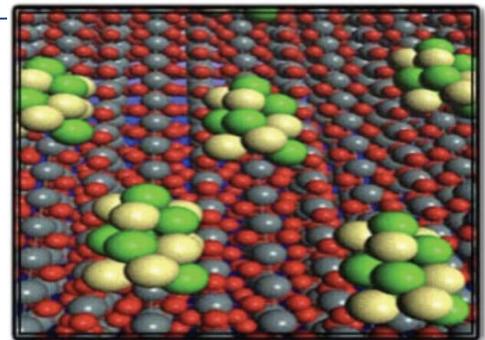
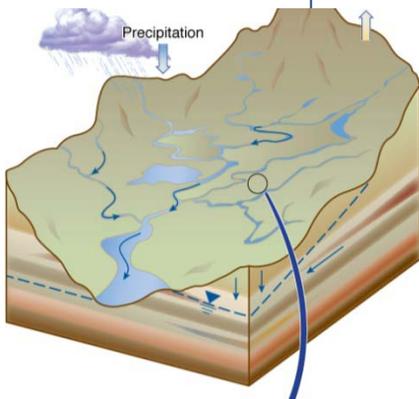
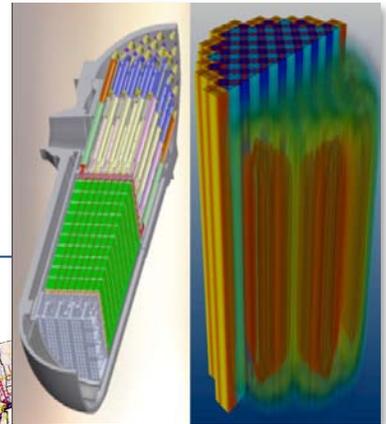
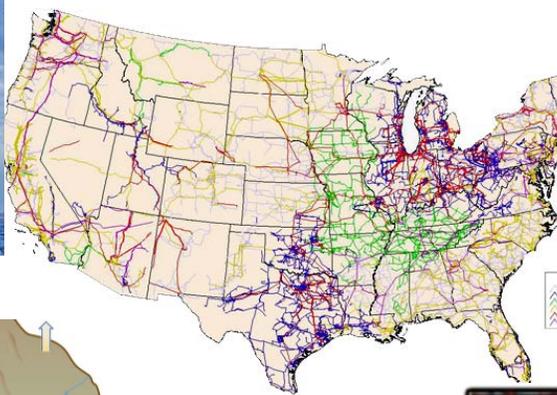


A Multifaceted Mathematical Approach for Complex Systems



Report of the DOE Workshop on Mathematics for the
Analysis, Simulation, and Optimization of Complex Systems

September 13–14, 2011

DISCLAIMER

This document was prepared as an account of work sponsored by the United States Government. While this document is believed to contain correct information, neither the United States Government nor any agency thereof, nor the Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or the Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof or the Regents of the University of California.

Cover images:

Upper left: offshore wind turbine, source: Large-Scale Offshore Wind Power in the United States, NREL/TP-500-40745, September 2010.

Center: electric power grid, source: FEMA.

Upper right: nuclear reactor core, courtesy of A. Larzelere.

Lower left: subsurface model, courtesy of S. Hubbard.

Lower right: catalyst, courtesy of M. Pederson.

This work was supported by the Applied Mathematics Program in the Office of Advanced Scientific Computing Research of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

A Multifaceted Mathematical Approach for Complex Systems

Report of the DOE Workshop on Mathematics for the
Analysis, Simulation, and Optimization of Complex Systems

September 13–14, 2011

Prepared by:

F. Alexander (LANL)
M. Anitescu (ANL)
J. Bell (LBNL)
D. Brown (LBNL)
M. Ferris (U. Wisc.)
M. Luskin (U. Minn.)
S. Mehrotra (Northwestern)
B. Moser (U. Texas)
A. Pinar (SNL)
A. Tartakovsky (PNNL)
K. Willcox (MIT)
S. Wright (U. Wisc.)
V. Zavala (ANL)

Table of Contents

Executive Summary	1
1 Introduction.....	5
1.1 Purpose and Scope	5
1.2 Motivating Problems and Workshop Overview.....	5
1.3 Previous Successes of the DOE Applied Mathematics Program	7
1.4 Aims and Structure of the Report	7
2 Vision of Integration.....	9
2.1 Context.....	9
2.2 The Path Forward for Multifaceted Mathematics	10
3 Research Needs in Topical Areas	12
3.1 Modeling Systems Described by Diverse Phenomena and/or Scales	12
3.2 Optimization Theory and Algorithms	13
3.3 Stochastics and Uncertainty.....	15
3.4 Integrating Data and Modeling	16
4 Combining Technologies to Solve Problems.....	19
5 Case Studies of the Potential Impact of Multifaceted Mathematics	22
5.1 Catalysis Discovery for Next Generation Fuels.....	22
5.2 Power Grid.....	23
5.3 Nuclear Reactor Core Reloading	24
5.4 Remediation	25
6 Conclusions.....	28
Acknowledgment	29
Bibliography	30
Appendix: Workshop Participants	31

A Multifaceted Mathematical Approach for Complex Systems

Executive Summary

The Applied Mathematics Program of the U.S. Department of Energy (DOE) traces its roots back to the early 1950s when the program was founded as part of the Atomic Energy Commission at the request of John von Neumann. The program has a long, rich tradition of developing the mathematical tools needed to address key scientific and technical needs of the department. Researchers supported by the program have made contributions ranging from creation of the mathematical foundation of key areas such as fluid mechanics and optimization to the development of numerical methods used throughout the department to the design of software that is ubiquitous within the computational science community.

The Department of Energy's mission requirements in energy, environment, and national security as well as its role in fostering US leadership in science are driving a need to address increasingly complex scientific and technical challenges. The nation faces a societal imperative, where the nation's energy needs, competitiveness, and security concerns require the scientific community to find answers to key scientific and technological questions. There is a growing sense of urgency to addressing these questions that demands a shift in the way we do science. This shift in focus will require researchers to look for solutions to problems within the context of broader systems. Finding the answers the nation needs for critical DOE applications, such as designing next generation engines and power plants, materials and chemistry by design, developing safe fission reactors, understanding the behavior of the electrical power grid, improving techniques for environmental remediation, and maintaining a safe and secure nuclear stockpile, will require adoption of this type of complex systems perspective.

To understand how applied mathematics could most effectively contribute to solving these types of problems, the DOE Office of Advanced Scientific Computing Research (ASCR) Applied Mathematics Program sponsored a Workshop on Mathematics for the Analysis, Simulation and Optimization of Complex Systems on September 13–14, 2011. The workshop had approximately 50 participants from both the national labs and academia. The specific goal of the workshop was to identify new research areas in applied mathematics that will complement and enhance the existing DOE ASCR Applied Mathematics Program efforts that are needed to address problems associated with complex systems. The workshop began with a plenary panel session that included presentations from Basic Energy Sciences (BES) within the Office of Science and from a number of the DOE Applied Offices, including Energy Efficiency and Renewable Energy, Electricity Delivery and Energy Reliability, Nuclear Energy, and Environmental Management. Each presentation provided an overview of their problem areas and described the major challenges facing their programs.

The presentations served to underscore the fact that meeting DOE's mission requirements will require the ability to model systems of unprecedented complexity. Dealing with these systems will require new types of models that couple a broad array of diverse processes spanning a wide range of spatial and temporal scales. The models will likely combine both deterministic and stochastic elements; they are likely to switch between alternative representations of processes as a function of dynamic properties of the

problems. Furthermore, the goal in studying these types of systems will not be to simply produce a simulation of the system but rather to answer difficult questions regarding design, risk analysis, or optimization, with quantified bounds on the error. Answering these types of questions will require combinations of new mathematical tools from several different areas. Workshop participants were asked to identify the key research directions in applied mathematics that will need to be pursued to address complex systems arising in DOE mission areas. Four topical areas for research were identified:

- modeling systems characterized by diverse phenomena and/or scales
- optimization theory and algorithms
- stochastic systems and systems with uncertainty
- integration of data and modeling.

While advances in each of these research areas would have clear impact on DOE needs, the workshop concluded that developing these tools independently with the hope that they will be able to work together to meet the challenges of complex systems is not the most effective strategy. A more effective approach is to pursue integrated strategies that consider a complex system as a whole and bring together all of the mathematical elements that are needed to answer critical questions about the system. We refer to this strategy as a *multifaceted mathematics* approach. With this type of approach, we bring together not only all of the mathematical tools needed to solve the problem, but we also consider all aspects of the problem-solving process ranging from the mathematical formulation to the development of approximate models to the implementation of solver technologies for emerging computer architectures. This type of holistic view of the problem-solving process will enable the applied mathematics community to find solutions to problems involving complex systems of interest of DOE.

Priority research directions identified by this workshop for multifaceted mathematics to treat complex systems include:

- algorithms and analysis of methodologies for multiscale/multiphysics systems that couple stochastic and deterministic models
- systematic approaches to constructing surrogate or reduced-order models with quantifiable fidelity for optimization, uncertainty quantification, and decision making
- algorithms and analysis for stochastic and robust optimization that address model uncertainties with limited information on parameter distributions, considering alternative models of uncertainty and risk, while addressing nonlinearities and special structures
- approaches for propagation of uncertainty through the different scales in a multiscale model and through the components of a multiphysics model
- improved algorithms for data assimilation and model/data fusion for use in system identification, control, and bounding of model uncertainty.

