Basic research needs for

INVERSE METHODS
FOR COMPLEX SYSTEMS UNDER UNCERTAINTY

***ASCR WORKSHOP** JUNE 10TH – 12TH 2025*

U.S. DEPARTMENT OF ENERGY | Office of Science

Basic Research Needs for Inverse Methods for Complex Systems under Uncertainty

Prepared for the U.S. Department of Energy Office of Science Advanced Scientific Computing Research Program

Co-Chairs	Jeffrey Donatelli	Lawrence Berkeley National Laboratory
	John Jakeman	Sandia National Laboratories
	Michael Shields	Johns Hopkins University
Organizing Committee	Anne Gelb	Dartmouth College
	Felix Herrmann	Georgia Institute of Technology
	Sanket Jantre	Brookhaven National Laboratory
	Jeffrey Larson	Argonne National Laboratory
	Juliane Mueller	National Renewable Energy Laboratory
	Assad Oberai	University of Southern California
	Cosmin Petra	Lawrence Livermore National Laboratory
	Brendt Wohlberg	Los Alamos National Laboratory
	Guannan Zhang	Oak Ridge National Laboratory
Other Contributors	Somayajulu Dhulipala	Idaho National Laboratory
	Bamdad Hosseini	University of Washington
	Drew Kouri	Sandia National Laboratories
	Sven Leyffer	Argonne National Laboratory
	Michael McCann	Los Alamos National Laboratory
	Teresa Portone	Sandia National Laboratories
	Juan Restrepo	Oak Ridge National Laboratory
	Stefan Wild	Lawrence Berkeley National Laboratory
	Timothy Wildey	Sandia National Laboratories
DOE Points of Contact	Hal Finkel	Advanced Scientific Computing Research
	Todd Munson	Advanced Scientific Computing Research
	David Rabson	Advanced Scientific Computing Research

DOI: https://doi.org/10.2172/2583339

Contents

Executive Summary

Inverse problems, which aim to infer unknown properties of a system using experimental and observational data, are central to addressing many of the U.S. Department of Energy's (DOE) most critical scientific and engineering challenges. Accurate, computationally efficient, and data-efficient solutions to inverse problems are essential for advancing DOE mission-critical science drivers, including analyzing data from large-scale experimental facilities, optimizing fusion reactor performance, accelerating materials discovery, enhancing geophysical imaging, improving wildfire predictions, and enabling autonomous systems and digital twins.

However, these problems are becoming increasingly complex, often involving nonlinear, high-dimensional, and interconnected systems and models that span multiple physics and scales, while relying on data with varying quantity, quality, and information content. Compounding these challenges is the uncertainty inherent in DOE-relevant systems, where errors in inputs, noise in data, incompleteness of data, and discrepancies between models and reality constrain the accuracy and precision of solutions. At the same time, the convergence of recent scientific computing trends—scientific machine learning, artificial intelligence, and computing advances such as exascale computing—is creating unprecedented opportunities for tackling these challenges.

The cross-cutting nature of inverse problems, combined with their growing complexity and rapidly evolving data and algorithmic demands, strongly motivates the formulation of a prioritized research agenda to maximize their capabilities and impact. In response to this need, DOE's Advanced Scientific Computing Research (ASCR) program in the Office of Science convened the Workshop on Basic Research Needs for Inverse Problems for Complex Systems Under Uncertainty in June 2025. This workshop brought together experts across disciplines to identify grand challenges and major opportunities in the field. Through collaborative discussions, the workshop defined transformative research directions aimed at addressing the mathematical, statistical, and computational challenges posed by inverse problems under uncertainty.

As a result of these efforts, four priority research directions (PRDs) were identified to guide future research and development in this area. These PRDs, summarized below, represent a roadmap for advancing the foundational science and mathematics of inverse problems, enabling robust, scalable, and uncertainty-aware solutions that are critical for DOE applications.

PRD 1: Discovering, exploiting, and preserving structure

Unlocking the next generation of DOE scientific discoveries will require innovative approaches for solving inverse problems that address escalating computational demands, overcome greater degrees of ill-posedness, and efficiently utilize high-performance computing and data resources. To meet this need, there are substantial opportunities to exploit new forms of mathematical, physical, and data structures inherent in emerging complex systems to reduce problem size and complexity, introduce informative priors and constraints, improve robustness and tractability, and accelerate computational performance. Achieving this will require fundamentally new methods for solving inverse problems that better leverage known structures, discover new exploitable structures, preserve essential mathematical and physical properties, and map computational and communication patterns to these structures to optimize performance on available computing and data infrastructure.

PRD 2: Identifying and overcoming model limitations

Forward models used in inverse problems are often constrained by simplified assumptions, structural discrepancies, incomplete representations of physical processes, imperfect likelihoods, limited simulator fidelity, or inadequate priors. These limitations, also known as misspecifications, compromise the reliability, efficiency, and interpretability of solutions while introducing significant uncertainty into the inference process. Addressing these challenges requires the development of mathematical frameworks to diagnose, quantify, and mitigate model inadequacy during inference. An important component of this effort will be the design of methods that provide robust uncertainty estimates consistent with the data, account for both known and unknown limitations—such as those represented by a range of possible distributions—and enhance the accuracy of inverse solutions across varying problem scales.

PRD 3: Integrating disparate multimodal and/or dynamic data

Recent advances in data acquisition technologies have enabled measurements of complex phenomena and structures across vastly different sources, physics, domains, and scales. However, current inversion algorithms often struggle to integrate such disparate multimodal or dynamic data, which can be heterogeneous, inconsistent, and vary in fidelity, volume, and sparsity, while also contending with the use of multiple models that have conflicting representations of physics and data, or the storage and processing of data across different facilities. Effectively utilizing disparate data, models, and facilities to solve inverse problems at DOE-relevant scales will require algorithms that capture common features in the data, filter noise, fuse multiple fidelities and scales, bridge distinct representations, and quantify correlated uncertainties. Methods are also needed to dynamically assimilate data, update rapidly evolving posterior distributions, and address abrupt changes in nonlinear, multiscale, multiphysics, or chaotic systems.

PRD 4: Solving goal-oriented inverse problems for downstream tasks

Inverse problems have traditionally focused on inferring model parameters from data, but they are increasingly being integrated into downstream tasks such as control, design, certification, and decision support. This shift creates a pressing need for goal-oriented inverse problems (GIPs), which reformulate inverse problems to directly align with the objectives of downstream tasks, enabling improved predictions, reduced uncertainty, and enhanced decision-making for complex systems on actionable timescales. Realizing the potential of GIPs will require the codesign of scalable algorithms for inverse problems and downstream tasks, including methods to quantify uncertainty in task outcomes, exploit parameter-to-task relationships, autonomously steer experimental data collection, efficiently allocate computational resources, adapt model structure or fidelities, support "what-if" decision-making, and implement risk-aware approaches tailored to stakeholder needs.

These four priority research directions represent a cohesive vision for advancing the science of inverse problems under uncertainty. Together, they address the critical challenges of discovering, exploiting, and preserving physical and mathematical structure, overcoming model limitations, integrating disparate multimodal and/or dynamic data, and tailoring the solution of inverse problems to downstream tasks. While each PRD focuses on a distinct aspect of inverse-problem research, their interconnected nature highlights the importance of a holistic approach that leverages progress across all areas to achieve transformative solutions.

By identifying the challenges inherent in solving inverse problems under uncertainty and the opportunities to overcome them, this report establishes a foundation for transformative advances that will enable robust, scalable, and uncertainty-aware solutions to complex systems. The proposed research agenda spans mathematics, statistics, and computer science disciplines, guided and enriched by rapid advances in artificial intelligence, high-performance computing, and experimental facilities. Together, these efforts aim to unlock new capabilities, maximize scientific impact, and address the growing demands of inverse problems, which are central to DOE's mission. These advances are expected to drive progress across the breadth of DOE mission-critical science drivers, including energy sciences, materials discovery, autonomous systems, and beyond, ensuring that the DOE remains at the forefront of scientific innovation and technological breakthroughs.

1. Introduction

Inverse problems, which aim to infer unknown properties of a system using experimental and observational data through the use of a forward model, are central to many of the U.S. Department of Energy's (DOE's) most critical scientific and engineering challenges. The cross-cutting nature of inverse problems, coupled with their increasing complexity and rapidly evolving data and algorithmic requirements, underscores the urgent need for transformative advances in the mathematical, statistical, and computational foundations for solving inverse problems under uncertainty. These approaches must rigorously account for uncertainties and limitations in data, models, and assumptions, and systematically propagate this information to fully characterize its impact on solutions. Recent breakthroughs in artificial intelligence (AI), high-performance computing (HPC), and advanced experimental facilities offer unprecedented opportunities, but realizing their full potential in inverse problems will require innovative mathematical strategies. Such advances are crucial for unlocking new capabilities, maximizing scientific impact, and meeting the expanding needs of DOE science initiatives.

