertainty with mean-variance risk measure, using Taylor approximation and variance reduction techniques [3], along two lines: (1) multiobjective optimization under uncertainty, which we have applied to optimal design of metamaterials [2]; and (2) chance-constrained optimization under uncertainty, with application to optimal control of subsurface flow [1].

We posed multiobjective optimization under uncertainty subject to multiple PDE constraints as

$$\min_{z} \mathcal{J}(\{u_n\}_{n=1}^{N}, z) \quad \text{subject to} \quad \mathcal{R}_n(u_n, m, z) = 0,\ i = 1, \ldots, N, \tag{1}$$

where $\mathcal{R}_n$, $n = 1, \ldots, N$, are N PDE constraints (e.g., for the metamaterial design problem, these are acoustic wave equations with N different incident waves and frequencies), with u_n representing the scattered wave for each incident wave. The cost functional $\mathcal{J}$ consists of a (weighted) linear

combination of N objectives. In the application of optimal design of acoustic metamaterial for cloaking, z represents the deterministic (infinite-dimensional) design field—the bulk modulus of the material—while m represents the manufacturing uncertainty modeled by an infinite-dimensional random field with Gaussian distribution $\mu = \mathcal{N}(0, \mathcal{C})$. The cost functional is given by

$$\mathcal{J}(\{u_n\}_{n=1}^N, z) = \sum_{n=1}^{N} \mathbb{E}_{m\sim\mu}[|u_n|^2] + \beta \mathsf{Var}_{m\sim\mu}[|u_n|^2] + \mathcal{P}(z), \tag{2}$$

where $\mathcal{P}(z)$ is the L_1 penalization of the design z to promote sparsity of the cloak material. We employed Taylor approximation and variance reduction [3] to approximate each scattered wave u_n with respect to the uncertainty m, in which the trace of the covariance-preconditioned Hessian of u_n with respect to m that arises is approximated by the sum of dominant eigenvalues computed by randomized SVD. To solve the infinite-dimensional optimization problem constrained by the N sets of eigenvalue problems and governing state and adjoint acoustic wave equations, we proposed a quasi-Newton algorithm with the Hessian approximated by that of a low-order Taylor approximation. We demonstrated that the cost of our method, measured by number of PDE solves, is independent of the uncertain parameter and design variable dimension. The optimal design of the cloak under uncertainty and the corresponding wavefields with and without the optimal metamaterial cloak are shown in Figure 15, from which we can observe effective cloaking of the obstacle from multiple incident waves. Further details on multiple frequencies, more complex obstacle geometries, and the advantage of design under uncertainty are presented in [2].

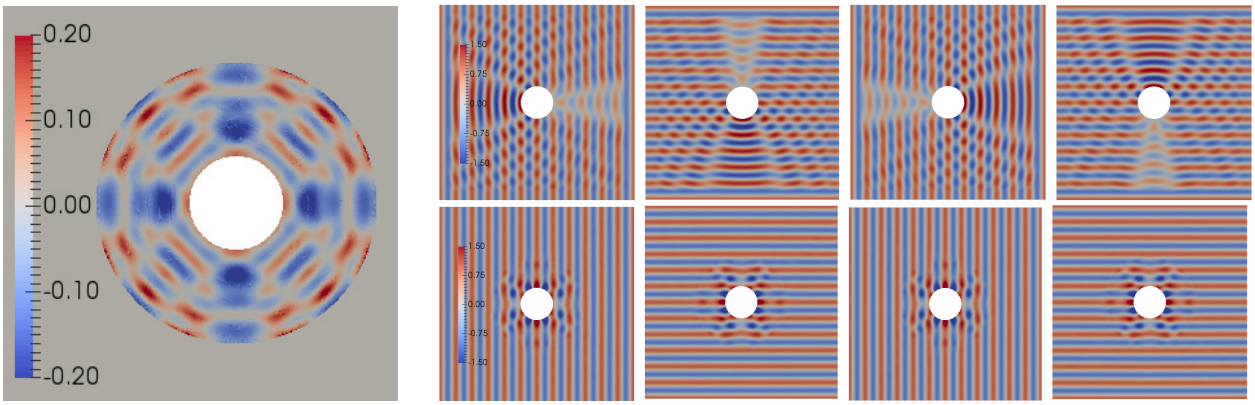


Figure 15: Left: optimal cloak obtained by multiobjective optimal design (minimizing scattered waves with four incident angles) under uncertainty. Right: the real part of the total wave without (top) and with (bottom) the presence of cloak.