The type of research activities envisioned for multifaceted mathematics represents a paradigm shift from the traditional single-investigator model for applied mathematics research. Multifaceted mathematics by definition will draw on expertise from a variety of different mathematical subdisciplines; it cannot be pursued using a single-investigator model. Projects would be focused around a research team within an organizational structure that supports mechanisms for making linkages to single-investigator research that is relevant to the larger project. In addition, multifaceted mathematics will have a different model for how

mathematicians engage with domain experts. Multifaceted mathematics will require a long-term engagement with domain experts and mathematicians working together from problem inception to final solution. This type of paradigm shift will enable applied mathematics to make a significant impact on scientific and engineering challenges that face DOE.

1 Introduction

1.1 Purpose and Scope

DOE's mission requirements in energy, environment, and national security, as well as its role in fostering US leadership in science, are driving a need to address increasingly complex scientific and technical challenges [1]. Challenges of importance to DOE include designing next generation engines [2] and power plants [3], materials and chemistry by design [4, 5], developing safe fission reactors [6], understanding the behavior of the electrical power grid [7], improving techniques for environmental remediation [8, 9], and maintaining a safe and secure nuclear stockpile [10]. These challenges come at a critical time for our nation. We face a societal imperative, where the nation's energy needs, competitiveness, and security concerns require attention to end-goal considerations for the scientific enterprise in general and for the DOE mission areas in particular.

Applied mathematics has an important role to play in developing the tools needed for the analysis, simulation, and optimization of these types of complex problems. These efforts require the development of the mathematical foundations for scientific discovery, engineering design, and risk analysis based on a sound integrated approach for the understanding of complex systems. However, maximizing the impact of applied mathematics on these challenges requires a novel perspective on approaching the mathematical enterprise. Previous reports [11–14] that have surveyed the DOE's research needs in applied mathematics have played a key role in defining research directions with the community. Although these reports have had significant impact, accurately assessing current research needs requires an evaluation of today's challenges against the backdrop of recent advances in applied mathematics and computing.

To address these needs, the DOE Applied Mathematics Program sponsored a Workshop on Mathematics for the Analysis, Simulation and Optimization of Complex Systems on September 13–14, 2011. The workshop had approximately 50 participants from both the national labs and academia. The goal of the workshop was to identify new research areas in applied mathematics that will complement and enhance the existing DOE ASCR Applied Mathematics Program efforts that are needed to address problems associated with complex systems. This report describes recommendations from the workshop and subsequent analysis of the workshop findings by the organizing committee.

1.2 Motivating Problems and Workshop Overview

To frame the classes of problems of importance to DOE, the workshop began with a plenary panel session that included presentations from representatives of a range of DOE offices. The panel included presentations from Basic Energy Sciences (BES) within the Office of Science and from a number of the DOE Applied Offices, including Energy Efficiency and Renewable Energy, Electricity Delivery and Energy Reliability, Nuclear Energy, and Environmental Management. The BES presentation focused on systems arising in a more fundamental science context with an emphasis on the challenges in controlling matter and energy at electronic, atomic, and molecular levels that are needed for the “deterministic design and discovery of new materials and chemical assemblies with novel structures, functions, and properties.” Several presentations discussed issues related to understanding the behavior of the electric power grid and controlling its behavior to prevent cascading failures, and the challenges associated with developing wind

and solar electricity generation capabilities and integrating that capacity into the power grid. Another presentation discussed the modeling requirements for nuclear energy that are needed to address key operating issues associated with existing nuclear reactors and to provide the tools needed to design the next generation of reactors. The environmental management presentation focused on the need to develop modeling and simulation tools for supporting environmental management situations that can provide critical assessment and risk analysis for regulatory compliance.

Although the systems described above arise in widely disparate contexts, they share a number of common characteristics that impact how they can and should be modeled. None of these systems can be analyzed at the required fidelity with monolithic mathematical models. Effective models must be hierarchical and include multiple sub-models that represent different phenomena with vastly differing scales. These models can range from the representation of detailed physical processes to the description of complex engineered systems that must be optimized to obtain desired performance. These hierarchical models will require new approaches for understanding the impact on system behavior of the interplay between sub-models at different physical resolutions or between decision points in an engineered system. Answering key questions about these systems may also require a much tighter coupling between data and simulation, and require new approaches to aggregate information from multiple sources.

In addressing DOE's needs for the understanding of complex systems, another common characteristic is the need to answer a wide range of questions about the system. Models that only focus on predicting system behavior under a specific set of conditions are no longer sufficient to address the important challenges that DOE faces. Instead, an expanded set of questions must be answered such as how reliable are the predictions, how can the behavior of the system be optimized in the face of uncertainties, and what are the inherent risks in the system. Answering these new questions will also require novel mathematical approaches, including methods that exploit large amounts of data from vastly different sources and that treat underlying uncertainties within data and model structure and definition.

Workshop participants were asked to examine these prototypical complex systems and identify key cross-cutting mathematical issues. They were then asked to assess the current state of the art and identify important research directions in applied mathematics required to address complex systems arising in DOE mission areas. Four topical areas for research were identified:

- modeling systems characterized by diverse phenomena and/or scales
- optimization theory and algorithms
- stochastic systems and systems with uncertainty
- integration of data and modeling.

While advances in each of these research areas will have clear impact on DOE needs, the workshop concluded more significantly that the highest potential for impact on DOE challenges would come from combinations of methodologies that exploit synergies among these research areas.

1.3 Previous Successes of the DOE Applied Mathematics Program

The Applied Mathematics Program has a long history, tracing its roots back to the early 1950s, when the program was founded as part of the Atomic Energy Commission at the request of John von Neumann. Researchers supported by the program are responsible for much of the progress in numerical methods for partial differential equations, numerical linear algebra, and optimization, to name a few significant areas.

Many of the numerical methods used to solve partial differential equations today have their origins in the Applied Mathematics Program. In fluid dynamics, fundamental work in the theory of hyperbolic conservation laws and shock waves by Lax, the development of projection methods by Chorin, and the development of adaptive mesh refinement all trace their origins to the program. Level set algorithms for the tracking interfaces and fast multipole methods also have their roots here. Researchers associated with the program played a pivotal role in developing the theoretical foundations of unconstrained optimization and developed MINPACK, the first reliable software for unconstrained minimization, solution of nonlinear systems, and nonlinear least squares problems. Building on a long tradition of excellence in numerical software for linear algebra for broad scientific use, researchers in numerical linear algebra have discovered new scalable methods and made them broadly available to the scientific community in packages such as ScaLAPACK, Trilinos, and PETSc. As an indication of the broad contribution that this work has had to energy science and technologies, in 2011 the core developers of PETSc were awarded the first Ernest Orlando Lawrence Award ever bestowed upon computational scientists.

The Applied Mathematics Program has maintained an excellent balance between its commitment to research with a longer-term perspective and its agility at funding new program areas. For example, recent thrusts within the Applied Mathematics Program have laid the groundwork for multiscale modeling, the analysis of complex distributed systems, and uncertainty quantification. In addition, the program has also made initial inroads into the mathematical analysis of data, looking at the data deluge problem from a mathematical perspective, and understanding challenges that lie at the interface between applied mathematics and computer science on emerging architectures.

1.4 Aims and Structure of the Report

The research directions discussed in this report are intended to build upon and complement the core strengths of the Applied Mathematics Program in differential equations, optimization, and numerical linear algebra, as well as the early results from these newer thrust areas. Our ability to impact complex systems of interest to DOE will not only rely on advances in the topics discussed here but will also need to marshal advances across multiple core mathematical areas.

Section 2 provides an overview of the vision for using the integration of different research topics discussed above to attack complex systems of interest to DOE. Section 3 identifies some of the research tasks within the various topical areas that are needed to support this vision. Section 4 discusses the need for approaches based on holistic problem-solving strategies that establish new integrated mathematical frameworks for scientific discovery, design, optimization, and risk assessment for complex systems. Section 5 illustrates how this approach would work in several potential application areas. Section 6

provides some concluding remarks and discusses some of the sociological issues related to changing how the applied mathematics community addresses the challenges of complex systems.

2 Vision of Integration

2.1 Context

Historically, progress in computational mathematics has been carried out primarily by a careful classification of the problem at hand, followed by a deep dive into the analysis of the problem, which in turn has led to algorithms and finally software. Insofar as the resulting model interacted with other mathematical paradigms through software implementations, this was typically done via a generic interface with few structural attributes. This type of multi-use, modular approach sits at the base of many numerical libraries that have been created over the past decades and are essential workhorses in the landscape of DOE mission-driven computational science. While this point of view has been enormously successful and remains essential to discovery and decision-making, an increased emphasis on multidisciplinary research has also exposed its limitations. One example is the implementation of a high performance code for simulation of physics problems arising in a particular context, such as fluid dynamics, solid mechanics, magnetohydrodynamics, or neutron transport, when the end use may be for optimal design or uncertainty analysis. In this case, many of the functionalities that would make optimization or uncertainty analysis more efficient, such as existence of derivatives or the ability to define random forcing, are absent, thus making it extremely difficult to obtain the desired solution. In addition, an opportunity to advance understanding at the boundary between disciplines in a realistic mission-driven context is lost through this approach.