1.1 Scope

Recognizing the urgency and importance of developing innovative methods for solving inverse problems, DOE's Advanced Scientific Computing Research (ASCR) program in the Office of Science convened the Workshop on Basic Research Needs for Inverse Problems for Complex Systems under Uncertainty in June 2025. This workshop brought together leading experts from diverse disciplines to identify grand challenges and major opportunities in the field. Through collaborative discussions, the workshop focused on defining transformative research directions that address the mathematical, statistical, and computational challenges posed by inverse problems under uncertainty. These discussions were informed by position papers submitted by researchers, which articulated key challenges and opportunities and helped shape the workshop agenda.

This report synthesizes the insights and discussions from the workshop, presenting a detailed exploration of the priority research directions (PRDs) that were identified. These PRDs provide a strategy for advancing the field and are designed to enable robust, scalable, and uncertainty-aware solutions to inverse problems, which are essential to address the evolving needs of mission-driven science priorities within the DOE. The report offers a comprehensive discussion of the existing and anticipated challenges posed by solving inverse problems for complex systems under uncertainty, complemented by short science vignettes that illustrate these challenges and the transformative impact of addressing them. It highlights opportunities for overcoming these challenges and recent innovations that demonstrate the readiness of the research community to successfully execute these PRDs. By summarizing recent progress and establishing the foundations for transformative advances, the report aims to drive progress in solving complex inverse problems to support DOE's mission and ensure its continued leadership in scientific innovation and technological breakthroughs.

1.2 Definitions

This section defines several key terms used in the report that are essential for describing the challenges and opportunities in solving inverse problems for complex systems under uncertainty.

A *forward model* is a mathematical or computational representation of a (partially) observable system that predicts the system's output based on its inputs and parameters. An *inverse problem*

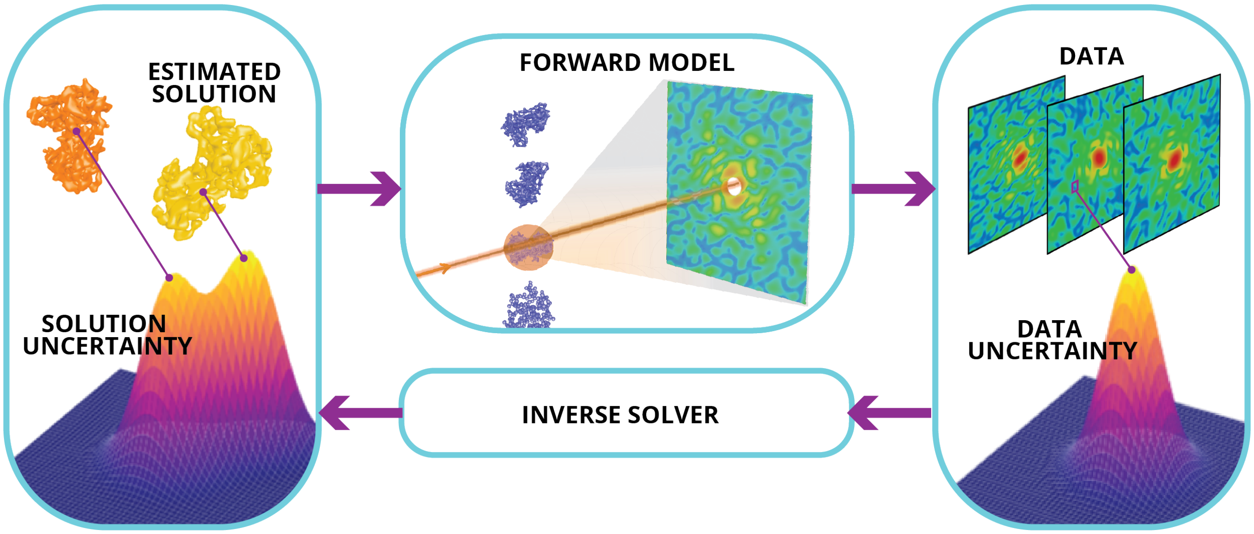

Figure 1: Conceptual illustration of a prototypical inverse problem. Starting from an initial solution estimate, the inverse solver infers unobserved quantities of a system—the solution—from observed data by ensuring that the forward model predictions align with the data, prior knowledge, and constraints. The problem's ill-posedness, along with uncertainties in prior knowledge, models, and data, can result in multiple plausible solutions. In this example, the inverse problem aims to reconstruct a 3D protein structure from a large number of 2D single-particle diffraction images obtained in an X-ray free-electron laser experiment. Image courtesy of Jeffrey Donatelli, Lawrence Berkeley National Laboratory.

seeks to infer the inputs or parameters of a system from observed outputs, aiming to identify input values that make the forward model outputs consistent with the observations. A problem is considered *ill-posed* if there is insufficient data or modeling information to yield a unique solution that is stable with respect to small perturbations in the data or model. *Priors*, *constraints*, and *regularization*, which mathematically encode prior knowledge about the solution or other characteristics of the system, are often used to constrain and stabilize the solution, thus mitigating ill-posedness.

Complex systems are sets of interconnected components whose relationships and collective behavior are not easily predictable from the properties of individual parts. Such systems can be physical, computational, biological, social, or engineered, and are often characterized by strongly coupled components, multiple interacting processes or physical laws operating across widely varying spatial and temporal scales, nonlinear interactions, and sensitivity to changing conditions. These characteristics frequently lead to algorithmic challenges such as nonconvexity, ill-posedness, and high dimensionality, making the solution of inverse problems for complex systems particularly difficult.

Uncertainty refers to the lack of complete knowledge about or predictability of an outcome, measurement, or state of a system, model, or process. Uncertainty stems from a variety of sources, including errors in system inputs, noise or gaps in experimental and observational data, and problem *misspecification*, such as incorrect or inaccurate priors, noise models, or forward models that fail to fully capture real-world phenomena. Such uncertainties can be *aleatoric*, arising from inherent randomness or variability in the system and considered irreducible, or *epistemic*, stemming from incomplete knowledge or information about the system and considered reducible. In the context of inverse problems, these uncertainties can lead to multiple plausible solutions, significant errors, and reduced reliability and interpretability, all of which are significantly magnified when such problems are ill-posed.

The process of solving an inverse problem is conceptually illustrated in Figure 1. It involves estimating the inputs of a forward model so that its outputs are consistent with observed data while being consistent with prior knowledge. The particular formulation of an inverse problem depends on the application: for example, model-parameter inference, which uses observed data to estimate unknown parameters in complex mathematical models, such as coefficients and initial/boundary conditions in differential equations; computational imaging, which reconstructs high-resolution 2D or 3D images of objects from indirect measurements, such as scattering or microscopy; and data assimilation, which estimates the evolving state of a physical system by combining partial or noisy observational data with incomplete mathematical models, as in weather forecasting and digital twin updates.

Inverse problems for complex systems under uncertainty can be solved with both *probabilistic* approaches, which characterize all possible solutions through probability distributions (such as Bayesian inference), and *deterministic* methods (such as maximum likelihood estimation), which aim to identify a single best solution. Examples of probabilistic approaches for solving inverse problems include *simulation-based inference* (SBI), *generative AI* (genAI) methods, *variational inference*, *transport maps*, and *Markov Chain Monte Carlo* (MCMC) methods. These methods can also be integrated into *goal-oriented inverse problems* (GIPs), which tailor the solution process to downstream tasks, such as optimal design or control, by adjusting inputs to achieve specific desired outcomes.

Additionally, many inverse problems relevant to the DOE are inherently *large-scale*, involving multiple challenging aspects. Forward models often have high-dimensional state spaces and can be computationally expensive to evaluate. These problems may include a large number of observations, each of which could be high-dimensional, such as high-resolution spatio-temporal data, and solutions that themselves may span high-dimensional spaces. The exact notion of "large-scale" is problem-dependent, but high-dimensionality can easily involve $O(10^6)$ variables, and a single simulation of a computationally expensive forward model can take up to $O(10^6)$ CPU hours. Consequently, it is essential to develop *scalable* methods—approaches designed to efficiently handle increasing problem size and complexity while maintaining computational feasibility and accuracy—that can address these multifaceted challenges effectively.

The following section describes several examples of inverse problems that are important to applications spanning the entirety of the DOE mission space.

1.3 Science Drivers

Accurately solving inverse problems under uncertainty is vital for achieving DOE goals across a wide range of complex systems within its mission space. In this section, a series of vignettes are presented that illustrate the broad importance of inverse problems and the enduring impact of solving them accurately. These vignettes highlight how advances in solving inverse problems can drive progress in areas ranging from energy security to materials discovery. Together, they underscore the critical role of developing foundational, broadly applicable mathematical methods for solving inverse problems in shaping the future of DOE research and development.