We also studied chance-constrained optimization under uncertainty in the abstract formulation

$$\min_z \mathcal{J}(u, z) \quad \text{subject to } \mathcal{R}(u, m, z) = 0 \text{ and } P(f(u) \leq 0) > \alpha, \tag{3}$$

where f denotes a (chance) functional of the state u, depending implicitly on m and z through the PDE $\mathcal{R}(u, m, z) = 0$, and P denotes the probability of $f \leq 0$ with respect to the random field m. To address the discontinuous inequality chance constraint in an infinite-dimensional random parameter space, we proposed an adaptive combination of (i) smooth approximation of the discontinuous inequality constraint, (ii) quadratic approximation of the objective functional and the chance functional, (iii) low-rank approximation of the quadratic term of the chance functional, (iv) sample average approximation of the chance probability, and (v) an external penalization of the inequality constraint by a differentiable penalty function. We demonstated the accuracy, convergence, and efficiency of the proposed adaptive algorithm in a chance-constrained optimal control of subsurface flow problem [1].

References

[1] P. Chen and O. Ghattas, *Taylor approximation for chance-constrained optimization under uncertainty*, in preparation, (2019).

[2] P. Chen, M. R. Haberman, and O. Ghattas, *Optimal design of acoustic cloak under uncertainty*, preprint, (2019).

[3] P. Chen, U. Villa, and O. Ghattas, *Taylor approximation and variance reduction for PDE-constrained optimal control under uncertainty*, Journal of Computational Physics, 385 (2019), pp. 163–186.

[4] P. Chen, K. Wu, J. Chen, T. O'Leary-Roseberry, and O. Ghattas, *Projected Stein variational Newton: A fast and scalable Bayesian inference method in high dimensions*, arXiv preprint arXiv:1901.08659, to appear in Advances in Neural Information Processing Systems, (2019).

[5] R. Choksi and X. Ren, *On the derivation of a density functional theory for microphase separation of diblock copolymers*, Journal of Statistical Physics, 113 (2003), pp. 151–176.

[6] J. Daws, A. Petrosyan, H. Tran, and C. G. Webster, *A Weighted ℓ_1-minimization approach for wavelet reconstruction of signals and images*, IEEE Transactions on Signal Processing, (2019). Submitted (arXiv:1909.07270).

[7] J. Daws and C. G. Webster, *A polynomial-based approach for architectural design and learning with deep neural networks*, in Advances in Neural Information Processing Systems 32, Curran Associates, Inc., 2019. Submitted (arXiv:1905.10457).

[8] A. Dereventsov, A. Petrosyan, and C. G. Webster, *Greedy shallow networks: A new approach for constructing and training neural networks*, in Advances in Neural Information Processing Systems 32, Curran Associates, Inc., 2019. Submitted (arXiv:1905.06448).

[9] A. Dereventsov and C. G. Webster, *The Natural Greedy Algorithm for reduced bases in Banach spaces*, Foundations of Computational Mathematics, (2019). Submitted (arXiv:1905.06448).

[10] G. Detommaso, T. Cui, Y. Marzouk, A. Spantini, and R. Scheichl, *A stein variational Newton method*, in Advances in Neural Information Processing Systems, 2018, pp. 9187–9197.

[11] N. Dexter, H. Tran, and C. G. Webster, *A mixed ℓ_1 regularization approach for sparse simultaneous approximation of parameterized PDEs*, ESAIM: Mathematical Modelling and Numerical Analysis, (2019). Accepted (arXiv:1812.06174).

[12] ——, *Reconstructing high-dimensional Hilbert-valued functions via compressed sensing*, IEEE Signal Processing, (2019). Accepted (arXiv:1905.05853).

[13] C. Eck, *Homogenization of a phase field model for binary mixtures*, Multiscale Modeling & Simulation, 3 (2005), pp. 1–27.

[14] G. H. Fredrickson, V. Ganesan, and F. Drolet, *Field-theoretic computer simulation methods for polymers and complex fluids*, Macromolecules, 35 (2002), pp. 16–39.

[15] M. Gunzburger, M. Schneier, C. G. Webster, and G. Zhang, *An improved discrete least-squares/reduced-basis method for parameterized elliptic PDEs*, Journal of Scientific Computing, 86 (2018), pp. 1–16.

[16] G. Kabacaoğlu and G. Biros, *Machine learning acceleration of simulations of stokesian suspensions*, Physical Review E, 99 (2019), p. 063313.

[17] J. Lei and P. Bickel, *A moment matching ensemble filter for nonlinear non-gaussian data assimilation*, Monthly Weather Review, 139 (2011), pp. 3964–3973.

[18] A. Petrosyan, H. Tran, and C. G. Webster, *Reconstruction of jointly sparse vectors via manifold optimization*, Applied Numerical Mathematics, 144 (2019), pp. 140–150.

[19] B. Radhakrishnan, S. B. Gorti, J. A. Turner, R. Acharya, J. A. Sharon, A. Staroselsky, and T. El-Wardany, *Phase field simulations of microstructure evolution in in718 using a surrogate nifenb alloy during laser powder bed fusion*, Metals, 9 (2018).

[20] V. Reshniak and C. G. Webster, *Robust learning with implicit residual networks*, in Advances in Neural Information Processing Systems 32, Curran Associates, Inc., 2019. Submitted (arXiv:1905.10479).