Moreover, we find ourselves in the midst of two very disruptive technological currents: the computing architecture revolution and the data deluge. The computing architecture revolution is prompted by the power-limited nature of emerging high-end architectures and has two important design consequences. Power constraints make it increasingly difficult to increase microprocessor speed; consequently, increases in computational power need to come from an increase in parallelism, resulting in a rapidly increasing number of cores per processor. The development of low-power, high-concurrency multicore processors leads, in turn, to having computational costs, in terms of both performance and power, be dominated by data movement and storage. As a result, optimizing performance for emerging architectures will require decreasing memory and data movement per achievable floating point operation. Such trends imply that existing algorithms and software will exhibit a dramatically increasing departure from optimal performance. Designing new algorithms that will run near peak performance on power-limited architectures, including those up to the exascale requires (1) a comprehensive analysis of numerical algorithms in light of emerging metrics that emphasize data movement over floating point operations, (2) exploring novel mathematical formulations from a high-level viewpoint that includes all known classes of techniques and combinations of them, and (3) inventing radically new modeling formulations and paradigms that offer application scientists new functionality characterized by maximizing the ratio of computations to data movement.

The second challenge, data analysis, is a key component of the scientific endeavor. However, the combination of exponentially increasing experimental and observational data and advances in scientific computing are fundamentally changing the way we think about data. With the massive amounts of data that are available, it is often no longer feasible to store the data for subsequent analysis; at least some

reduction of the data in a streaming mode is essential for reducing data volume. Furthermore, it is not simply the case that the volume of data has increased dramatically; the complexity of the systems under consideration has led to an increase in the complexity of the data. To address this growing complexity, models are needed that can utilize data realized in real time, in conjunction with aggregate historical data and analytical information about the underlying components of the system. In this setting, it is not only the size of the data but also the nature of the streaming flow of information that provides the challenges to the underlying mathematical models. Conceptual questions that link data of varying forms, such as the integration of sensor data with computational models, must be tackled using rigorous tools to combine, analyze, and deploy information for decision making and scientific discovery.

The advent of these two fundamental technological shifts is occurring at an important societal juncture. There is a growing sense of urgency in DOE's key mission areas [1] that has resulted in an increased focus on end-goal priorities in energy sciences and technology on time scales on the order of five to ten years. The confluence of changing paradigms in computer technology and changing attitudes toward energy security of the nation provides an enormous opportunity for the applied mathematics community to demonstrate the unique role it can play in this dynamic science and technology landscape. Through approaches discussed in this document, applied mathematicians can create a formalism in which to classify, analyze, and refine the models originating from the target areas; identify performance and complexity bottlenecks; remove artificial barriers imposed by improper solution abstraction; and thus open the road for scalability of the computational solution, which in turn accelerates the transition of innovative scientific ideas into applications. However, accomplishing such bold conceptual goals while providing insights and significantly contributing to the solution of essential and complex DOE challenges on moderate time scales requires a novel perspective on approaching the mathematical enterprise.

2.2 The Path Forward for Multifaceted Mathematics

This discussion highlights the tremendously complex problems within DOE's mission areas of energy, environment, and security, as well as the architectural and data-driven technological issues that will govern their computational solution space. These problems pose grand challenges in applied mathematics because they are multifaceted: they can involve complex continuous and discrete phenomena with both stochastic and deterministic elements; they can involve coupled models from multiple scientific disciplines; they require integration of simulation and data; they are often characterized by pervasive uncertainties; and they can also involve multiple stakeholders and decisions at multiple levels. The multifaceted nature of these problems demands a new mode of operation for applied mathematicians, a mode that cuts across the boundaries of traditional development of methods and tools. To contribute towards solving problems of major importance to DOE, we must consider new research directions not just within the mathematical topic areas considered in this report (modeling, optimization, stochastics, and data) but also at the intersections of these areas: *a multifaceted mathematics approach*.

This crosscutting view is not new. For example, PDE-constrained optimization is a field that has arisen at the intersection of modeling and optimization. The development of optimization algorithms that exploit PDE problem structure, efficient PDE solvers, and adjoint capabilities has led to solution of a new class of optimization problems that were previously intractable. However, problems of major importance to

DOE demand new levels of crosscutting research in applied mathematics: research that can define, analyze, and refine new problem classes; and research that can optimally leverage current and emerging computing architectures. This crosscutting view of applied mathematics will lead to fundamental mathematical questions whose answers will pave the way for addressing new application areas.

We believe that now is the time for an increased and consistent emphasis on multifaceted mathematics. Such an endeavor would build on long-range, fundamental research in the DOE Applied Mathematics Program and the technology transfer of that basic mathematical research to important DOE applications accomplished through the SciDAC program and through its embodiment in high-quality mathematical software. It will leverage new concepts developed by the emerging community of DOE-fostered applied mathematicians, focusing on multiphysics and multiscale methods for bridging space and time scales, for simulation of rare events, and for phase space sampling. The pursuit of novel, integrated problem-solving strategies for complex systems will also contribute to the success of extreme-scale computing by developing new methodologies designed for emerging computer architectures.

Multifaceted mathematics provides an opportunity to enhance the impact of applied mathematics on solving problems of critical importance to the nation. It will also result in high-impact, lasting contributions to applied mathematics. The process of developing integrated problem-solving strategies for complex systems will lead to the creation of a far richer, deeper set of mathematical tools in areas ranging from basic analysis to software development. These tools will, in turn, lead to improved theoretical understanding and computational methodologies for complex systems. We will provide examples of such questions in Section 5.

3 Research Needs in Topical Areas

Combinations of the different research topical areas have the potential to make significant impact on a number of challenges facing DOE; however, realizing that potential will require major research advances within each area. In this section we discuss the state of the art and the necessary advances in each area. For each area we identify priority research directions needed for DOE applications within that topical area. Priority research directions that combine activities from different topical areas in significant ways are discussed in Section 4.

3.1 Modeling Systems Described by Diverse Phenomena and/or Scales

One common characteristic of complex systems typical of critical DOE applications is that they cannot be treated at adequate fidelity with monolithic models. Models may include multiple sub-models with vastly differing characteristics. Physics-based systems that incorporate a diverse collection of phenomena are referred to as *multiphysics* models or, more generally, *multidisciplinary* models. Models may also need to span a wide range of length and time scales, with different representations for phenomena as a function of scale. Examples of these types of models, typically referred to as *multiscale* models, include combining continuum models with molecular models in fluid mechanics and material modeling, or addressing how long-term investment decisions in the power grid affect reliability and security issues that occur at the second or quicker time-scale. As we move to increasingly complex hierarchical models with an expanded feature space, new mathematical approaches will be needed to understand the impact on system behavior of the interplay between sub-models of different physical processes and at different scales or between decision points in an engineered system.

Multiphysics modeling is central to a host of DOE problems such as combustion, materials science, nuclear energy, and fusion, to name a few. As model complexity grows, new approaches will be needed to accurately model these types of systems. Accurate treatment of the individual components in the model will not be sufficient; accurate simulation of multiphysics phenomena will require the development of new, high-accuracy coupling strategies. This is not simply a numerical analysis issue; one needs to start with an examination of the mathematical structure of the overall problem to analyze the interrelationship between the components and what information needs to be exchanged between them. This underlying structure needs to be reflected in the discretizations or coarse-graining used to approximate the overall model, and in the solver technology needed for effectively solving the resulting systems of equations. Care needs to be taken to understand how the accuracy of the components affects accuracy of the overall system and how uncertainties propagate through the system.

New theoretical and computational tools that bridge time and space scales are also needed to enable the predictive and efficient simulation of multiscale problems. As with multiphysics coupling, multiscale decompositions require advances in mathematical analysis and modeling to correctly formulate the flow of information between scales, such as the initialization of probability distributions at the micro-scale in non-equilibrium situations.

Many DOE applications are complicated by having both multiscale and multiphysics aspects. For example, to design new materials or to improve the properties of existing materials through a rational design methodology, the prediction of material strength requires the coupling of models at multiple scales: molecular dynamics at the microscale to simulate chemistry and defect nucleation; mesoscale models such as dislocation dynamics and single crystal plasticity at the grain scale to capture longer-range elastic and plastic effects; and macroscale continuum models to average the effects of grain texture and microstructure. The accuracy of combined multiscale and multiphysics methods is not understood, reminiscent of the early days of computational fluid dynamics before the foundations were laid by the research of mathematicians such as John von Neumann and Peter Lax. A multifaceted approach involving theoretical analysis, algorithm and code development, and benchmark computational studies is needed to develop reliable and efficient computational tools able to simulate challenging multiphysics and multiscale DOE applications.

Modeling needs stretch even further beyond the development of these fundamental simulation capabilities. For example, new approaches to *surrogate and reduced-order modeling* are needed to accelerate computations, thus making tractable optimization and quantification of uncertainties to support better decisions. Existing methods for surrogate and reduced-order modeling have focused largely on single-physics, single-scale deterministic problems; new approaches are needed for surrogate modeling in the context of challenging multifaceted DOE applications. New methods are needed to quantify the limits on predictive power of the models, referred to as *model inadequacy*, and to *formalize the selection of models* appropriate for the task at hand. This becomes even more critical as we consider complex multiphysics problems that push the boundaries of modeling assumptions. In addition, new methods are needed in situations where there are hierarchies of models at different levels, referred to as *multi-level models*, and where there are families of models with differing fidelity with respect to different quantities of interest, referred to as *multi-fidelity* models. For these classes of problems, systematic approaches are needed to exploit and synthesize estimates from a broad range of modeling options.