Reconstructing 3D Structure from Light-, Electron-, and Neutron-Source Experiments

DOE's investments in advanced scientific user facilities are enabling exploration across materials science, chemistry, physics, and biology. Extracting insights from experiments often requires solving inverse problems, such as reconstructing 3D macromolecular structures from single-particle diffraction data [1] (see Figure 2). Challenges include sparsely sampled data, low signal-to-noise ratios, and high computational demands. New initiatives under the American Science Cloud initiative, including Integrated Research Infrastructure (IRI) [2] and High-Performance Data Facility (HPDF) [3], offer opportunities to integrate multimodal data and enable autonomous experimentation, but current methods need advancement to fully leverage these capabilities. Progress in inverse problem-solving is key to unlocking the full potential of DOE facilities to drive scientific breakthroughs.

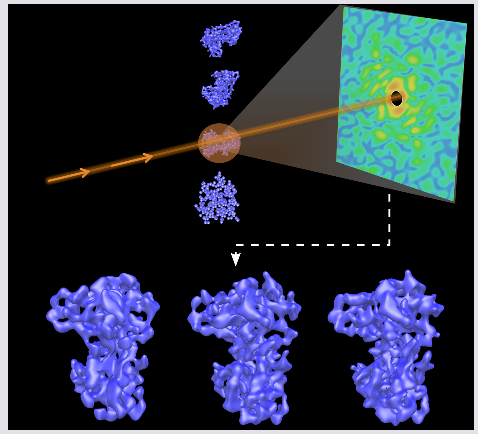

Figure 2: 3D reconstruction of a protein from a large number of single-particle diffraction images measured in an X-ray free-electron laser experiment. Image courtesy of Jeffrey Donatelli, Lawrence Berkeley National Laboratory.

Inferring Plasma States of Fusion Energy Systems

Fusion energy, exemplified by systems like tokamaks (see Figure 3) and the National Ignition facility, has the potential to address the increasing energy demands of the U.S. in the coming decades. Achieving this requires solving inverse problems to infer plasma properties, such as current distributions, from observational data like magnetic measurements, enabling precise control of fusion systems and minimizing instabilities [4]. However, achieving large-scale magnetic fusion energy and inertial fusion energy presents significant challenges, including the need to efficiently and repeatedly solve inverse problems to estimate system states and enable real-time autonomous control, while simultaneously modeling high-dimensional plasma phenomena across extreme scales, addressing strong nonlinear interactions, and overcoming uncertainties in data and material properties. Addressing these challenges is essential for designing reliable fusion systems, preventing costly reactor damage, and advancing U.S. energy independence.

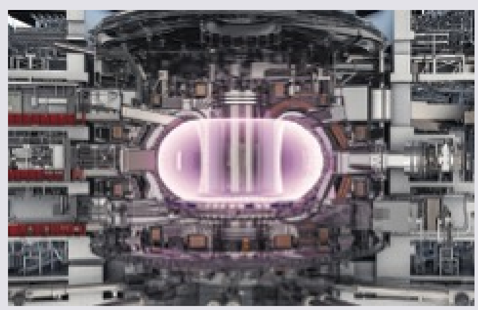

Figure 3: An illustration of a tokamak containing plasma. Achieving stable fusion requires accurately inferring plasma states from observational data and leveraging this information to enable reliable, real-time autonomous control of the reactor. Image credit: Oak Ridge National Laboratory, shared under the Creative Commons license: https://en.wikipedia.org/wiki/en:Creative_Commons.

Seismically-Informed Digital Twins for Underground Energy Production

Underground energy exploration and production—exemplified by oil and gas exploration, enhanced oil recovery (EOR, see Figure 4), and geothermal production—play a critical role in meeting the U.S.'s growing energy demands. Realizing low-risk energy production in these settings, such as CO_2-based EOR [5] in conventional oil fields and fracking, requires the inference of high-resolution subsurface characterizations from limited surface seismic data governed by complex wave physics. The challenges include developing multiphysics digital twins [6] based on goal-oriented inverse problems that integrate real-time data assimilation with monitoring, multiphase flow, risk assessment, decision making, and control, while rigorously quantifying uncertainty and accounting for model misspecification. Overcoming these challenges is essential to ensure efficient, sustainable, and low-risk underground energy production.

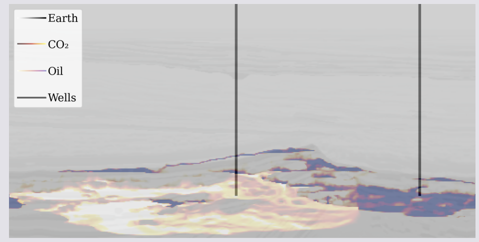

Figure 4: Multi-flow simulations for CO_2 and oil concentrations. The Earth properties model (gray) is overlain with the simulated concentrations. Multiphysics digital twins are used to infer these concentrations from time-lapse seismic data to enhance CO_2-based oil recovery [5]. Image courtesy of Georgia Institute of Technology's Seismic Laboratory for Imaging and Modeling, shared under the Creative Commons license: https://en.wikipedia.org/wiki/en:Creative_Commons.

Inferring New Fundamental Particle Properties

High Energy Physics (HEP) research has led to numerous groundbreaking discoveries, from identifying the Higgs boson to revealing the accelerating expansion of the Universe, as well as significant technical advancements, particularly in medicine. Enabling future high-impact insights requires solving inverse problems to infer the properties of particles produced in collisions from signals captured by detectors surrounding the collision events [7] (see Figure 5). A key challenge is assimilating high-dimensional data at extremely small time scales, often on the order of sub-microseconds or less. Advancing the solutions to HEP inverse problems will pave the way for new foundational discoveries about the Universe and potential applications spanning medicine, technology, and national security.

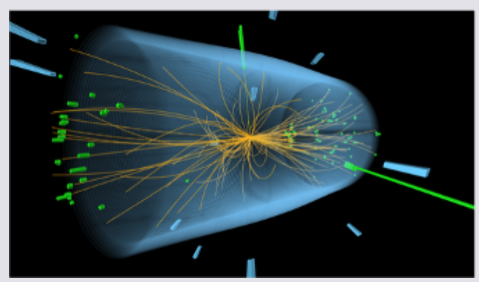

Figure 5: Inference of Higgs boson properties from observed high-energy particle-collision events. Image credit: Conseil Européen pour la Recherche Nucléaire (CERN).

Inverse Design of Materials

The discovery of new materials is essential for advancing technologies such as medical devices, power generation, microelectronics, quantum computing, and imaging systems. Materials discovery relies on solving multiobjective and often constrained inverse problems across vast and complex search spaces to identify manufacturable compositions with target properties like minimal cost, maximal strength, or specific optical performance [8–10] (see Figure 6). This process is particularly challenging due to the reliance on sparse data spanning multiple modalities, the computational expense of numerical simulations, and the huge cost and time required for experimental investigations [11], all of which is compounded by the sequential multistep nature of the process, where the solutions to one inverse problem often feed into another. Overcoming these challenges will accelerate the discovery of new materials capable of transforming manufacturing, transportation, healthcare, energy, and computing technologies.

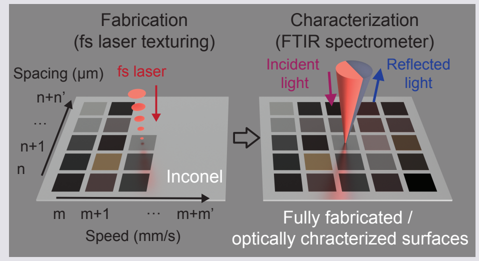

Figure 6: High throughput femtosecond laser fabrication and optical property characterization using Fourier Transform Infrared spectrometer. The process involves inferring laser fabrication parameters corresponding to desired target optical properties [8]. Image courtesy of Juliane Mueller, National Renewable Energy Laboratory.

Reconstructing Nuclear Structure

Nuclear physics, particularly the study of atomic nuclei structures, has the potential to deepen our understanding of fundamental interactions and inform applications in energy, medicine, and national security. Central to this understanding is the solution of inverse problems to infer nuclear properties, such as neutron densities, from experimental data like proton scattering measurements, enabling precise modeling of nuclear systems and improving predictive capabilities [12]. However, developing a global description of nuclei that is valid across the nuclear chart poses significant challenges, including the integration of computationally expensive simulations based on density functional theory (see Figure 7) with noisy experimental data, all while rigorously quantifying uncertainty. Addressing these challenges is essential for refining energy density functionals, advancing nuclear theory, and supporting critical applications in science and technology [13].

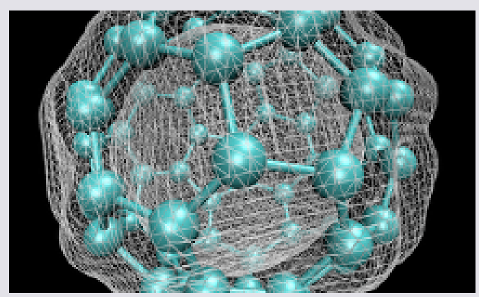

Figure 7: Depiction of a ground-state electron density calculated with density functional theory. Image credit: Isaac Tamblyn, University of Ottawa, shared under the Creative Commons license: `https://en.wikipedia.org/wiki/en:Creative_Commons`.