[21] A. Spantini, R. Baptista, and Y. Marzouk, *Coupling techniques for nonlinear ensemble filtering*, arXiv preprint arXiv:1907.00389, (2019).

[22] K. Wu, P. Chen, and O. Ghattas, *A Stein variational Newton method for expected information gain in Bayesian optimal experimental design*, in preparation, (2019).

[23] X. Xie, F. Bao, T. Maier, and C. G. Webster, *Analytic continuation of noisy data using adams bashforth resnet*, in Advances in Neural Information Processing Systems 32, Curran Associates, Inc., 2019. Submitted (arXiv:1905.10457).

[24] X. Xie, F. Bao, and C. G. Webster, *Evolve filter stabilization reduced-order model for stochastic Burgers equation*, Fluids, 3 (2018), pp. 84–12.

[25] X. Xie, G. Zhang, and C. G. Webster, *Non-intrusive inference reduced order model for fluids using deep multistep neural network*, Mathematics, 7 (2019), pp. 1–15.

[26] G. Zhao, X. Qian, B.-J. Yoon, F. Alexander, and E. Dougherty, *Model-based robust filtering and experimental design for stochastic differential equation systems*, IEEE Transactions on Signal Processing, submitted (2019).

[27] A. Zollanvari and E. Dougherty, *Optimal bayesian classification with autoregressive data dependency*, IEEE Transactions on Signal Processing, 67 (2019), pp. 3073–3086.

A. Organizational chart

DRIVING SCIENTIFIC APPLICATION AREA: ADVANCED MANUFACTURING & MATERIALS	
additive manufacturing testbed (Turner)	materials self-assembly testbed (Alexander and Oden)
INTEGRATIVE RESEARCH THRUSTS	
Thrust 1: Learning predictive models via Bayesian inference & optimization (Webster & Willcox)	**Thrust 2: Optimizing** experiments, processes, & designs under uncertainty (Alexander & Ghattas)

RESEARCH SUB-THRUSTS							
large-scale Bayesian inference (Marzouk)	predictive multiscale models & inadequacy (Moser)	learning from scientific data (Webster)	low-dimensional & reduced modeling (Willcox)	multifidelity methods for OUU (Gunzburger)	large-scale Bayesian OED (Ghattas)	optimal operator design (Dougherty)	optimal control under uncertainty (Biros)

Table 1: AEOLUS organizational chart. Research is organized under two integrative research thrusts, each featuring four sub-thrusts that aim to overcome outer loop challenges in learning and optimization for complex uncertain models. Specific application to advanced manufacturing and materials systems will be carried out under two application testbeds.

B. Work breakdown

The following page lists all AEOLUS personnel during Year 1 of the center (September 2018–September 2019), their positions and institutional affiliations, and the percentage of their time dedicated to each one of the major AEOLUS research tasks.

AEOLUS WORK BREAKDOWN STRUCTURE

	Additive manufacturing	Materials self-assembly	Inference, inverse problems & learning from data	Low-dimensional modeling & reduced models	Optimization under uncertainty	Optimal experimental design	Predictive multiscale models & inadequacy	Administrative tasks
George Biros, UT co-PI	20		20	30	20			10
Yuanxun Bao, UT postdoc	60			20	20			
Will Ruys, UT PhD student			80	20				
Gokberk Kabacaoglu, UT PhD student				80	20			
Omar Ghattas, UT PI and center co-director	10	10	15	15	15	15		20
Peng Chen, UT research associate	10	10	20	20	20	20		
Nick Alger, UT postdoc			50	50				
Ilona Ambartsumyan, UT postdoc			40	40			20	
Josh Chen, UT PhD student		20	30	30	20			
Dingcheng Luo, UT PhD student		25		25	50			
Tom O'Leary Roseberry, UT PhD student			40	40		20		
Keyi Wu, UT PhD student			40	20		40		
Robert Moser, UT co-PI			50				50	
Todd Oliver, UT research scientist			50				50	
Teresa Portone, UT PhD student							100	
J. Tinsley Oden, UT co-PI	10	40	5		15		30	
Danial Faghihi, UT research associate, now assistant professor, U. of Buffalo	25	20	10	5		25	15	
Lianghao Cao, UT PhD student	10	70			20			
Prashant Jha, postdoc	10	70	10			10		
Karen Willcox, UT PI and center co-director	10		20	40		10		20
Parisa Khodabakhshi, UT postdoc	30			70				
Shane McQuarrie, UT PhD student			10	90				
Youssef Marzouk, MIT co-PI			60			30		10
Daniele Bigoni, MIT research scientist			80	20				
Jakob Zech, MIT postdoc			100					
Ricardo Baptista, MIT PhD student			70			30		
Francis Alexander, BNL PI		25	25			25	25	
Gyorgy Matyasfalvi, BNL postdoc (now at Princeton)								
Anthony DeGennaro, BNL staff scientist		30		30		10	30	
Clayton Webster, ORNL PI		20	20	20	20	10		10
Max Gunzburger, ORNL co-PI		30	10	30	30			
Olena Burkovska, Householder Fellow, ORNL		40	20	30	10			
Anton Dereventsov, ORNL postdoc		40	20	30	10			
Armenak Petrosyan, ORNL postdoc			50	50				
Viktor Reshniak, ORNL postdoc			50		50			
Joe Daws, UTK PhD student			30	30	10	30		
Edward Mitchell, UTK PhD student				50	50			
Guannan Zhang, Senior Scientist, ORNL	10		30	30	10	20		
John Turner, ORNL co-PI	50							50
Balasubramaniam Radhakrishnan, ORNL senior scientist	100							
Steven DeWitt, ORNL staff scientist	100							
Edward Dougherty, TAMU PI			25	15	30	30		
Xiaoning Qian, TAMU associate professor			25	25	25	25		
Byung-Jun Yoon, TAMU associate professor			30		35	35		
Shahin Boluki, TAMU PhD student			50		50			
Omar Maddouri, TAMU PhD student			50		50			
Guang Zhao, TAMU PhD student			50			50		