Priority research directions to address these challenges include:

- formulate and analyze methods to bridge a wide range of space and time scales
- develop rigorous coupling techniques for multiphysics/multidisciplinary models
- develop methodologies for constructing and using multi-fidelity models needed to support multi-level, multidisciplinary modeling, including approaches for model selection
- develop methods to quantify and account for model inadequacy, especially in the context of multi-level and multidisciplinary systems.

3.2 Optimization Theory and Algorithms

Optimization research has made major strides in the past fifty years and is now routinely used in a wide variety of disciplines as a critical tool for design, prediction, strategic planning, and operational decision making. Optimization provides a rich and broad framework for scientific discovery, ameliorating and treating uncertainty, knowledge mining, low-loss information compression, and multiscale model calibration. Optimization necessitates an abstraction and understanding of the underlying system so that enhancements can occur at many levels using controls, designs, or additional constraints on operational

practice. Modern optimization algorithms are capable of routinely delivering solutions for linear problems with tens of thousands of variables and constraints. However, they are often limited to delivering only local solutions to nonlinear problems of this magnitude. Significant work remains to successfully address many problems of interest to DOE that will require nonlinear optimization approaches. Many of these challenges are discussed below.

Reaching the best decisions and outcomes for systems with rapidly increasing complexity that characterize DOE applications requires the development of advanced high-fidelity optimization technology that is capable of exploiting real-time information arising from a collection of interacting models, while producing acceptable, robust solutions that are useable by a decision maker in a timely fashion. Generation of such solutions requires the development and solution of multiple connected models that incorporate uncertainties, work on different time scales, and engage stakeholders at different levels with differing priorities. Decision making with such high fidelity models raises fundamental new mathematical questions in optimization, and it requires further development of time, space, and energy efficient algorithms capable of exploiting rapidly evolving modern computer architectures.

Optimization models arising in the DOE context are nonlinear with continuous, integer, and categorical variables, resulting in significant research challenges, particularly when coupled to an effort to incorporate parameter uncertainty. For example, within fossil energy systems, decisions on plant startup, shutdown, and transitions involve the nonlinearity inherent in ramp-up efficiencies, integer variables such as processing units, and uncertain parameters such as consumer demand. There is little mathematical understanding of theory or solution methods for nonlinear optimization problems with general mixed integer variables under parameter uncertainty. DOE applications have a large number of uncertain parameters (e.g., wind conditions at different locations), and often their behavior must be characterized from only a small number of samples due to the large sample cost (as is the case, for example, for high-resolution weather forecasts needed in renewable energy problems). Mathematical challenges involve the development of methods that reliably incorporate limited information, systematically reduce both the dimension and the number variables needed to describe the distribution of the uncertain parameters and consistently produce solutions that are robust under sampling errors.

An additional set of mathematical challenges arise when problems have nonlinear, non-convex, non-differentiable, and disjunctive constraints. For example, such nonlinearities arise in energy markets when game theoretic approaches are used to model price and supply of electricity from independent system operators. Finding global solutions of such problems requires the development of theory and algorithmic frameworks that might work with suitable approximations to rapidly generate good solutions, with rigorous bounds on deviations from best possible solutions. Fast reliable heuristics must be developed to produce real time solutions under continuous data feed.

In a coupled multi-system setting, decisions made in one arena may significantly impact the performance of another arena. Added complications arise when a difference in relative importance is given to different objectives used in decision-making in the sub-systems. Such policy and operational problems are modeled using multilevel and multi-objective optimization. The multilevel structure arises with a certain decision hierarchy, e.g., when there are regulator-imposed constraints on operator environmental

characteristics. The multi-objective problems have several competing objectives, e.g., comfort versus room temperature and humidity; or speed versus gasoline consumption. These problems have so far evaded acceptable modeling or solution strategies and remain an unresolved challenge.

Optimization problems arising in applications are typically limited by the inability to process realistically sized models; and the demand for large scale solution increases as the methodology is adopted more widely. These problems frequently possess additional model structure, due to underlying physics, system design or operation, or mathematical modeling features. Efficient solution strategies should exploit such problem structures, for example by combining domain-specific fast algorithms for the structured part of the problem with more robust general purpose algorithms for unstructured part of the problem to speed up overall computation. More research is needed to develop modeling paradigms that naturally allow and/or identify known structures in a problem. Lacking this information, which is typical in existing modeling environments, is the most common obstacle in demonstrating the benefit of advanced algorithms for complex applications.

Priority research directions for addressing the aforementioned challenges include:

- development of algorithms and analysis for finding global solutions of nonlinear non-smooth problems with mixed variable types
- development of theory and solution methods for the new challenges posed to nonlinear optimization problems by the treatment of uncertainty, such as: interaction of uncertainty with integrality constraints, efficient estimation and control of the sampling errors, handling of the sharply increased problem complexity by reduction methods and scalable algorithms
- development of methods capable of exploiting problem structures such as networks, variational, complementarity, polynomial, and conic constraints, as well as capable of generating solutions with desired (sparse) structure
- development of modeling tools that are capable of facilitating decisions for hierarchical, multi-level, multi-objective problems.

3.3 Stochastics and Uncertainty

Providing useful information about complex systems of interest to DOE requires not only the ability to model the system but also the ability to characterize the uncertainties in the model. Uncertainty can arise from imperfectly known or random inputs, inadequacies of the models used to describe the system, and errors or noise in the data acquired about the system. Furthermore, many systems are characterized by inherently stochastic dynamics. Stochastic processes can arise from random forcing or from the effects of phenomena not represented in a model of the system. This latter source of stochastic behavior is a common characteristic of multiscale or hierarchical models in which dynamics at a fine scale appears as stochastic forcing at a coarser scale. Stochastic models and uncertainty are not the same; however, they are discussed together here because uncertainty is commonly represented probabilistically, and the two areas share many mathematical tools. Furthermore, many systems contain aspects of both, with the stochastic behavior contributing to the uncertainty in the overall model.

Although there has been significant progress in the mathematical and computational treatment of both stochastic systems and uncertainty in recent years, the available tools are insufficient to meet DOE's needs. Computational tools for uncertainty quantification are currently limited to a modest number of random or uncertain inputs in models of modest computational complexity. Many existing stochastic methods rely on the assumption that the underlying process is Markovian with Gaussian noise. There have been a number of recent advances in numerical methods for stochastic differential equations such as kinetic Monte Carlo, the stochastic simulation method, and tau leaping; however, the state-of-the-art lags behind comparable methodologies for deterministic equations.

Addressing critical DOE problems will require advances across a wide range of topics in uncertainty and stochastic modeling. New approaches for uncertainty quantification are needed to treat systems with a large number of uncertain parameters and systems where the underlying model is highly complex. Of particular interest in this regard is the development of methodologies to effectively model the propagation of uncertainties in multiphysics and multiscale systems. We also need to develop a much richer toolset for modeling stochastic processes. Approaches are needed for systems where the underlying distributions are not Gaussian, including non-stationary or even parametric distributions, such as particle systems with small numbers of particles or the distribution of hydraulic properties in subsurface models. Techniques are also needed for processes that are non-Markovian. Non-Markovian models arise in systems that have memory effects and systems in which one can only observe part of the state. Similarly, we need to make significant improvements in methods for stochastic differential equations, which, in some cases, will require not only the development of numerical methods but also development of the basic fundamental theory. These types of stochastic differential equations often arise in mesoscale models in continuum mechanics where the stochastic forcing represents fluctuations at the microscale. For example, capturing the tail of the velocity autocorrelation function of a Brownian bead in a fluid requires a model for molecular fluctuations; it cannot be accurately predicted with deterministic fluid equations. Additionally, for many complex systems, we are not simply interested in computing mean behavior. Often we also want the ability to characterize rare events that may require knowledge of the entire probability density function (PDF) of the system of interest. For example, we would like to be able to predict cascading failure of the electric grid, assess worst-case scenarios for environmental remediation problems, or identify design flaws in nuclear reactors that lead to low-probability but serious system failures.

Priority research directions in this area are:

- algorithms to quantify uncertainty with respect to data that is high dimensional or continuous
- modeling methodology for general classes of stochastic processes, including models that are non-Markovian and processes where the underlying distribution is not Gaussian
- analysis and numerical methods for stochastic differential equations
- improved methodologies to characterize and identify rare events.

3.4 Integrating Data and Modeling

Data drives science. It comes from many sources and in many forms, for example, from high-resolution sensors, large-scale astronomical surveys, high-throughput experiments, and the output of high-performance computer simulations. DOE scientific user facilities such as the Advanced Photon Source

(APS) and the Spallation Neutron Source (SNS) currently give rise to terabytes of data. Large datasets also arise in astrophysics (for example, the planned Large-Scale Synoptic Telescope) and in geophysical/seismic observations. The amount of data generated by the Advanced Light Source (ALS) at Lawrence Berkeley Laboratory will soon approach multiple petabytes per week. Data is available—often in streaming form—from sensors that monitor systems such as the atmosphere, the power grid, and the real-time output of experiments.