Data Assimilation for Wildfire Mapping

Wildfires are catastrophic events that cause massive human and economic loss and have devastating impacts on communities in the wildland urban interface, home to over 44 million residences. The economic costs of the 2025 Los Angeles wildfires are estimated to be in the range of 100–250 billion USD [15]. Solving inverse problems to infer wildfire behavior from sparse and noisy observational data collected by satellites, drones, and ground-based sensors is critical for improving predictions of wildfire spread and enabling better emergency response [14, 16] (see Figure 8). These problems are highly challenging due to their nonlinear, chaotic dynamics, high dimensionality, and the need to integrate multimodal data in real time. Advancing inverse modeling methods will improve wildfire predictions, quantify fire risk, and optimize emergency response strategies to save lives and minimize destruction.

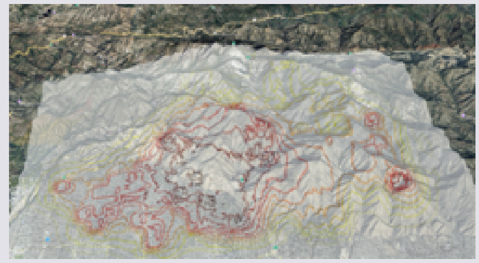

Figure 8: Inference of fire perimeters at 4-hour intervals over the first 48 hours of the 2025 Eaton Fire in Los Angeles. Earlier times are denoted by darker colors. These perimeters, derived from sparse satellite observations [14], provide a basis for evaluating fire progression and may support the planning of firefighting operations. Image courtesy of Assad Oberai, University of Southern California.

Inferring Quantum System Properties

Quantum information science (QIS) promises transformative technologies in computing, sensing, and networking, with applications such as secure communication, quantum-enhanced simulation for materials discovery, and advanced manufacturing [18]. Solving inverse problems to infer quantum states, system parameters, and operational characteristics from indirect, noisy, and incomplete measurements is essential for enabling these breakthroughs. These problems are challenging due to their ill-posed nature, high-dimensional complexity, nonlinear dynamics, and sparse data. Addressing these challenges will unlock the full potential of QIS, driving advances in cybersecurity, materials discovery, and manufacturing design and certification across the DOE mission space [19].

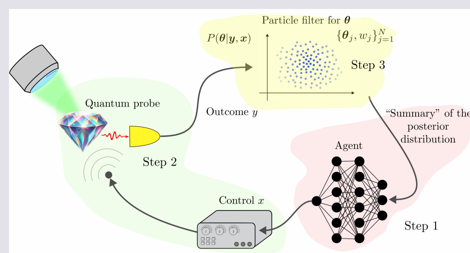

Figure 9: Optimizing quantum metrology and estimation through optimal experimental design. Image from [17] and shared under the Creative Commons Attribution 4.0 International (CC BY 4.0) license https://creativecommons.org/licenses/by/4.0/.

These vignettes underscore the critical role inverse problems play across a spectrum of DOE science areas. While inverse problems are central to understanding and optimizing complex systems, their growing complexity—driven by multiphysics, multiscale models, multimodal data, high-dimensional state spaces, and model limitations, such as missing physics in simulation models—demands transformative innovations in their formulation and solution. The following section outlines four priority research directions to overcome these challenges by improving the accuracy, computational efficiency, and uncertainty quantification of inverse problem solutions, ensuring that the DOE is equipped with the tools and knowledge needed to tackle the most pressing scientific and engineering challenges of the coming decades.

2. Priority Research Directions

This section describes the PRDs identified at the Workshop on Basic Research Needs for Inverse Problems for Complex Systems under Uncertainty. These PRDs lay the groundwork for transformative advances that will shape the future of methodologies for solving inverse problems, directly addressing emerging DOE mission-critical science needs. The convergence of ASCR investments into extensive computational capabilities and infrastructure, rapid advances in scientific machine learning and AI, and the increasing complexity of scientific data and analysis across DOE Office of Science programs make this a uniquely opportune time to tackle longstanding, emerging, and future challenges in inverse problems.

Tackling these PRDs is now feasible in large part due to the extensive capabilities established by decades of ASCR research initiatives across mathematics, computer science, computational science, supercomputing, and data management. For example, ASCR's pioneering research in large-scale modeling of complex systems [20] enables the use of sophisticated forward models for inverse problems, while progress in randomized algorithms [21] and data reduction [22] provides fundamental tools for efficiently tackling otherwise computationally intractable problems. ASCR initiatives in scientific machine learning [23] and AI for science [24] have created new pathways for incorporating novel data-driven insights into inverse problem frameworks. Moreover, ASCR's advancement of exascale computing [25, 26], scientific data management [27–29], codesign [30], and integrated research infrastructure [2, 3] offer the potential to scale inverse solvers to unprecedentedly large problems and integrate novel combinations of data and computing across different facilities. Finally, the expansive multidisciplinary and multi-institutional partnerships fostered by the Scientific Discovery through Advanced Computing (SciDAC) program create unique opportunities to align the development of new inverse problem capabilities with mission-critical needs across the DOE science programs and to rapidly deploy these advancements.

Fully leveraging these capabilities for inverse problems will require a coordinated effort to advance fundamentally new mathematical, statistical, and computational research to close key foundational gaps in the field. For instance, many inverse problems remain ill-posed or computationally intractable—even with abundant data and computing power—necessitating innovative approaches that exploit underlying mathematical and physical structures of the problems. Moreover, the incompleteness and inaccuracy of models are seldom well understood and can severely limit inference accuracy, highlighting the need for transformative methods that detect, quantify, and mitigate uncertainty by leveraging all available information during the inversion process. Current inversion methods struggle to integrate highly disparate and complex multimodal or dynamic datasets, calling for new strategies that bridge vast differences in scale, fidelity, and physics, and accommodate rapidly evolving data. Furthermore, most existing inversion approaches prioritize solution accuracy over the needs of downstream tasks, which can significantly limit overall performance. Overcoming this limitation will require foundational advances that reformulate inverse problems to directly address task-specific objectives. Addressing these challenges will revolutionize how inverse problems are solved, enabling robust, scalable, and uncertainty-aware solutions critical to advancing DOE scientific discovery.

PRD 1 Discovering, exploiting, and preserving structure

Motivation and Impact. As DOE science problems grow in scale and complexity, inverse problems face significant barriers, including prohibitive computational and data costs, pervasive noise, measurement limitations, lack of prior knowledge, heightened sensitivity to model parameters, and the demand for accurate results on actionable timescales. These issues often render inverse problems computationally intractable or ill-posed, challenges that cannot be resolved solely through advances in computing power or data collection. However, many complex inverse problems possess rich structure in their underlying mathematics, physics, and data—such as relationships encoded in governing equations, physical laws, numerical representations, and properties of solution spaces—that can be exploited to reduce solution error and simplify computation. Examples include reduced-order models, low-dimensional latent representations, domain decompositions for parallelization, rank structure and sparsity in large operators, multilevel and multifidelity decompositions, and constraints or priors on solution and data spaces. However, open questions remain on how to optimally exploit these structures in emerging complex systems to enhance scalability, computational performance, and regularization of ill-posedness. There are also significant opportunities to discover and utilize new forms of hidden structure; for example, through advances in AI surrogates, AI data reduction, and genAI.

New methods for discovering, exploiting, and preserving structure within inverse problems will significantly reduce computational complexity to achieve near-real-time results and allow the solution of problems that are currently computationally intractable even on large HPC machines. These structures can also be used as additional priors or constraints to enable the solution of inverse problems that are currently too ill-posed to provide a unique and stable solution; for example, due to data collection limitations or noise. Furthermore, to ensure physically meaningful and stable solutions, reduced representations and approximations must preserve key physical structures (e.g., conservation laws, nonnegativity, or physical bounds) and mathematical structures (e.g., governing equations, positive definiteness, or equivariance), while carefully estimating, propagating, and controlling the errors and uncertainties introduced by these approximations.

Thrust A: How can we better exploit and preserve the structure inherent in emerging complex systems?

Challenges and Gaps. While many types of structure are known, it is often unclear how to best exploit them in complex systems, and existing approaches often have limited performance, introduce excessive error, or fail to preserve important properties. For example, current methods for reduced-order modeling, sketching, and compression struggle to handle highly nonlinear or high-rank systems [31], are unable to overcome the intractability of many likelihood functions [32], often fail to preserve key physical or mathematical properties [33–36], and can accumulate significant errors within iterative inverse solvers [37]. Additionally, existing approaches to regularizing ill-posed problems by enforcing structures as priors or constraints are often insufficient for highly data-deficient systems, may be inconsistent with the true solution, and struggle to balance multiple competing priors and information sources, leading to bias and inaccuracy [38, 39]. Furthermore, scaling inverse solvers on advanced compute and data platforms remains challenging, especially when forward operators are tightly coupled across space, time, or physical processes, or when architectures are complex [40–43]. These factors make domain decomposition and optimization of communication and computation difficult. Finally, uncertainty propagation through these structures is often not well understood, and establishing rigorous error bounds and adapting reduced models to evolving speed and accuracy requirements remain persistent challenges [44, 45].