C. AEOLUS publications, September 2018 – September 2019

C.1. Publications appeared or accepted

1. P. Chen, K. Wu, J. Chen, T. OLeary-Roseberry, O. Ghattas, Projected Stein variational Newton: A fast and scalable Bayesian inference method in high dimensions, NeurIPS 2019. https://arxiv.org/abs/1901.08659

2. N. Dexter, H. Tran, and C. G. Webster. Reconstructing high-dimensional Hilbert-valued functions via compressed sensing. IEEE Signal Processing, 2019. Accepted (arXiv:1905.05853).

3. N. Dexter, H. Tran, and C. G. Webster. A mixed l1 regularization approach for sparse simultaneous approximation of parameterized PDEs. ESAIM: Mathematical Modelling and Numerical Analysis, 2019. Accepted (arXiv:1812.06174).

4. Fritz, Marvin; Lima, E. A. B. F.; Nikolic, Vanja; Oden, J. T.; and Wohlmuth, B. Local and Nonlocal Phase-Field Models of Tumor Growth and Invasion Due to ECM Degradation,M3AS (Mathematical Models and Methods in Applied Sciences) Accepted. To Appear, 2019.

5. Amir Gholami, Kurt Keutzer, and George Biros, ANODE: Unconditionally Accurate Memory-Efficient Gradients for Neural ODEs , Proceedings of the International Joint Conferences on Artificial Intelligence , Macao, China (19% acceptance rate), August 2019, https://arxiv.org/abs/1902.10298

6. Amir Gholami, George Biros et al, ANODEV2: A Coupled Neural ODE Framework, to appear in NeurIPS 2019, Vancouver, December 2019 (Acceptance rate 21%), https://arxiv.org/abs/1906.04596

7. G. Kabacaolu and G. Biros, Machine learning acceleration of simulations of Stokesian suspensions, Physical Review E 99(6), 2019 https://link.aps.org/doi/10.1103/PhysRevE.99.063313

8. M. McKerns, F. J. Alexander, K. Hickmann, T. J. Sullivan, and D. Vaughan, Optimal Bounds on Nonlinear Partial Differential Equations in Model Certification, Validation, and Experiment Design, in Advanced Analysis Solutions for Leading Experimental Techniques (2019 to appear)

9. A. Petrosyan, H. Tran, and C. G. Webster. Reconstruction of jointly sparse vectors via manifold optimization. Applied Numerical Mathematics, 144:140150, 2019.

10. K. G. Reyes and F. J. Alexander, Autonomous Experimental Design and Execution, in Advanced Analysis Solutions for Leading Experimental Techniques (2019 to appear)

11. Scarabosio, Laura; Faghihi, Danial; Wohlmuth, B.; Oden, J.T. Goal-Oriented Adaptive Modeling of Random Heterogeneous Media and Model-Based Multilevel Monte Carlo Methods. Computers & Mathematics with Applications, Accepted; To Appear, 2019.

12. X. Xie, G. Zhang, and C. G. Webster. Non-intrusive inference reduced order model for fluids using deep multistep neural network. Mathematics, 7(8):115, 2019.

13. Chenhan Yu, Severin Reiz, and George Biros, Distributed O(N) Linear Solver for Dense Symmetric Hierarchical Semi-Separable Matrices, IEEE 13th International Symposium on Embedded Multicore/Many-core Systems-on-Chip, Nanyang Technological University, Singapore, October 2019

14. Zollanvari, A., and E. R. Dougherty, Optimal Bayesian Classification with Autoregressive Data Dependency, IEEE Transactions on Signal Processing, 67(12). 3073-3086, 2019.

C.2. Publications submitted

1. I. Ambartsumyan, W. Boukaram, T. Bui-Thanh, O. Ghattas, D. Keyes, G. Stadler, G. Turkiyyah, and S. Zampini, Hierarchical Matrix Approximations of Hessians Arising in Inverse Problems Governed by PDEs, submitted, 2019.