A DOE Workshop on Mathematical Analysis of Petascale Data in June 2008 [15] described the many application domains in which data plays a key role, and outlined the role of mathematics and computation in analysing data for various purposes. A Workshop on Exascale Data Management, Analysis, and Visualization in February 2011 [16] focused on data issues at extreme scales. The October 2011 ASCR/BES Data Workshop examined the data and communication needs in Basic Energy Sciences. All of these workshops make clear that mathematical analysis of data informs a large part of the DOE mission. While we discussed many aspects of data in the present workshop, our focus was mainly on data as it relates to complex systems, in particular, the construction of practical, valid models.

In data-driven modeling, observational data may be used to guide construction and assembly of the model and to adjust and identify the parameters that define a model. This process improves the fidelity of the model and allows it to be used for its intended purposes, namely those of prediction or analysis. One type of methodology in uncertainty quantification uses sampled data to capture the inherent variability in a phenomenon of interest. The resulting estimated distribution can be propagated through models to yield a distribution of outcomes, which can then be analysed or optimized. Feature extraction can also be useful in reducing the cost of future data collection and sensing, by identifying where to take measurements to optimize the utility of the data. Similarly, analysis of data can lead to reduction in the complexity and size of the models that drive them, thus speeding computations without damaging the quality of predictions.

Reconstruction of complex metabolic networks and other “reverse engineering” methods require new algorithms for solving large-scale inverse problems and for performing parameter identification, integrating data with available theoretical understanding. Control of systems is also driven by data in an explicit feedback fashion. The emerging field of statistical learning is centered on data analysis; there is great potential to adapt techniques from this field to development and validation of models. Optimization algorithms of all kinds—discrete and continuous; linear, nonlinear, and nonsmooth; stochastic optimization—form a core technology in data analysis and its applications in complex systems. Indeed, the demands of these data analysis areas have sparked new fundamental research directions in optimization and are leading to exciting collaborations across disciplines.

Sampling methodologies and other techniques from applied probability are vital in dealing with large data sets. Such issues arise, for example, in environmental monitoring with large-scale sensor networks. Some algorithms may need to sample from a static or streaming data collection, as it may be impossible to process the full data set, due to computational limitations and/or power/energy constraints. Bayesian methods such as Markov Chain Monte Carlo (MCMC) are powerful tools for analysis of complex data and integration with models. There is a need for ever more efficient algorithms to sample stationary distributions.

In short-term weather prediction for energy systems management, more efficient and accurate state estimation methods are required, including variants of particle filtering. For inverse problems, data assimilation, and parameter estimation, there are a number of challenges to be faced: reducing memory requirements for adjoint computation, better inference from sparse and noisy data, and better global optimization algorithms, to name a few.

Specific high-priority research directions for data as it relates to complex systems, particularly in the areas of constructing, reducing, and validating models, include:

- Algorithms for data-driven model validation, reduction, and control. Algorithms from machine learning, optimization and model predictive control are all relevant. Algorithms that will efficiently utilize advanced computer architectures are especially needed.
- Robust inversion/optimization/learning algorithms for noisy data.
- Sampling methodologies for data reduction; selective discarding of streaming data.
- Methods that economize on sampling using randomization and feedback in the sampling strategy (leveraging techniques from compressed sensing, active learning, and adaptive control).

4 Combining Technologies to Solve Problems

As discussed in Section 2, solving the problems that are critical to DOE mission requirements in energy, environment, and national security will require a concerted strategy that marshals mathematical techniques from a number of different areas. Section 3 identifies four of these areas and discusses a number of priority research directions where advances are needed to address important DOE problems. Although advancing these specific areas is a necessary ingredient in solving these problems, these advances in isolation will not be sufficient. Making significant progress on answering key questions about complex systems will require that we also develop an integrated effort that provides the mathematical framework needed to combine mathematical advances in several areas. This new framework will need to address all aspects of the problem-solving process, ranging from the formulation and analysis of mathematical models, to the development of new discretization approaches, to the development of solver methodologies and software.

Meeting DOE's mission requirements will require the ability to model systems of unprecedented complexity. Dealing with these systems will require new types of models that couple a broad array of diverse processes spanning a wide range of spatial and temporal scales. The models will likely combine both deterministic and stochastic elements; they are likely to switch between alternative representations of processes as a function of dynamic properties of the problems. Furthermore, the goal in studying these types of systems will not be to simply produce a simulation of the system but rather to answer difficult questions regarding design, risk analysis, or optimization with quantified bounds on the error. Answering these types of questions will require combinations of new mathematical tools from several different areas. However, developing these tools independently, with the hope that they will be able to work together effectively to meet the challenges of complex systems, is unlikely to be successful. One needs to pursue integrated approaches that bring together all of the mathematical components needed to solve a problem—a multifaceted mathematics approach. For this type of integrated approach to be successful, we need to consider all aspects of the problem-solving process, ranging from the mathematical formulation, to the development of approximate models, to the implementation of solver technologies for emerging architectures. With the flexibility to simultaneously consider different aspects of the problem-solving process, we can hope to find the best match of formulation, algorithms, and solvers. Furthermore, this development needs to look at the problem holistically from a mathematical perspective so that the requirements for the individual components and how they are coupled together can be elucidated. The basic idea here is similar in spirit to the notion of co-design, in which one simultaneously and holistically considers a specific application, the supporting algorithms, computer science, and computer architecture in an integrated approach to the design of extreme computing systems. Here, however, the intent is to use that type of approach at a more fundamental level where we simultaneously examine, in an integrated form, all of the mathematical elements needed to make significant advances in the analysis, simulation, and optimization of a complex system.

The starting point in the problem-solving process is the mathematical formulation. Analysis of the mathematical formulation provides a characterization of the structure of the problem. There are fundamental gaps in our current understanding of the mathematical structure of complex systems. Increasing model complexity places a premium on understanding the mathematical structure of the

constituent processes and how they couple to other components of the model. Without this understanding, models will either need to treat the system using first principles, which is typically infeasible for a complex system, or rely on ad hoc coupling or reduction strategies that may not represent the system being modeled adequately and predictively. Mathematical analysis can elucidate the relationship between scales and the key features that determine how processes interact. It can also be used to understand how uncertainties propagate within the model. As we begin to ask more complex questions, such as how to optimize performance of the system or minimize risk from low-probability events, the structure of the model for the system, as well as the question being asked, informs the requirements for the optimization algorithm. A second role of fundamental mathematical analysis is to develop completely new ways to formulate problems that will provide approaches that are more amenable to emerging architectures or offer other significant computational advantages than traditional approaches.

Translating a mathematical model of a complex system into a computational model requires a discrete representation of the system and algorithms for manipulating that representation. Dealing with system complexity accurately will require a computational model that reflects the underlying mathematical structure of the problem. In this integrated framework, the discrete representation might include characterization of uncertainty, a model for merging data with the simulation, or an overarching optimization algorithm. Mathematical formulations that exploit the structure of the problem are typically more complex than traditional generic formulations and require considerably more sophisticated numerical treatments. Computational models for complex systems will need to incorporate advances from the topical areas discussed in section 3 along with advances in traditional areas of applied mathematics. In addition, these models will need to be augmented with approaches for coupling algorithms from different areas to work together. We will need to develop approaches for characterizing and controlling the uncertainty in the algorithmic components and how those uncertainties interact in the overall model.

Computational models of complex systems that exploit the underlying mathematical structure of the system will lead to requirements for efficient solution of linear and nonlinear systems of equations that reflect that structure. These types of systems will place severe requirements on the linear and nonlinear solver methodologies. The development of customized solution strategies that exploit the specialized structure of these systems is another key element of the solution process. Solver methodologies need to take advantage of hierarchical structures within the problem and reflect the mathematical couplings between elements to effectively utilize the architecture. Approaches based on more generic approaches are likely to suffer from significant degradation in performance, particularly as we move to new architectures that place a premium on minimizing data movement.

This holistic approach to problem solving is not a process of working from formulation to discretization to algorithm but rather an iterative consideration of all aspects of the solution process. It is through this iteration that we can arrive at the combination of formulation, discretization, and solver technologies will enable us to weave together the elements of the different components to construct the tools needed to answer important questions about complex systems while making efficient use of emerging computer architectures.

Priority research directions in multifaceted mathematics that combines diverse methodologies to treat complex systems include:

- algorithms and analysis of methodology for multiscale/multiphysics systems that couple stochastic and deterministic models
- systematic approaches to construct surrogate or reduced-order models with quantifiable fidelity for optimization, uncertainty quantification, and decision making
- algorithms and analysis for stochastic optimization that address model uncertainties with limited information on parameter distributions, considering alternative models of uncertainty and risk, while addressing nonlinearities and special structures
- approaches for propagation of uncertainty through the different scales in a multiscale model and through the components of a multiphysics model
- improved algorithms for data assimilation and model/data fusion, for use in system identification, control, and bounding of model uncertainty.

5 Case Studies of the Potential Impact of Multifaceted Mathematics

The combination of technologies from the different topical areas discussed above has the potential for significant impact on complex systems of interest to DOE. In this section we consider several concrete examples to illustrate how this type of applied mathematics research can be used to solve important DOE applications problems.

5.1 Catalysis Discovery for Next Generation Fuels

An important problem for the nation's energy future is the discovery of novel fuels that can release their energy efficiently and sustainably. A critical challenge towards that end is the determination of new and more efficient catalysts that can enable the physical processes needed to make fuel from sunlight and to produce low carbon fuels and biomass-based fuels for transportation.