Why Now? Several emerging needs and recent advances make research into new methods for exploiting structure in inverse problems especially timely. New imaging technologies can capture high-resolution, multicomponent, and dynamic information, but are hampered by measurement limitations, noise, and other inaccuracies, resulting in highly underconstrained and ill-posed inverse problems that current methods do not adequately address [46–50]. Cutting-edge experiments enabled by upgrades to DOE scientific user facilities involve increasingly complex physics, which may be computationally infeasible to simulate within inverse problems without more effective and targeted model reduction strategies [51–59]. Progress in compute and data infrastructures—including edge, autonomous, cloud, federated, exascale, and quantum platforms—further necessitates inversion algorithms that can handle compressed and streaming data while optimally leveraging complex hardware architectures [2]. Novel model reduction techniques, such as those inspired by quantum algorithms [60–65] or advanced machine learning architectures [66–75], offer substantial computational gains for forward modeling. However, these approaches often fail to preserve mathematical structures and physical conservation laws, leading to unsustainable error accumulation in iterative inverse solvers. Recent developments in randomized numerical linear algebra [76–79] and sketching [80] have demonstrated promise in efficiently exploiting high-dimensional structure, yet generalizing these methods to nonlinear problems and ensuring the preservation of essential physical and mathematical properties remains challenging. Advances in optimization over manifolds [81] and other nonsmooth forms [82–87] pave the way for efficiently imposing a broader class of regularization priors, but have yet to be fully realized.

Research Opportunities. New research is needed to fully exploit structure in inverse problems to address emerging challenges in computation, ill-posedness, and the effective use of computing and data resources. Developing new structure-preserving model reduction techniques and AI architectures, or methods to correct for structure violations, could dramatically reduce computational costs while maintaining accuracy. These approaches should be integrated with rigorous error guarantees, methods to propagate uncertainty through the reduced models, and adaptive reduction, compression, and multifidelity computing strategies that adaptively balance computational cost and accuracy. Improving utilization and scalability on advanced computing and data platforms will require new strategies that align computation and communication patterns with the inherent structures of inverse problems, admissible problem/domain decompositions, and the underlying hardware architecture. To overcome pervasive ill-posedness, research is needed on mathematically formulating new, expressive forms of priors and constraints on solution and data spaces, estimating their agreement with the true solution and data, and balancing them using quantified uncertainties from all available information sources. Central to these efforts will be the development of new nonsmooth and potentially derivative-free optimization algorithms, inexact gradient methods, higher-order methods, and solution strategies that decompose inverse problems along their substructures and available hardware, all designed to provide provable convergence rates and fully leverage priors, surrogate models, and reduction schemes while mitigating their induced biases.

Thrust B: How can we discover and utilize new exploitable structures hidden in complex systems?

Challenges and Gaps. Many structures in complex inverse problems are inherently unknown, constantly evolving, or too complex to identify and exploit a priori. This demands new data-driven approaches that can automatically discover and leverage such hidden or difficult-to-formulate structures. For example, while many problems could be simplified by operating in low-dimensional latent spaces, current methods often struggle to robustly identify these spaces and reformulate operations in them in the context of inverse problems, particularly for nonlinear cases [88]. Additionally, traditional hand-crafted priors are frequently insufficient to address the ill-posedness of emerging data-deficient inverse problems, and much of domain knowledge lacks a natural mathematical formulation for exploitation. Although learning priors from simulations, previous measurements, and databases holds promise [89], integrating these into inverse problems efficiently remains challenging due to the computational cost of AI architectures, issues with generalizability, potential bias, and the lack of specialized architectures for inverse problems [90, 91]. Structures in data (e.g., sparsity, low-rank, tensor patterns, and relationships between multiple models and data sources) are often not known in advance, and distinguishing true structure from noise is difficult, leading to either overfitting or missed opportunities to improve performance [92]. Furthermore, understanding how noise and uncertainty affect latent features, reliably separating true features from artifacts, and ensuring the generalizability and unbiasedness of discovered structures remain ongoing challenges [93, 94].

Why Now? Rapid increases in the complexity of observable systems and the volume of data they generate are surpassing our ability to identify exploitable structures a priori, creating an urgent need for methods that can automatically detect and leverage unknown or hidden structures. Moreover, as systems become increasingly nonlinear, high-dimensional, and interconnected, traditional approaches relying on expert-crafted priors and predefined models are proving insufficient, underscoring the necessity of innovative, data-driven techniques to address these challenges effectively. Recent advances in AI architectures and genAI techniques show great promise for uncovering low-dimensional latent spaces in models and data [95–100], creating significant opportunities to leverage these techniques within inverse problems. While operator-learning [101, 102] and machine-learning surrogate models [103–106] have enabled efficient representations of forward models, exciting avenues remain for inverse problems, such as scaling to high-dimensional parameter spaces, capturing intricate nonlinearities, and preserving information essential for accurate inversion. Emerging research in learned priors has demonstrated their potential to enhance the accuracy and stability of solutions to ill-posed problems by capturing solution manifolds beyond traditional priors [107–110] but can fail to extract the most salient information, mitigate bias, leverage representations from large models efficiently, or generalize across multiple problem domains. Additionally, for well-posed inverse problems, amortized SBI [111] variants offer the potential to reduce inference costs and scale to large datasets by learning mappings from observations to posteriors across many simulations, but can require an impractical amount of training data.

Research Opportunities. New research directions are needed to unlock the full potential of structure discovery in inverse problems. Approaches that adaptively learn low-dimensional latent spaces and reformulate more computationally efficient versions of forward operators, solutions, and data in these reduced spaces could greatly accelerate a wide range of applications. Effectively and reliably leveraging learned priors will require new statistical and AI frameworks capable of capturing salient, generalizable prior information; strategies for efficiently employing large foundation-like models as priors across domains; robust methods for assessing their trustworthiness and bias; and seamless integration into broader inverse problem frameworks. There is also a need for inversion methods capable of automatically detecting and adapting computation to hidden structures in data, such as sparsity, rank structure, and relationships between multiple data sources and models. Advances in operator learning, graph-learning, sparsity-promoting techniques, and hybrid symbolic-statistical approaches could further expand the toolkit for identifying and exploiting meaningful structures. Furthermore, algorithms that align discovered structures of the inverse problem with the hardware layout of advanced compute and data architectures would create an automated approach to performing computationally effective domain decomposition, reducing required communication, and allocating resources, thus making scaling on emerging complex hardware more accessible. Ultimately, addressing these challenges will benefit from hybrid approaches to inverse problems that integrate AI with physics-based models and mathematical operations, together with theoretical advances that rigorously quantify uncertainty, provide robust error bounds, and ensure convergence guarantees when using these discovered structures.

PRD 2 Identifying and overcoming model limitations

Motivation and Impact. As DOE science confronts unprecedented complexity in phenomena and measurements while demanding ever-greater accuracy, the limitations of priors, likelihoods, and forward models are emerging as critical bottlenecks hindering scientific progress in many fields. Common sources of limitations, also known as *misspecification*, include oversimplifications, omitted variables, unmodeled noise, inaccurate priors, incomplete understanding of the underlying physics, poorly calibrated models, incorrect assumptions, and statistical bias. Moreover, the nature or even existence of misspecification is frequently unknown a priori and cannot always be overcome by simply improving the models, necessitating its direct treatment when solving inverse problems. Misspecifications can have dire consequences for inverse problem solutions—causing identifiability problems, introducing bias to the solution, and making the magnitude of uncertainties unwieldy or mathematically ill-defined—and are under-addressed in research to date. While progress has been made in addressing model misspecification for forward simulation and uncertainty quantification, these efforts primarily focus on improving predictive accuracy or quantifying uncertainty in forward models alone. In contrast, inverse problems pose unique challenges associated with propagating model errors through the inference pipeline, where they interact with misspecification in complex ways.

Fundamental advances in identifying, quantifying, and mitigating model limitations within inverse problems will transform the extraction of reliable insights from real-world systems that are too complex or uncertain to model faithfully. New methods capable of rigorously attributing uncertainty to sources, such as forward model errors, prior assumptions, out-of-distribution observations, and computational approximations, are urgently needed. These methods must also track the propagation of this uncertainty through the inference process and use this information to adaptively correct the model and solution. Such advances would revolutionize the accuracy and reliability of inverse-problem solving in domains where model limitations have long posed significant barriers. These advances would also substantially improve estimates of uncertainty and risk in solutions, thereby reducing reliance on overly conservative assumptions that may impede decision-making while supporting the safe and reliable operation of autonomous systems. The timeliness of this effort is underscored by the growing need of modern science drivers to study complicated phenomena that are not fully captured by existing models, coupled with increasing system complexity, both of which make the identification and mitigation of model misspecifications and limitations in inverse problems more challenging than ever. This urgency is further compounded by the growing availability of and need to leverage fast and expressive approximations—such as genAI—which can themselves introduce model misspecifications that adversely impact the accuracy and reliability of inverse problem solutions.