2. Bigoni, D., Zahm, O., Spantini, A., Marzouk, Y.M. Greedy inference with layers of lazy maps. Submitted (2019). arXiv:1906.00031.

3. P. Chen and O. Ghattas, Sparse polynomial approximations for affine parametric saddle point problems, submitted. https://arxiv.org/abs/1809.10251

4. P. Chen and O. Ghattas, Sparse polynomial approximation for optimal control problems constrained by elliptic PDEs with lognormal random coefficients, submitted. https://arxiv.org/abs/1903.05547

5. J. Daws and C. G. Webster. A polynomial-based approach for architectural design and learning with deep neural networks, Submitted (arXiv:1905.10457), 2019.

6. A. Dereventsov and C. G. Webster. The Natural Greedy Algorithm for reduced bases in Banach spaces. Foundations of Computational Mathematics, 2019. Submitted (arXiv:1905.06448).

7. A. Dereventsov, A. Petrosyan, and C. G. Webster. Greedy shallow networks: A new approach for constructing and training neural networks. Submitted (arXiv:1905.06448), 2019.

8. Khristenko, Ustim; Constantinescu, Andrei; Le Tallec, Patrick; Oden, J Tinsley; and Wohlmuth, Barbara. A Statistical Framework for generating microstructures of two-phase random materials: application to fatigue Analysis, arXiv preprint arXiv:1907.02412, July 2019 (in review, Multiscale Modeling and Simulation: A SIAM Interdisciplinary Journal).

9. Musolas, A., Massart, E., Hendrickx, J., Absil, P.-A., Marzouk, Y.M. Low-rank multi-parametric covariance estimation. Submitted (2019).

10. T. OLeary-Roseberry, N. Alger, O. Ghattas, Inexact Newton Methods for Stochastic Non-Convex Optimization with Applications to Neural Network Training, submitted. https://arxiv.org/abs/1905.06738

11. V. Reshniak and C. G. Webster. Robust learning with implicit residual networks. Submitted (arXiv:1905.10479), 2019.

12. Spantini, A., Baptista, R., Marzouk, Y.M. Coupling techniques for nonlinear ensemble filtering. Submitted (2019). arXiv:1907.00389

13. Swischuk, R., Kramer, B., Huang, C. and Willcox, K. , Learning physics-based reduced-order models for a single-injector combustion process. Oden Institute Report 19-13. Submitted to AIAA Journal.

14. Tong, X. T., Morzfeld, M., Marzouk, Y.M. MALA-within-Gibbs samplers for high-dimensional distributions with sparse conditional structure. Submitted (2019). arXiv:1908.09429

15. U. Villa, N. Petra, and O. Ghattas, hIPPYlib: An Extensible Software Framework for Large-Scale Inverse Problems Governed by PDEs; Part I: Deterministic Inversion and Linearized Bayesian Inference, submitted, 2019. http://arxiv.org/abs/1909.03948

16. Wang, Z., Cui, T., Bardsley, J., Marzouk, Y.M. Scalable optimization-based sampling on function space. Submitted (2019). arXiv:1903.00870

17. X. Xie, F. Bao, T. Maier, and C. G. Webster. Analytic continuation of noisy data using adams bashforth resnet. Submitted (arXiv:1905.10457).

18. Zhao, G., Qian, X., Yoon, B-J., Alexander, F. J., and E. R. Dougherty, Model-based Robust Filtering and Experimental Design for Stochastic Differential Equation Systems, IEEE Transactions on Signal Processing (in review).

D. AEOLUS invited presentations, September 2018 – October 2019

D.1. Major keynote and plenary lectures at international conferences and other distinguished lectures

- Willcox gives a keynote invited talk "Towards efficient multifidelity modeling for engineering design under uncertainty: From model reduction to scientific machine learning" at the European Numerical Mathematics and Advanced Applications (ENUMATH) Conference, Egmond aan Zee, The Netherlands, October 2019.

- Willcox gives the Kalman Lecture on "Predictive data science for physical systems: From model reduction to scientific machine learning" at the University of Potsdam, Germany, September 2019.

- Willcox gives the opening plenary talk on "From nonlinear partial differential equations to low-dimensional models: Physics-based model reduction" at the 15th U.S. National Congress on Computational Mechanics, July 2019.

- Ghattas gives a keynote invited talk on "Large-scale stochastic PDE-constrained optimization" at the International Congress on Industrial and Applied Mathematics (ICIAM 2019), July 2019.

- Willcox gives a keynote invited talk on "Predictive data science for physical systems: From model reduction to scientific machine learning" at the International Congress on Industrial and Applied Mathematics (ICIAM 2019), July 2019.

- Marzouk gives an invited plenary lecture on "Sampling via Transport: Preconditioning and Low-Dimensional Structure" at the 12th International Conference on Monte Carlo Methods and Applications (MCM2019, http://www.mcm2019.unsw.edu.au/), Sydney, Australia, July 2019.