Understanding and predicting the properties of novel catalysts and designing them for achieving the full potential of novel fuels is a grand challenge of science, mathematics, and computation. Catalysis involves primarily surface effects. Surface structures thus need to be resolved at levels of unprecedented fidelity, a goal possibly unachievable by either experiments or simulation alone. Rather, this requires a process that melds inverse problems that assimilate information from experiments with forward quantum and molecular dynamics calculations that refine the atomic positions and assess the stability of the predicted compound. To this end, multiple mathematical areas need to interact, such as computational methods for dynamical systems, partial differential and integral operators; optimization methods for inverse problems and energy minimization; and data analysis techniques. More is needed than just application of established methodologies; in many cases the path forward requires the development of new mathematical methods that are faster, more robust, and can handle new classes of problems, and perhaps more importantly, approaches that are explicitly designed to work together.

To assess the technological feasibility of the resulting catalyst at the accelerated pace required by the nation's energy needs, modeling and simulation, assisted by data analysis from high-resolution experimental capabilities, must achieve predictability on multiple coupled scales, from the surface structure to reactor level. This requires exchanging information and coupling mathematical models with strikingly different natures at the quantum, atomistic, mesoscopic, and macroscopic scales. Such mathematical models may include

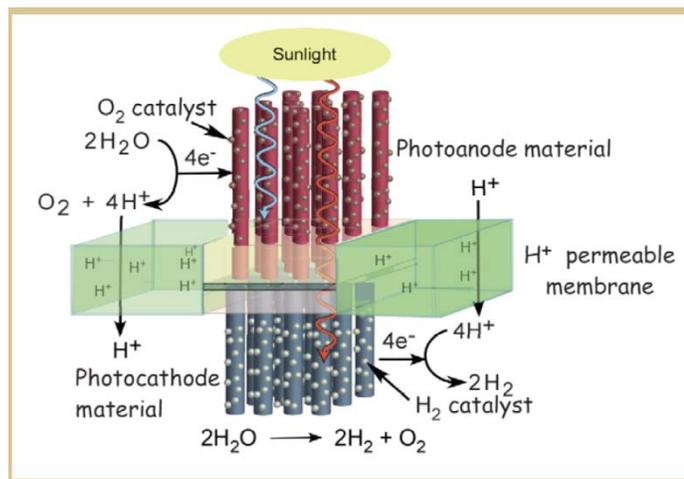


Figure 1. Schematic of a photoelectrochemical cell being designed to harness sunlight for generating chemical fuel. Image courtesy of Nate Lewis and Liz Santori.

quantum mechanics, molecular dynamics, kinetic rate theory, mean field models, and computational solid and fluid dynamics.

A salient feature of such an effort is integration of such models on vastly different spatial and temporal scales. For example, at the atomistic scale, the molecular dynamics time scale is on the order of 1 fs (the time step needed to accurately integrate a vibration period), whereas the catalyst action time scale is on the order of 1ns or larger. The complexity of the water-splitting system described in Figure 1 requires making sure that one captures the right chemical reaction pathways that simulate the dynamics for long enough (10^6 – 10^8 time steps using electronic structure forces, for one instance) to actually see the reaction events. For being responsive to an end-goal endeavor, the molecular dynamics needs to be accelerated for many more physical circumstances than it currently covers and in a multi-model, quantum-to-reactor context. Coarse-graining in space and time needed to study how surface and bulk defects affect the catalysis process requires techniques for resolving multiple spatial and temporal scales. A formal, systematic effort for achieving mathematical model integration on vastly different spatio-temporal scales for the situation where multiple mathematical disciplines are involved is an important necessity for achieving the full potential of modeling and simulation to the end of discovering new catalysts and enabling novel fuels.

5.2 Power Grid

The US electricity grid is the world’s largest engineered system. It consists of a dynamic collection of interacting components, operated under an enormous range of physical, reliability, economic, social, and political constraints that need to be satisfied over time scales ranging from seconds for relay-action and closed-loop control to decades for transmission siting and construction (Figure 2). It is required to meet a varying demand for energy that is hard to accurately predict using an evolving mixture of resources.

Carbon emission and sustainability concerns demand increasing renewable energy penetration, leading to increasing supply uncertainty. Absence of suitable storage solutions mitigates hedging strategies, resulting in increased risk of unreliability, demand shedding, or even blackouts. Smart grid infrastructure will alleviate some of these issues: mathematics to garner information from smart meters, phasors, weather forecasting systems, and other high resolution information networks, and process the vastly larger and more complex state models in real time, are imperative to guide decision makers and instantiate appropriate incentives.

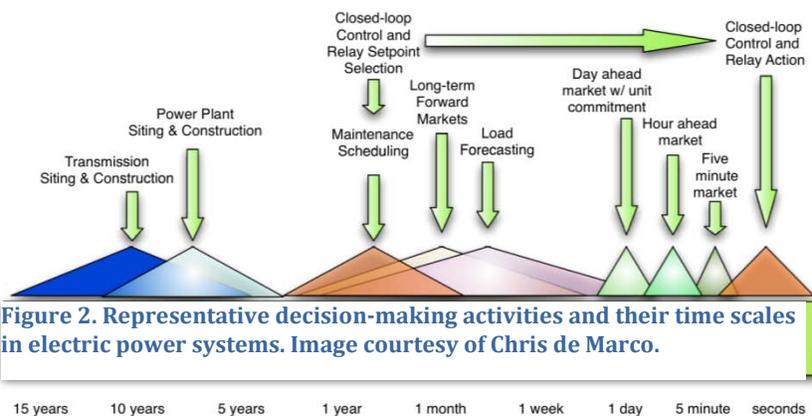


Figure 2. Representative decision-making activities and their time scales in electric power systems. Image courtesy of Chris de Marco.

Multifaceted mathematics has a unique role to play in creating the decision framework that realizes the full potential of novel power grid infrastructure, while satisfying the needs and constraints of all its

stakeholders. It will create new optimization algorithms and models that exploit the multiscale patterns in the power grid and the additional information in new high resolution data sources. It will facilitate adaptivity of the system, provide for control of the components of the grid, allow for treatment and mitigation of uncertainties in the system, and provide understanding of the interaction effects. Developing mathematical interfaces and exploiting the hierarchical structure will gain flexibility for decision makers and will improve overall solution speed so that complex mathematical models can be deployed in both strategic and operational settings, facilitating understanding of the localized effects of broader decisions while generating improved performance of the system.

A multifaceted mathematics approach will have significant impact on a number of fundamental power grid tasks. For example, *stochastic unit commitment and dispatch*—the decision process by which the Independent System Operators decide what resources are committed to satisfy the demand—involves a model that includes stochastic, nonlinear, and integer features in a real time environment that aim to incorporate effects due to wind and solar supply, to mitigate effects of changing weather patterns and to exploit the potential of newly developed technologies, such as new battery technologies and electric powered vehicles. The underlying models involve networks of continental scale, and must provide dynamical models to capture load profiles, market bidding, ramping effects, and varieties of failures. Interfacing these models on longer time scales to enable day-to-day co-optimization or even longer scales for transmission line siting and generator expansion will require exploiting another facet or decision axis, and will force the linking of physical models, operational decisions, and information from a wide variety of data conduits. Models will move from a reactive environment to a real-time predictive one, using new mathematical constructs integrating optimization with nonlinear, stochastic, and integer features with multiscale dynamics and data analysis.

5.3 Nuclear Reactor Core Reloading

Nuclear reactors offer a rich source of potential application of a multifaceted mathematical approach. One important part of nuclear reactor plant management is core fuel reloading. A nuclear reactor core is composed of a few hundred regularly aligned assemblies that each contain an array of a few hundred individual fuel pins surrounded by coolant. As the fuel is consumed, the most depleted assemblies need to be replaced. However, the fuel depletion is not spatially uniform, and placing new fuel rods at the location of the depleted extracted rods is not optimal from either an economic or a safety perspective. For this reason at the end of each cycle, the most depleted assemblies are discarded, fresh assemblies are inserted, and the remaining assemblies are reshuffled in an attempt to maximize fuel utilization while satisfying safety limitations and operational constraints. We refer to the determination of the desired arrangement of assemblies as the reactor core reload pattern design.

The objective of shuffling is to distribute the fuel elements in a way to keep the reactor critical for as long as possible (e.g., maximize fuel utilization) while at the same time making sure that no individual fuel rod is overheated. Maximizing fuel utilization tends to favor clustering of fresh assemblies, and thermal considerations favor distributing fresh assemblies. Thus, there is a very delicate trade-off between safety and efficiency. The design of core loading patterns is further complicated as designers utilize burnable

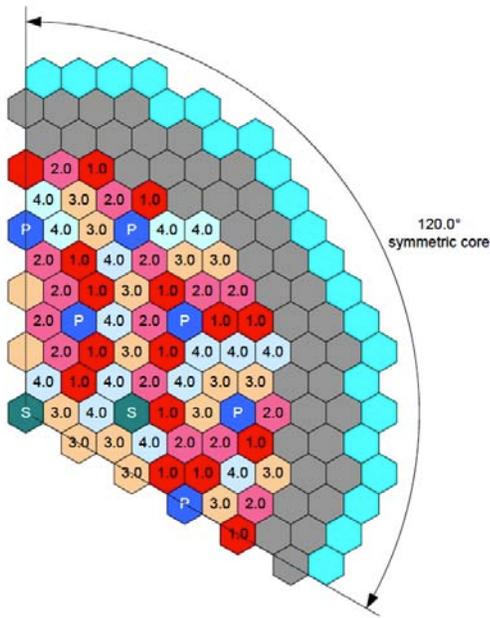


Figure 3. The optimal disposition of nuclear fuel rods from four cycles in an ABR-1000 reactor. Image courtesy of Taek K. Kim, Argonne National Laboratory.

absorbers within selected fresh fuel assemblies to control the evolution of pin power distributions as the core is depleted.