Thrust A: How can we adaptively correct, calibrate, and refine models during inversion to minimize misspecification?

Challenges and Gaps. Model limitations can significantly degrade the accuracy and reliability of solutions to the inverse problem. However, traditional methods for correcting model limitations often rely on linear, Gaussian, or additive corrections [112], which are typically inadequate for generating accurate solutions in complex systems. Additionally, these approaches primarily focus on constructing corrections to observation models, which cannot be used to improve predictions of unobserved quantities [113]. *Embedded enrichments*, or discrepancy models, add state-dependent modeling degrees of freedom to capture behaviors missing in the original model, improving consistency with data and uncertainty quantification of unobservable quantities [113–115]. However, these approaches have primarily been applied to forward modeling rather than solving inverse problems, and their formulation often requires significant subject-matter expertise, limiting their general applicability. Furthermore, even simple corrections introduce additional uncertain parameters which, when combined with poorly specified priors, can lead to identifiability issues [116, 117]. Moreover, ensuring the stability of dynamical systems or adherence to physical constraints when applying model corrections remains a significant challenge [118]. Measurement artifacts and calibration errors can significantly degrade solutions, but these effects can rarely be reliably identified and fully corrected. Although some issues can be readily identified and modeled, others—such as nonstatic, nonstationary, and nonuniform detector gain responses, fluctuating background, common modes, and correlated pixels—are notoriously difficult to detect, model, and correct, yet can significantly degrade solution quality [119–122]. Attempts to precorrect these measurement issues prior to inversion or postcorrect their effect on the solution after inversion are often inadequate [123, 124].

Why Now? The growing demand for higher resolution and accuracy in DOE science drivers is increasingly constrained by the limitations of imperfect models, highlighting the need for innovative approaches to correct, calibrate, and refine models when solving inverse problems. The rapid advancement of high-performance computing systems, coupled with their integration with experimental and observational facilities [2], provides an unprecedented opportunity to refine models and ensure robust uncertainty estimates in the presence of additional misspecifications. These advances, along with those in high-dimensional approximation techniques, have enabled the design of deterministic and stochastic embedded enrichments [118, 125], as well as learned forward model corrections [126], which improve forward prediction. Similarly, model discovery methods, such as symbolic approaches [127, 128] and neural operators [129, 130], have been used to improve constitutive models and closures. Despite progress in forward modeling, many of these tools remain underdeveloped for inverse problems [126], where uncertainties arising from limitations can be unexpectedly amplified and transformed. Although calibrating/correcting forward models prior to inversion is sometimes possible, recent work [131, 132] has shown that building these steps into the inversion process can be far more effective, and often necessary, particularly for complex systems or when the data is noisy/incomplete. However, such combined calibration/inversion may be unstable, and has been largely unexplored beyond a limited range of problems, leaving considerable room for further technique innovations. Although recent work has shown that isolating and modeling unstable components of forward models as additional degrees of freedom can significantly improve solution accuracy [133], this approach has so far been limited to cases where the forward operator can be linearly decomposed into separate stable and unstable terms. Additionally, deep imaging priors have been used to effectively correct localized, predictable measurement artifacts [134, 135], but their effectiveness diminishes in complex, nonlinear problems with global, correlated effects. Deep learning techniques have also been explored for precorrecting the data [135, 136] or postcorrecting the solution [137], but they lack integration with priors or forward models to assess correction quality, often leading to suboptimal performance.

Research Opportunities. Overcoming pervasive and emerging model limitations in inverse problems will require innovative research focused on correcting, calibrating, and refining models within inverse solvers. Methods that can learn model discrepancies or embedded enrichments that adhere to mathematical structures or physical principles, such as stability and conservation of energy, will be critical for ensuring accuracy during both model calibration and deployment. To produce reliable uncertainty estimates that do not underestimate errors in corrections, these efforts could be coupled with effective procedures for prior elicitation that constrain embedded parameters and align with a priori knowledge of the system, as well as generalized loss functions that introduce robustness into downstream tasks or focus on matching essential features that are insensitive to model misspecification. Additionally, methods are required to detect and adapt or switch between corrections when changes in the underlying data occur, enabling real-time inference and dynamic data assimilation, especially when using disparate multimodal data. Furthermore, holistic algorithmic frameworks are required to identify, model, and correct complex measurement artifacts—including nonstatic, nonstationary, and nonuniform detector gain responses, common modes, and spatially correlated pixel fluctuations—within the inversion process, enabling artifact correction to be guided by how it improves consistency between the data, forward model, and priors of the inverse problem. New methods are also needed to discover and isolate unstable or unneeded components of forward models, potentially by representing stable and unstable parts in separate latent spaces to produce parsimonious models that can balance the tradeoff between added expressivity and increased uncertainty. These advances must be supported by the development of standardized benchmarks for complex systems and rigorous theoretical foundations, including universal approximation guarantees and sample complexity guarantees for models with corrections, to improve trustworthiness and facilitate the verification and validation of solutions.

Thrust B: How can we design inverse solvers that quantify and mitigate the effects of pervasive model misspecification?

Challenges and Gaps. Correcting model limitations can improve accuracy in inverse problems; however, it does not eliminate them entirely or address other forms of misspecification. For example, likelihood distributions used in Bayesian inference for solving inverse problems may fail to fully characterize sophisticated noise in observed data, such as nonadditive, non-Gaussian, or dependent noise [138]. Similarly, prior distributions may inadequately represent the true characteristics of the solution space, such as physical constraints [139, 140]. For instance, when using generative models, model misspecification can lead to highly inaccurate posteriors [141]. Additionally, detecting and mitigating out-of-distribution experimental and observational data, which can degrade the reliability and accuracy of inverse solutions even when the problem is well-specified, remains challenging, especially outside of simplified settings [142, 143]. Furthermore, priors constructed from limited or biased datasets may distort the inverse solution or assign very low or even zero probability to the "true" data-generating parameters, thereby preventing their recovery during inversion [144].

Why Now? Rapid advances in experimental facilities and sensor technology are enabling the observation of increasingly complex phenomena, often exceeding the capabilities of current models to fully capture the underlying phenomenology. Additionally, AI surrogates [145, 146] and genAI [147–149] are increasingly used to accelerate high-fidelity models and construct priors, but can amplify misspecification and reduce robustness to misspecification, resulting in biased estimates or underestimated uncertainty. The shift toward autonomy, exemplified by digital twins, makes reliable solutions to inverse problems under misspecification increasingly critical. Algorithmically, recent efforts have demonstrated promising pathways for detecting and mitigating misspecification and validating uncertainty estimates [150–153], yet extending these techniques to high-dimensional, large-scale problems with computationally expensive forward operators remains elusive. These advances include methods that conservatively formulate, approximate, or sample posteriors to account for misspecification [152, 154], such as approaches based on robust summary statistics [149, 155], which focus on key features of the data that are less sensitive to misspecification, and marginal likelihood consistency-based regularization techniques [156]. However, constructing information-preserving summary statistics for high-dimensional, multimodal datasets remains challenging, as does balancing the relative contributions of the prior and likelihood, both of which critically impact their effectiveness. Generalized Bayesian inference [157, 158] provides robust alternatives to traditional posterior updates by modifying the likelihood function to account for model misspecification. However, these approaches face challenges such as the lack of a standardized methodology, reliance on problem-specific tuning, and increased computational complexity. Moreover, the promises of all the aforementioned directions have yet to be developed when mathematical or physical constraints must be enforced.

Research Opportunities. Ensuring reliable inferences from the observation and modeling of increasingly complex phenomena in the presence of model misspecification and out-of-distribution data will require transformative advancements in identifying and mitigating these sources of systematic and probabilistic error. For example, likelihood-free approaches that bypass the need for explicit knowledge of the true likelihood distribution will be critical to enable robust inference in scenarios where the data-generating process is complex or intractable, and allow inferences in the presence of known unknowns and unknown unknowns. Additionally, novel scalable methods that formulate and leverage robust summary statistics or alternative generalized loss functions could enable solutions to high-dimensional problems from complex systems that are currently out of reach. Techniques are needed to construct generative priors and apply them to inverse problems while guaranteeing nontrivial probability for the true data-generating parameters and ensuring robustness in the presence of prior misspecification. Methods that overcome all forms of misspecification while enforcing physical constraints will be vital for maintaining consistency with the underlying physics of the system while remaining adaptable to evolving model limitations, such as improving posterior inference by incorporating gradient information from differentiable simulators. Risk-aware approaches that tailor robustness measures to the risk preferences of inverse problem stakeholders are also needed to ensure solutions align with application-specific priorities. Importantly, all these methods must be accompanied by strong theoretical guarantees extending beyond linear and Gaussian settings to instill confidence in their foundational rigor. Furthermore, novel robust validation frameworks, generalizing existing posterior predictive diagnostics, and benchmarks will enable the systematic evaluation of methods, guide algorithm innovation, and foster greater trust in inverse solutions.