- George Biros gives the Argyris Lecture at the University of Stuttgart titled "Towards direct numerical simulation of blood flow in microcirculation," June 2019.

- Ghattas gives the Oden Lecture at the XVIth Conference on the Mathematics of Finite Elements and Applications (MAFELAP 2019), London, UK, June 2019.

- Willcox gives a keynote talk "Predictive Data Science for Physical Systems: From Model Reduction to Scientific Machine Learning" at the New York Scientific Data Summit, New York, NY, June 2019.

- Clayton Webster presents a keynote lecture on "Learning high-dimensional systems from incomplete data by optimal non-linear approximations," Isaac Newton Institute for Mathematical Sciences, University of Cambridge, Cambridge, United Kingdom, April 2019.

- Willcox gives the Charlemagne Distinguished Lecture "Projection-based model reduction: Formulations for Physics-based Machine Learning" at RWTH Aachen, April 2019.

- Ghattas gives a plenary talk entitled "Physics-Based Learning of Complex Models from Large-Scale Data: A Scalable Bayesian Inversion Approach," at the International Conference on Big Data in the Geosciences, China University of Geosciences, Wuhan, China, January 2019.

- Webster presents a Distinguished Lecture at the Pacific Institute for the Mathematical Sciences (PIMS) , IRMACS Theatre, Simon Fraser University, Burnaby, BC, December 2018.

- Ghattas gives the Charlemagne Distinguished Lecture at RWTH Aachen University, Germany. The lecture was entitled "Large-scale Bayesian inversion with applications to flow of the Antarctic ice sheet," October 2018.

D.2. Other conference and workshop invited talks

October 2019

- Biros gives two invited talks at the European Numerical Mathematics and Advanced Applications Conference The first talk is "PDE-constrained optimization for importance sampling of rare events"; and the second talk is "Reduced order models for the simulation of microfluidic devices for biological fluids"

- Marzouk gives an invited lecture at RWTH Aachen University.

September 2019

- Tinsley Oden gives the Keynote lecture at cellMath, a workshop on tumor growth modeling held at the Technical University of Munich (TUM) on September 10, 2019, on "A Review of Multiscale Models of Tumor Growth".

- Biros gives an invited talk titled "Machine learning-accelerated simulations of complex fluids", at the 10th International Workshop on Meshfree Methods for Partial Differential Equations

July 2019

- Tinsley Oden gives an invited Keynote lecture to Annual Meeting and Summer School 2019; IRTG 2379 in Austin on "Phase-Field Models of Phase Change of Complex Systems: Block Copolymers and Growth of Vascular Tumors", July, 2019.

- Biros gives an invited talk on "BIMC: The Bayesian Inverse Monte Carlo method for goal-oriented uncertainty quantification", at the Applied Inverse Conference in Grenoble, France.

- Ghattas gives an invited talk entitled "Fast methods for Bayesian inverse problems governed by PDE forward models with random coefficient fields," at the Applied Inverse Problems Conference in Grenoble, France.

- Francis J. Alexander gives a talk On the Connection between Optimal Uncertainty Quantification and the Mean Objective Cost of Uncertainty, presented at US National Congress on Computational Mechanics (USNCCM), Austin, Texas, July 29, 2019

- Chen gives an invited talk "Optimal design of acoustic cloak under uncertainty" at the 15th U.S. National Congress on Computational Mechanics, Austin, US

- Chen gives an invited talk "Stein variational methods for Bayesian optimal experimental design" at the Applied Inverse Conference in Grenoble, France.

- Chen gives an invited talk "Stein variational, reduced basis Bayesian inversion" at the International Congress on Industrial and Applied Mathematics (ICIAM 2019)

- Chen gives an invited talk "A scalable method for PDE-constrained optimization under high-dimensional uncertainty" at the International Congress on Industrial and Applied Mathematics (ICIAM 2019)

June 2019

- George Biros gives an invited talk titled "Reduced order models for nonlinear boundary value problems with moving interfaces" at the Sandia National Labs (NM).

May 2019

- Ghattas gives an invited talk entitled "Large-scale Optimal Experimental Design for Bayesian Inverse Problems" at the 5th International Conference on Design of Experiments (ICODOE 2019) in Memphis.

April 2019

- AEOLUS graduate student Teresa Portone gives an invited talk "A Stochastic Operator Representation of Model-Form Uncertainty" at Rising Stars in Computational & Data Sciences 2019.

- Marzouk gives invited talks at INRIA, Laboratoire Jean Kuntzmann. Grenoble, France; Schlumberger-Doll Research. Cambridge, MA; Auburn University, Department of Mathematics and Statistics, Colloquium. Auburn, AL; Worcester Polytechnic Institute, Department of Mathematical Sciences, Colloquium. Worcester, MA.