From a mathematical perspective, nuclear core reloading is a daunting challenge, as it encapsulates a combinatorial optimization problem whose constraints require high-accuracy, complex multiphysics simulations. Merely verifying a single proposed core loading pattern requires advances in multiphysics computations such as improved methods for solid, fluid, and chemical interactions, coupling between discrete and continuum scales, and coupling between deterministic and stochastic models. In addition, there is a need to include quantification of uncertainty in both predicted objective functions and constraints from the core simulations, across all domains. Solving the optimization problem requires efficiently traversing a large combinatorial space, where high-fidelity simulation at each point would require a complete complex multiphysics simulation, which is often prohibitively expensive, even

with the availability of extreme-scale computational resources. Thus, surrogate models are needed for the purpose of optimization. This calls for a change in the approach to surrogate modeling, since such models are traditionally designed to capture the physical phenomena being simulated, while optimization algorithms can permit a relaxation that can bound the feasible space. Moreover, the integrated treatment of uncertainty in the reloading pattern design process will require advanced stochastic optimization tools that accommodate both continuous and discrete variables.

While the core reload optimization problem is extremely difficult, its importance and need for computational tools is indisputable. Decisions are being made in a complex landscape where a 1 percent improvement in fuel utilization translates into monetary savings of \$1M per year for a medium sized power plant, and incorrect placement strategies can lead to decreased stability and unacceptable levels of risk for nuclear plant safety. Mathematical techniques that transition to optimization guided decision support tools can significantly transform operations and lead to increased energy production, improved safety, and reduce nuclear waste.

5.4 Remediation

One of the main environmental challenges that DOE and the nation face is massive groundwater contamination. Within the DOE complex, there are more than 10,000 contaminated sites, with contaminants ranging from organic contaminants such as gasoline and commercial solvents to radioactive contaminants. Cost-effective remediation strategies are required to clean up the subsurface and to protect the country's water supply. Accurate models for subsurface flow and transport are needed to design optimal and efficient remediation strategies. Such models are also important in other environmental and

engineering applications, including management of water resources, petroleum and natural gas production, geothermal energy, and geological CO₂ sequestration.

Flow and transport in porous media can be described on different scales. On the pore scale, subsurface flow and transport are described by a combination of Navier-Stokes and advection-diffusion equations. On the field scale, the porous media can be treated as a continuum, and dynamics of the system can be described by Darcy's law and advection-dispersion equations. The pore-scale description of transport is based on fundamental conservation laws and is able to accurately simulate flow and transport under a wide range of conditions. However, solving the microscale equations is infeasible in most practical applications where one is mostly interested in field to regional scale behavior.

For field-scale problems, macro-scale transport equations are necessary. When there is a clear scale separation, e.g., when microscale variables vary slowly on the macroscopic scales, accurate empirical or analytical closures can be developed. However, when these scale separation conditions are even partially violated, the assumptions underlying analytic closures break down, and we need to rely on empirical closures that are less accurate. Furthermore, because of their empirical nature, the a priori accuracy of these empirical models is impossible to establish. Examples of reactive transport problems where scale separation does not exist and traditional advection-dispersion equations may become inaccurate include multi-component reactive transport, mineral precipitation and/or microbial growth, mixing controlled reactions, nonlinear surface chemistry, etc. In these cases, accurate simulation of macroscopic behavior may require explicit simulation of pore- and sub-pore-scale processes. Recently, a number of hybrid and multiscale methods have been proposed to couple pore-scale and macro-scale models to avoid simulating pore-scale processes over the entire spatial and temporal simulation domain. However, the development of these techniques is still in its infancy, and there is a strong need for significant advances in multiscale/multiphysics models for subsurface.

Another challenge in flow and transport models is insufficient and often inaccurate data characterizations of the subsurface and uncertain initial and boundary conditions. Often the data represents different types of information from different sources. For a given site, the knowledge about geology and physical processes can be so limited that there are several competing conceptual models. As a result, subsurface modeling can be highly uncertain, and characterization of the uncertainty becomes a crucial issue for decision-making and risk analysis in the design of remediation strategies. New optimization techniques and data assimilation strategies for parameter estimation from noisy sparse data are needed to make optimal use of available information.

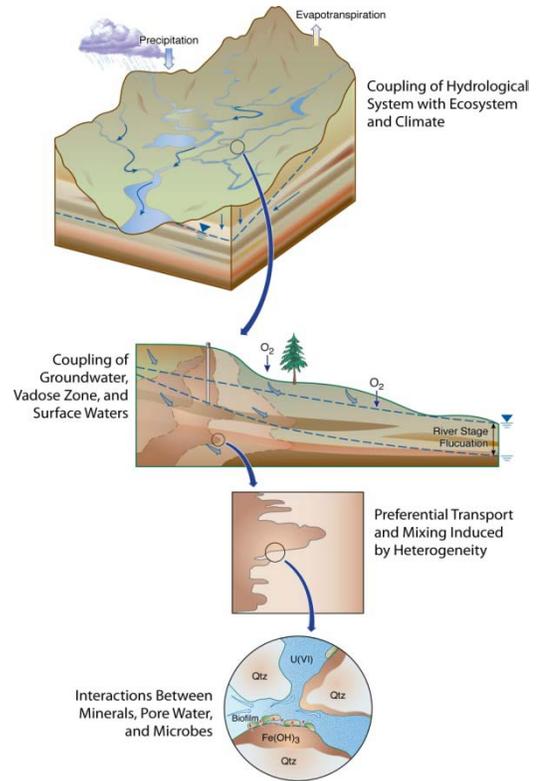


Figure 4. Illustration of multiscale character of subsurface flow. Image courtesy of S. Hubbard.

The need to make predictions in the face of the unknown variability of the subsurface makes uncertainty quantification an essential part of subsurface modeling. Existing UQ theory provides a rigorous probabilistic framework for propagating uncertainty but is limited in the degree of complexity that can be treated, restricting its applicability to relatively simple problems. Subsurface modeling requires new approaches that can deal with the “curse of dimensionality” for highly variable random fields such as hydraulic conductivity that can require thousands of independent random variables for an accurate description. The resulting methodology must also integrate seamlessly with the multiscale nature of remediation problems.

The target for a given remediation site is to find a cost-effective remediation strategy that provides assurances that environmental concerns have been adequately mitigated. Developing these types of strategies will require the development of optimization strategies that can be used to design effective strategies that respect the complex and uncertain nature of subsurface flow. Elements of a suitable approach for these types of problems would need to address optimization in the context of high-dimensional noisy data and include methodology for estimating rare events within the system.

6 Conclusions

DOE's mission requirements are driving a need to address increasingly complex scientific and technological challenges. Meeting these challenges will require the ability to model systems of unprecedented complexity. Furthermore, the goal in studying these types of systems will not be to simply produce a simulation of the system but rather to answer difficult questions regarding design, risk analysis, or optimization with quantified bounds on the error. Answering these types of questions will require combinations of new mathematical tools from several different areas. The workshop identified four key research directions in applied mathematics where advances are needed to address complex systems arising in DOE mission areas:

- modeling systems characterized by diverse phenomena and/or scales
- optimization
- stochastic systems and systems with uncertainty
- integration of data and modeling.

While advances in these research areas will have clear impact on DOE needs, the workshop concluded that developing these tools independently with the hope that they will be able to work together to meet the challenges of complex systems is not the most effective strategy. A more effective strategy is to pursue integrated approaches that consider a complex system as a whole and bring together all of the mathematical elements that are needed to answer critical questions about the system. This strategy is referred to as a *multifaceted mathematics* approach. The promise of such an endeavor is to provide new insights into problems of major DOE importance that exhibit features that can be characterized only by mathematical models that combine multiple mathematical subdisciplines. Examples of such applications were discussed in Section 5.

Multifaceted mathematical research of the type described here has not been carried out to date on the scale and with the focus that this report identifies as necessary. Pursuing this type of research activity represents a paradigm shift in the way applied mathematics research is conducted and will require new organizational structures to support a new model for research. Specific recommendations to facilitate this type of research include:

- teams of multiple researchers that combine expertise from different areas of mathematics working together toward a common goal
- early and deep engagement between mathematicians and domain experts from target application areas
- long-term support with the flexibility to adjust research directions as projects evolve.

Single investigator research remains an important source of innovative mathematical ideas. The organizational structure that supports multifaceted mathematics research needs to include mechanisms that enable single investigators to couple to larger efforts in order to provide a conduit for rapid incorporation of innovative breakthroughs to mission critical problems.

Acknowledgment

The organizers are grateful to the participants to the workshop and to several other scientists (Kord Smith, Laurence Marks, Michael Ortiz, and Art Voter) who have provided input and feedback, particularly on the DOE mission area examples in Section 5 of this report.