PRD 3 Integrating disparate multimodal and/or dynamic data

Motivation and Impact. Emerging DOE advances in experimental facilities and integrated research infrastructure, as exemplified by the American Science Cloud initiative, are creating unprecedented opportunities to collect and integrate increasingly complex data from multiple facilities across a broad range of spatial and temporal scales, physics regimes, and measurement modalities. However, existing inversion techniques often struggle to integrate the resulting multimodal or dynamic data when such data are exceedingly disparate. Multimodal data refers to information collected from multiple distinct sources, sensors, or measurement techniques that capture different aspects or properties of a system. Such data often encompass fundamentally different physics, scales, fidelities, and computational representations, making their integration in inverse problems exceptionally challenging. Additionally, assimilating dynamic data—updating inverse problem solutions with new observations—faces significant barriers in systems with high-dimensional nonlinear dynamics, abrupt state changes, and time-varying data quality, especially when rapid analysis is required for time-sensitive applications, such as digital twins or autonomous systems.

Resolving these challenges in handling multimodal and/or dynamic data will unlock new scientific capabilities by enabling the integration of novel combinations of data and models that were previously too disparate to bridge. By extracting greater information from heterogeneous data across domains and scales, these methods will achieve unprecedented solution accuracy and reveal patterns that emerge only through integrated analysis. For example, combining experimental measurements from scattering, microscopy, fluorescence, or spectroscopy with physical models has shown considerable promise for revealing new insights into material properties [159–162], while integrating diagnostic and simulation data is becoming increasingly important for predictive control of fusion devices [163, 164]. Furthermore, these advances will enable predictive models capable of capturing the extreme complexity and dynamics of natural and engineered systems—critical for DOE inverse problems ranging from extreme event prediction [165–167] to additive manufacturing [168]. Ultimately, these capabilities will help realize the full potential of the DOE's American Science Cloud initiative by enabling seamless analysis across multiple data sources and facilities, and will support real-time control and optimization of complex systems, particularly for digital twins that continuously learn and evolve in response to new data.

Thrust A: How can we effectively integrate multimodal data and models spanning vastly different scales, fidelities, sources, or underlying physics?

Challenges and Gaps. Solving inverse problems with disparate multimodal data poses unique challenges that cannot be resolved by combining standard unimodal techniques. Solution and data spaces, as well as forward models, often involve fundamentally different underlying physics, approximations, and computational representations across modalities, which makes unifying them into a single coherent mathematical formulation of the inverse problem a major open problem [169]. Furthermore, data from multiple modalities often require registration—such as alignment, rescaling, or geometric transformation—to match corresponding features, but this is challenging when data are noisy, incomplete, or measure fundamentally different information [170]. In such cases, performing registration independently of the inverse problem is often insufficient, as it overlooks valuable information inherent in the problem [171], and it is often unclear how to define an appropriate registration metric between heterogeneous data [172–174]. Balancing the contributions of modalities that differ in fidelity, volume, sparsity, and noise is another major hurdle, often causing either higher-quality data to dominate the analysis or lower-quality data to corrupt results [175]. Traditional methods based on loss or likelihood functions that use fixed weights or treat all modalities equally are unable to leverage the evolving and context-dependent relevance of each data source as the solution and system are updated. Additionally, inconsistent information or incompatible background and noise sources across modalities make it difficult to identify a solution that is consistent across all data [176]. Relationships between modalities are often poorly understood or difficult to model from first principles, further limiting the ability to exploit intermodal connections and accurately capture correlated uncertainties [177]. These challenges render many modalities too disparate to be robustly integrated by current inversion methods, leading to missed opportunities to extract new scientific insights that are only accessible when using multiple modalities [178].

Why Now? Advances in data collection and computing are opening the door to integrating previously unexplored combinations of multimodal data and models, while new ideas in data fusion are creating pathways to tackle multimodality in inverse problems. For instance, the DOE's American Science Cloud initiative, which includes the Integrated Research Infrastructure [2] and High Performance Data Facility [3] programs, will facilitate seamless workflows that connect data sources and computational resources across multiple facilities, creating unprecedented opportunities to integrate novel combinations of data and models. Recent developments in multimodal learning—including variational autoencoders [179, 180], self-supervised encoders [181], vision transformers [182, 183], and representation learning [184–187]—have demonstrated the ability to construct unifying latent spaces or embedding spaces that capture meaningful structure across heterogeneous modalities. However, the potential of these AI approaches remains largely untapped for general multimodal inverse problems, where more research is needed to scale these methods within iterative solvers and leverage additional information from the forward model, prior, and other constraints. Additionally, recent developments in feature-based [173] and deep-learning-based [188] registration have been instrumental in registering multimodal data during a preprocessing step. However, most of these techniques have yet to be directly integrated within inverse solvers, where leveraging information from the forward model and evolving solution estimates could enable more accurate registration. Notably, incorporating similar steps within the inversion process—rather than as a preliminary data fusion or registration stage—has been shown to be more efficient and accurate [171, 189, 190], though further research is needed to establish effective approaches for integrating the advances described above.

Research Opportunities. Realizing the full potential of emerging opportunities to integrate multimodal data and models will require several fundamental innovations. Modalities with fundamentally different physics or computational representations could finally be integrated through new unified computational frameworks capable of learning efficient mappings between representations, identifying correlations across modalities, and leveraging these relationships throughout the inversion process. Novel information obscured by noise and discrepancies between modalities could be revealed by new methods that harness all available data, models, and priors to identify shared structure, such as latent spaces, across modalities, and project data, priors, and operators into these spaces. Current limitations in multimodal registration could be overcome through innovative approaches that identify and match common features in heterogeneous data (e.g., via new registration metrics or AI approaches), optimize the registration process to maximize accuracy and stability of the inversion, and perform registration jointly with inversion to leverage information from evolving solution and system state estimates. New strategies are also needed to automatically balance data sources and priors according to fidelity, information content, and consistency, with promising approaches including new information-theoretic methods that dynamically quantify trust and consistency among information sources and update modality weights and constraints accordingly. Rigorous error bounds and convergence guarantees could be achieved through new information-theoretic and graph-based approaches to analyze and quantify the propagation of uncertainty and information throughout the intricate network of interconnected and correlated relationships present in complex multimodal inverse problems. Finally, new federated algorithms—which enable analysis across multiple sites without requiring direct sharing of raw data—that incorporate these advances would enable rapid large-scale integration of multimodal data from geographically dispersed experimental facilities, which is essential for time-sensitive applications such as autonomous experimentation and digital twins.

Thrust B: How can we dynamically assimilate disparate data from high-dimensional, nonlinear, rapidly evolving, and computationally demanding systems on actionable timescales?

Challenges and Gaps. Although the need for data assimilation is longstanding, rapidly evolving scientific drivers with greater complexity, larger data volumes, and diverse modalities necessitate a fundamental rethinking of current approaches. Kalman filters [191], including their ensemble variants, are computationally efficient, but rely on Gaussian assumptions to accelerate computations, which are often unrealistic for complex systems in practice. While particle filters [192] relax some of these assumptions, they often suffer from particle degeneracy, where most samples have low probability, or overly confident posteriors [193], thus requiring an impractically large number of particles, which may be computationally intractable. Particle filters often fail to effectively balance observational data and model information near sharp features and struggle in high-dimensional settings [194], particularly when attempting to represent posterior distributions with multiple peaks [195]. Moreover, computational costs can quickly become prohibitive when assimilating data in real time, as required for in-situ process monitoring [196, 197], particularly when working with large datasets or computationally expensive forward models. These challenges are further compounded when working with multimodal data, which often requires integrating observations with dynamically varying quality—for example, visual data may degrade at night, necessitating updates to the weighting of data importance [198]. Similarly, abrupt changes in the observed system, such as the sudden formation of a crack in a material, can render the current posterior nearly irrelevant, as it assigns negligible probability to events that now become highly probable [199].