- Willcox gives the Keynote Talk "Data to Decisions: Computational Methods for the Next Generation of Engineering Systems" at the SIAM Central Valley Regional Student Conference, University of California, Merced, April 2019.

- Moser gives an invited lecture entitled "Making Reliable Computational Predictions: Is it Possible" at Computational & Data Enabled Science Days at SUNY Buffalo.

March 2019

- AEOLUS graduate student Teresa Portone gives an invited talk "A Stochastic Operator Approach to Representing Model-Form Uncertainty" at Sandia National Laboratories, Albuquerque, NM.

- Chen gives an invited seminar talk "Hessian in action for high-dimensional model reduction, stochastic optimization, and Bayesian inversion" at Department of Mathematics, Peking University, China

February 2019

- Ghattas gives an invited talk entitled "Large-scale Optimal Experimental Design for Bayesian Nonlinear Inverse Problems," at the SIAM Conference on Computational Science and Engineering, (CSE19) in Spokane, WA

- Willcox gives a Keynote Talk "Projection-based model reduction: Formulations for Physics-based Machine Learning" at the LANL Workshop on Machine Learning for Computational Fluid and Solid Dynamics, February 2019.

- Chen gives an invited talk "Towards Breaking the Curse of Dimensionality for PDE-constrained Optimization under High-dimensional Uncertainty" at SIAM Conference on Computational Science and Engineering
- AEOLUS graduate student Joshua Chen gives a talk "Dimension Adaptive Sparse Quadrature and Sparse Polynomial Parametrized Transport Maps for High Dimensional Bayesian Integration" at SIAM Conference on Computational Science and Engineering
- AEOLUS postdoc Ilona Ambartsumyan presented a poster "An Edge-preserving Method for Joint Infinitedimensional Bayesian Inversion" at SIAM Conference on Computational Science and Engineering

January 2019

- The AEOLUS team presented four posters at the Department of Energy ASCR PI Meeting.
- Clayton Webster presents the Tianyuan Distinguished lecture on Learning high-dimensional systems from incomplete data by nonlinear approximation and deep networks; School of Mathematics, Jilin University, Changchun, China.
- Marzouk and Webster are invited speakers at the Johns Hopkins University, USACM workshop on Uncertainty Quantification in Computational Solid and Structural Materials Modeling, Baltimore, MD.

December 2018

- Willcox visits LLNL and SNL Livermore to discuss research collaborations, and gives a talk "Lift & Learn: From nonlinear PDEs to low-dimensional polynomial approximations."
- Marzouk gives an invited talk at the University of Michigan Institute for Computational Discovery and Engineering, Ann Arbor, MI.

October 2018

- Ghattas gives an invited talk entitled "Learning from data through the lens of models: Scalable algorithms for Bayesian inverse problems," at the Workshop on HPC and Data Science for Scientific Discovery, at the Institute for Pure and Applied Mathematics (IPAM), UCLA, Los Angeles, CA
- Marzouk gives a colloquium talk at Duke University, Department of Civil and Environmental Engineering.
- Marzouk gives an invited talk at the Finnish Meteorological Institute, Helsinki, Finland.
- Chen gives an invited talk "Sparse quadrature for high-dimensional Bayesian inverse problems" at the 4th annual meeting of SIAM central states section, Oklahoma
- AEOLUS postdoc Ilona Ambartsumyan gives a talk "Bayesian inversion of fault properties in two-phase flow in fractured media" at the Annual Meeting of the SIAM Texas-Louisiana Section, LSU, Baton Rouge, Louisiana

September 2018

- Ghattas gives an invited talk entitled "Scalable algorithms for optimal training data for Bayesian inference of large scale models," at the Workshop on Big Data Meets Large-Scale Computing, Institute for Pure and Applied Mathematics (IPAM), UCLA, Los Angeles, CA

E. Other significant dissemination

Besides the traditional avenues for dissemination of research (publications and presentations), the AEOLUS team has been actively sharing the results of its research with the broader scientific community via other mechanisms, including public software releases, teaching of summer schools and short courses, development of new semester-long courses, co-organization of topical workshops and conference, and training of Ph.D. students and postdocs who move on to academic and laboratory positions.

E.1. Software

None to report yet.