Bibliography

- [1] *U. S. Department of Energy Strategic Plan*, May 2011, http://energy.gov/sites/prod/files/2011_DOE_Strategic_Plan_.pdf.
- [2] *A Workshop to Identify Research Needs and Impacts in Predictive Simulation for Internal Combustion Engines (PreSICE)*, March 2011, http://www1.eere.energy.gov/vehiclesandfuels/pdfs/presice_rpt.pdf.
- [3] *DOE/NETL Carbon Dioxide Capture and Storage RD&D Roadmap*, U.S. Department of Energy, National Energy Technology Laboratory, December 2010, http://www.netl.doe.gov/technologies/carbon_seq/refshelf/CCSRoadmap.pdf.
- [4] *Computational Materials Science and Chemistry: Accelerating Discovery and Innovation through Simulation-Based Engineering and Science*, July 2010, http://science.energy.gov/~media/bes/pdf/reports/files/cmssc_rpt.pdf.
- [5] *Materials Genome Initiative for Global Competitiveness*, June 2011, http://www.whitehouse.gov/sites/default/files/microsites/ostp/materials_genome_initiative-final.pdf.
- [6] *Nuclear Energy Research and Development Roadmap*, U.S. Department of Energy, April 2010, http://www.nuclear.energy.gov/pdfFiles/NuclearEnergy_Roadmap_Final.pdf.
- [7] *Proceedings of the Computational Needs for the Next Generation Electric Grid Workshop*, April 19-20, 2011, http://energy.gov/sites/prod/files/FINAL_CompNeeds_Proceedings2011.pdf.
- [8] *Advice on the Department of Energy's Cleanup Technology Roadmap: Gaps and Bridges*, Committee on Development and Implementation of a Cleanup Technology Roadmap, National Research Council, The National Academies Press, 2009.
- [9] *Status of Environmental Management Initiatives to Accelerate the Reduction of Environmental Risks and Challenges Posed by the Legacy of the Cold War*, U.S. Department of Energy, January 2009, <http://www.em.doe.gov/pdfs/NDAA%20Report-%2801-15-09%29a.pdf>.
- [10] *Nuclear Posture Review Report*, U.S. Department of Defense, April 2010, <http://www.defense.gov/npr/docs/2010%20nuclear%20posture%20review%20report.pdf>.
- [11] *Applied Mathematics at the U.S. Department of Energy: Past, Present and a View to the Future*, Lawrence Livermore National Laboratory, LLNL-TR-401536, May 2008.
- [12] Hendrickson, B.A. and Wright, M.H., *Mathematical Research Challenges in Optimization of Complex Systems*. December 2006.
- [13] Brase, J.M. and Brown, D.L., *Modeling, Simulation and Analysis of Complex Networked Systems*, Lawrence Livermore National Laboratory, LLNL-TR-412733, May 2009.
- [14] Brown, D.L. and Messina, P., *Scientific Grand Challenges: Crosscutting Technologies for Computing at the Exascale*, Pacific Northwest National Laboratory, PNNL-20168, February 2010, http://science.energy.gov/~media/ascr/pdf/program-documents/docs/Crosscutting_grand_challenges.pdf.
- [15] *Mathematics for Analysis of Petascale Data*, Report of a Department of Energy Workshop, June 2008, http://science.energy.gov/~media/ascr/pdf/program-documents/docs/Peta_scaled_at_a_workshop_report.pdf.
- [16] *Scientific Discovery at the Exascale: Report from the DOE ASCR 2011 Workshop on Exascale Data Management, Analysis, and Visualization*, February 2011, <http://science.energy.gov/~media/ascr/pdf/program-documents/docs/Exascale-ASCR-Analysis.pdf>.

Appendix: Workshop Participants

Last Name	First Name	Email	Organization
Alexander	Frank	fja@lanl.gov	Los Alamos National Laboratory
Anitescu	Mihai	b48560@zimbra.mcs.anl.gov	Argonne National Laboratory
Balakrishnan	Kaushik	kaushikb@lbl.gov	Lawrence Berkeley National Laboratory
Bell	John	jbbell@lbl.gov	Lawrence Berkeley National Laboratory
Berzins	Martin	mb@sci.utah.edu	University of Utah
Bienstock	Daniel	dano@columbia.edu	Columbia University
Bindewald	Gil	Gilbert.Bindewald@hq.doe.gov	U.S. Department of Energy
Brown	David	dlb@lbl.gov	Lawrence Berkeley National Laboratory
Christlieb	Andrew	christlieb@math.msu.edu	Michigan State University
Clark	Carlton	Charlton.Clark@ee.doe.gov	U.S. Department of Energy
Constantinescu	Emil	emconsta@mcs.anl.gov	Argonne National Laboratory
Darve	Eric	eric.darve@stanford.edu	Stanford University
Davis	Nora	nora.deram@gmail.com	University of Texas at Austin
Diachin	Lori	diachin2@llnl.gov	Lawrence Livermore National Laboratory
Dixon	Paul	p_dixon@lanl.gov	Los Alamos National Laboratory
Ferris	Michael	ferris@cs.wisc.edu	University of Wisconsin
Forest	Greg	forest@unc.edu	University of North Carolina, Chapel Hill
Gabriel	Steve	sgabriel@umd.edu	University of Maryland
Goodman	Jonathan	goodman@cims.nyu.edu	Courant Institute, NYU
Gunzburger	Max	gunzburg@fsu.edu	Florida State University
Hart	Chris	Chris.Hart@ee.doe.gov	Department of Energy
Hauck	Cory	haukc@ornl.gov	Oak Ridge National Laboratory
Hill	Judy	hilljc@ornl.gov	Oak Ridge National Laboratory
Hittinger	Jeff	hittinger1@llnl.gov	Lawrence Livermore National Laboratory
Hovland	Paul	hovland@msc.anl.gov	Argonne National Laboratory
Huang	Henry	zenhu.huang@pnnl.gov	Pacific Northwest National Laboratory
Katsoulakis	Markos	markos@math.umass.edu	University of Massachusetts, Amherst
Knoll	Dana	nol@lanl.gov	Los Alamos National Laboratory
Laird	Carl	carl.laird@tamu.edu	Texas A&M University
Landsberg	Sandy	Sandy.Landsberg@science.doe.gov	U.S. Department of Energy
Larzelere	Alex	alex.larzelere@nuclear.energy.gov	U.S. Department of Energy
Laviolette	Randall	randall.laviolette@science.doe.gov	U.S. Department of Energy
Livermore	Dave	lvrmr@math.umd.edu	University of Maryland
Lin	Guang	guang.lin@pnnl.gov	Pacific Northwest National Laboratory
Luedtke	Jim	jrleudt1@wisc.edu	University of Wisconsin-Madison
Luskin	Mitchell	luskin@umn.edu	University of Minnesota
Mehrotra	Sanjay	mehrotra@iems.northwestern.edu	Northwestern University
Meza	Juan	jcmeza@ucmerced.edu	University of California
Mitchell	John	luskin@umn.edu	Rensselaer Polytechnic Institute
Mitran	Sorin	mitran@unc.edu	University of North Carolina
Moser	Bob	rmoser@ices.utexas.edu	University of Texas at Austin
Moulton	David	moulton@lanl.gov	Los Alamos National Laboratory
Munson	Todd	tmunson@mcs.anl.gov	Argonne National Laboratory
Najm	Habib	hnnajm@sandia.gov	Sandia National Laboratories

Nash	Stephen	snash@gmu.edu	George Mason University
Oberai	Assad	oberaa@rpi.edu	Rensselaer Polytechnic Institute
O'Neill	Dick	richard.oneill@ferc.gov	FERC
Ozdaglar	Asu	asuman@mit.edu	Massachusetts Institute of Technology
Pang	Jong-Shi	jspang@illinois.edu	University of Illinois, Urban-Champaign
Pao	Karen	Karen.Pao@science.doe.gov	U.S. Department of Energy
Pederson	Mark	Mark.Pederson@science.doe.gov	U.S. Department of Energy
Pinar	Ali	alipinar@gmail.com	Sandia National Laboratories
Plechac	Petr	plechac@math.utk.edu	University of Delaware
Purnell	Jeannie	J.Robinson@orise.orau.gov	ORISE
Ralphs	Ted	ted@lehigh.edu	Lehigh University
Scheinberg	Katya	katyas@lehigh.edu	Lehigh University
Sen	Suvrajeet	sen.22@osu.edu	Ohio State University
Sethian	James	sethian@math.berkeley.edu	Lawrence Berkeley National Laboratory
Shadid	John	jnshadi@sandia.gov	Sandia National Laboratories
Shanbhag	Uday	udaybag@illinois.edu	University of Illinois at Urbana-Champaign
Shoemaker	Christine	cas12@cornell.edu	Cornell University
Singer	Mike	msinger@llnl.gov	Lawrence Livermore National Laboratory
Tang	Lisa	lmtang@wisc.edu	University of Wisconsin
Tartakovsky	Alexandre	alexandre.tartakovsky@pnl.gov	Pacific Northwest National Laboratory
Watson	Jean-Paul	jwatson@sandia.gov	Sandia National Laboratories
Willcox	Karen	kwillcox@mit.edu	Massachusetts Institute of Technology
Williamson	Mark	mark.williamson@em.doe.gov	U.S. Department of Energy
Wright	Steve	swright@cs.wisc.edu	University of Wisconsin-Madison
Zavala	Victor	zavala@zimbra.mcs.anl.gov	Argonne National Laboratory