E.2. Short courses/summer schools taught and long courses developed

- Willcox gave invited tutorials on "Multifidelity Models and Methods: Fusing models and data to achieve efficient design, optimization, and uncertainty quantification" and "Model Order Reduction: Approximate yet accurate surrogates for large-scale simulation" as part of the thematic program on Science at Extreme Scales: Where Big Data Meeting Large-Scale Computing, Institute for Pure and Applied Mathematics, September 2018.
- Chen lectured a short course "Approximation of high-dimensional parametric PDEs" at Department of Mathematics, Peking University, China, March 2019.
- Marzouk is an invited lecturer at the MASCOT-NUM Research School on Uncertainty in Scientific Computing, 2327 September 2019. Frjus, France.
- Chen is one main lecturer in the Cargese summer school on The Mathematics for Climate and Environment, Cargese, France, September 2019.
- AEOLUS co-PI Moser gave an invited talk titled Uncertainty, Validation and Prediction with Computational Models at 1st Computational Physics School for Fusion Research (https://sites.google.com/view/mit-psfc-cps-fr2019/home), CPS-FR, August 2019.
- AEOLUS co-PI Marzouk gave an invited talk titled Bayesian modeling and computation for inverse problems at 1st Computational Physics School for Fusion Research (https://sites.google.com/view/mit-psfc-cps-fr2019/home), CPS-FR, August 2019.
- AEOLUS co-PI Moser gave an invited talk title "Uncertainty, Validation & Prediction with Computational Models" at 1st Computational Physics School for Fusion Research (https://sites.google.com/view/mit-psfc-cps-fr2019/home), CPS-FR, August 2019.

E.3. Books published or in manuscript form

None to report yet.

E.4. Workshops and conferences co-organized

Team members served as co-chairs or co-organizers of the following workshops and conferences on AEOLUS-related themes.

- AEOLUS co-PI Clayton Webster served as the Co-Chair of Surrogate models for UQ in complex systems; Isaac Newton Institute for Mathematical Sciences, University of Cambridge, Cambridge, United Kingdom, November 2018.
- The Oden Institute and Sandia National Laboratories co-host Rising Stars in Computational & Data Sciences in April 2019, bringing 32 outstanding female PhD students and postdocs to Austin for a two-day intensive research workshop.

- AEOLUS co-PI Clayton Webster served on the Organizing Committee for the Workshop on Sparse Grid and Applications; Institute for Advanced Study, Technische Universitat Munchen, Munich, Germany, July 2019.

- Webster served as the Organizing Chair of The 2019 SIAM SEAS Annual Conference; The University of Tennessee, Knoxville, TN. Conference website: https://www.math.utk.edu/SIAM-SEAS/

- Webster served on the Scientific Committee of the 8th workshop on high-dimensional approximation; Seminar for Applied Mathematics, ETH Zurich, Zurich, Switzerland.

- Ghattas is co-organizing the Workshop on Big Data, Data Assimilation, and Uncertainty Quantification as part of the Trimester on Mathematics of Climate and the Environment, Institut Henri Poincare, Paris, France, November 12-15, 2019. Ghattas is co-teaching the mini-course Big data, data assimilation, and uncertainty quantification immediately preceding the workshop.

- Marzouk is co-organizing the MIT/Alan Turing Institute/Lloyds Register Foundation workshop on Data-Centric Engineering, to be held Dec 9-12 in Cambridge, MA. Ghattas and Willcox are giving invited talks at the workshop.

- Ghattas is co-organizing the Workshop on Mathematical Modeling in Glaciology at the Banff International Research Station, Banff, Canada, January 1217, 2020.

- Marzouk is co-organizer of the Oberwolfach workshop on Data assimilation: mathematical foundations and applications, April 2020.

- Willcox is co-organizer of the workshop on "Multilevel and multifidelity sampling methods in UQ for PDEs" at Erwin Schrödinger Institute, May 2020.

- Ghattas and Marzouk are co-organizing the Workshop on Mathematical Foundations of Data Assimilation and Inverse Problems, at the Conference on Foundations of Computational Mathematics (FoCM20), Vancouver, Canada, June 15-24, 2020.

E.5. Other dissemination

- Ghattas and former Oden Institute and DiaMonD (MMICC-1) members Tobin Isaac, Noemi Petra, and Georg Stadler received the 2019 SIAM SIAG on Computational Science & Engineering Best Paper Prize (for the period 2015–2018) for the paper "Scalable and Efficient Algorithms for the Propagation of Uncertainty from Data through Inference to Prediction for Large-scale Problems, with Application to Flow of the Antarctic Ice Sheet," published in 2015 in the *Journal of Computational Physics*. Tobin Isaac presented the prize lecture at CSE19.

- Ghattas received the 2019 SIAM Geosciences Career Prize, for "groundbreaking contributions in analysis, methods, algorithms, and software for grand challenge computational problems in geosciences, and for exceptional influence as mentor, educator, and collaborator."

- Ghattas and Marzouk are on the editorial board of the new journal *Foundations of Data Science*, published by the American Institute of Mathematical Sciences.

- Marzouk began a two-year term as program director of the SIAM Activity Group on Mathematical Issues in the Geosciences (SIAG/GS), January 2019.

- Ghattas began a two-year term as Chair of the SIAM Activity Group on Uncertainty Quantification (SIAG/UQ), January 2019
- A dedication ceremony was held September 2019 for the NSF-supported Frontera supercomputer at UT Austins Texas Advanced Computing Center. Ghattas is Chief Scientist, and George Biros and Robert Moser are on the Science Team, for this project. The nearly 50 petaflops Frontera is the fastest university-based supercomputer in the world, and 5th fastest overall.