Why Now? Emerging science drivers increasingly demand the assimilation of complex, rapidly evolving, dynamic data on actionable timescales. For example, advanced digital shadows and twins [6, 200] involve dynamically updating computational models of complex real-world systems by assimilating streaming data at a sufficient rate to support real-time monitoring, prediction, control, and decision-making. Additionally, the DOE's Grid Modernization Initiative calls for advanced computational techniques capable of assimilating dynamic grid data and interactions across transmission, distribution, energy technologies, and critical infrastructure to enhance grid resilience, improve situational awareness, and support reliable integration of multiple energy resources. Fortunately, recent algorithmic developments suggest several promising research paths to overcome current limitations in handling the complex dynamics, high-dimensional nonlinearities, multimodal information streams, rapidly evolving states, and strict time constraints inherent in these science drivers. For example, conditional generative algorithms, including diffusion models, flow matching, and stochastic interpolants, have shown potential to reduce the computational cost of solving static high-dimensional nonlinear probabilistic inference problems under complex settings [201–205], but their potential for dynamic assimilation has yet to be realized. Similarly, advances in transport maps [6, 191, 206, 207] offer a promising approach to reducing restrictive assumptions in particle-based assimilation by formulating data assimilation as transporting particles from the forecast density (generated by forward dynamics) to the filtered density (updated with current observations) [208–211]. Yet, the computational costs of learning and applying transport maps can grow recursively over time, posing significant scalability challenges. Efforts have been made to generalize the Gaussian ansatz of ensemble Kalman filters by directly learning filters from data with AI [212, 213] but have been limited to specific types of inverse problems, with scalable methods for high-dimensional, computationally expensive models still undeveloped.

Research Opportunities. Several promising research directions could help overcome current limitations in assimilating disparate data for complex systems. For example, research is needed to develop generative approaches for dynamic data assimilation that avoid the need for retraining at every time step, ensure posterior stability over time, and guarantee convergence as data volume and network complexity increase. Research is also required to design new hybrid approaches that integrate advances in generative modeling with novel strategies for particle- and sample-based data assimilation. These approaches could overcome the limiting assumptions of current methods regarding prior knowledge, model complexity, observation operators, and noise and error structures, enabling more complex and higher-dimensional real-world applications to be addressed. Scalable approaches to transport maps, including new strategies for selecting distance measures and novel learning techniques, could minimize the cost of recursive retraining, making these methods computationally feasible for high-dimensional systems and enabling them to model more complex dynamics and noise than current methods. Additionally, methods are needed to effectively demonstrate these benefits on multimodal datasets, posteriors with multiple modes, high-dimensional data, and complex state spaces, especially when forward operators are expensive to evaluate. Extensions of these advances are essential to adapt loss function weightings when multimodal data distributions undergo significant dynamic changes over time, ensuring that the assimilation process remains robust and accurate under evolving conditions.

PRD 4 Solving goal-oriented inverse problems for downstream tasks

Motivation and Impact. A broad spectrum of DOE programs increasingly demand solving inverse problems to improve predictions and quantify uncertainties in downstream tasks, such as model calibration, optimal control, design optimization, risk assessment, optimal experimental design (OED), and decision support. A prototypical example involves calibrating a digital twin—a numerical model of a specific physical system that evolves alongside the system over its lifetime—thus enhancing its ability to predict quantities of interest (QoIs) that differ from the observed quantities but are essential for decision-making, such as identifying optimal sensor placements to improve system performance. Traditionally, these goal-oriented inverse problems (GIPs), as defined in Section 1.2, are solved sequentially, first addressing inverse problems independently of downstream tasks by focusing solely on inferring model parameters from data, and then using the solution in downstream tasks. Existing methods are limited by restrictive assumptions, impractical computational demands, and an inability to fully extract information from available data, often resulting in biased predictions and overconfident uncertainty estimates. Furthermore, the growing reliance on digital twins, autonomous systems, and the rapid evolution of sensing technologies and data acquisition systems have made the need for novel, scalable solutions more urgent than ever. Without timely intervention, critical advancements in DOE mission areas—spanning energy systems, disaster response, and materials discovery—risk being delayed or remaining unattainable.

Meeting the needs of emerging GIPs requires a fundamental rethinking of how these problems are formulated and solved. Codesigning inverse problems with downstream tasks is essential to efficiently prioritize computational resources and data collection, even for competing objectives, and to deliver solutions within actionable timescales. Risk-aware approaches are needed to ensure downstream tasks are both informed by data and aligned with stakeholder preferences, helping to mitigate rare but high-consequence undesired events, such as system failures. Causal inference methods are crucial for disentangling correlation from causation, enabling the exploration of intervention outcomes, control actions, or experimental changes—key for evaluating "what-if" scenarios and ensuring robust decision-making. New approaches for solving GIPs that prioritize targeted data use to reduce uncertainty, optimize decision-making, and balance technical, economic, and safety objectives under stringent computational and resource constraints will significantly enhance efficiency and predictive accuracy across diverse fields, driving breakthroughs in energy systems, materials discovery, and autonomous systems.

Thrust A: How can we design goal-oriented inverse solvers that integrate downstream tasks?

Challenges and Gaps. While GIPs offer a transformative opportunity to align inverse problem solutions with downstream objectives, their formulation introduces significant challenges. Existing approaches often rely on restrictive linear and Gaussian assumptions [214, 215], which limit their ability to capture the complexity and nonlinear dependencies inherent in GIPs. Structures or data that are effective for improving solutions to traditional inverse problems may lose their utility when applied to GIPs, as the focus shifts from estimating model inputs to predicting QoIs [214, 216]. Even robust solutions to traditional inverse problems or model-error corrections frequently fail to deliver accurate QoI predictions, resulting in biased outcomes and overconfident uncertainty estimates [113]. In parallel, GIPs often result in highly nonlinear, nonconvex optimization problems, which remain a significant challenge—particularly when models fail to produce exact gradients, provide only approximate ones, or when objectives and constraints exhibit nonsmooth behavior [215, 217, 218]. These difficulties are further compounded when bilevel and mixed-integer optimization is required, such as in automated decision-making tasks like planning-operation or design-operation tasks. Automating the GIP solutions often requires risk-aware approaches to quantify and mitigate rare extreme events and uncertainties critical to decision-making. However, very little work has incorporated risk awareness [219], which amplifies computational costs and must adapt to evolving risk preferences. Disentangling correlation from causation is crucial to ensure predictions remain valid under interventions, such as control actions or experimental changes [220], yet existing methods often rely on overly restrictive assumptions and are typically tailored to simple downstream tasks, such as QoI prediction [214]. These challenges become even more pronounced when GIPs involve multiple downstream tasks, requiring the balancing of competing objectives that cannot be addressed by independently solving each task's GIP.

Why Now? The rapid advancement of sensing and data acquisition technologies, combined with the growing integration of scientific user facilities, data assets, and advanced computing—such as through the DOE's American Science Cloud initiative—could automate workflows centered around inverse problems. However, automating workflows, such as those used by large-scale materials simulations and experimentation, will require seamlessly and efficiently integrating inverse problem solutions with complex downstream tasks [200]. Recent advances in genAI [147, 148] hint at the possibility of bypassing traditional bottlenecks in estimating posterior distributions by directly conditioning uncertainty in downstream QoIs on observed data, but these methods must be fundamentally reimagined to address the complexities of emerging downstream tasks. Similarly, while risk-aware methods have been successfully applied outside of inverse problems, such as for uncertainty quantification in QoIs [221], they must be redesigned to address risk when solving GIPs. Multitask inference frameworks [222] and distributional inverse problems [223, 224] have shown that leveraging shared patterns across related systems can significantly enhance predictions, but similar breakthroughs are needed for GIPs. While these advances from other fields have been limited to simpler downstream tasks, such as prediction, they suggest that innovative approaches to solving GIPs could soon overcome existing challenges and unlock their full potential for improving decision-making and design optimization.

Research Opportunities. Meeting the critical needs of DOE mission challenges, which involve computationally expensive simulations and complex downstream tasks, requires a transformative rethinking of how inverse problems and downstream tasks are codesigned. Innovative approaches are needed to solve GIPs that regularize the inverse problem, uncover and exploit low-dimensional structures informed by data, and inform downstream objectives, thereby significantly reducing computational and data demands. Multitask methods capable of leveraging shared patterns across related systems are essential to enhance solution accuracy and precision for related, yet sometimes competing, downstream tasks. Additionally, new analogies of AI foundation models tailored to inverse problems could allow pretraining one model that powers many downstream GIP tasks. Furthermore, robust causal inference methods that can handle complex downstream tasks and ensure validity under varying conditions and distributional shifts will be essential for quantifying intervention outcomes, control actions, or experimental changes—critical for addressing "what-if scenarios." Risk-aware approaches are urgently needed to revolutionize decision-making and automation by accurately quantifying the data-informed impact of rare but high-consequence events while dynamically adapting to evolving risk preferences. Techniques for mitigating or quantifying model misspecification must be tailored to specific downstream tasks and risk-preferences to avoid overly conservative uncertainty estimates and ensure actionable solutions. Central to these efforts is the development of methods with strong theoretical error bounds and practical error estimation techniques. High-order optimization methods capable of efficiently solving problems that are multiobjective, risk-aware, and nonsmooth will also be vital to address these challenges and unlock the full potential of GIPs for DOE mission priorities.

Thrust B: How can we dynamically adapt data collection, fidelity, and computational resource allocation to improve the accuracy and efficiency of goal-oriented solutions?

Challenges and Gaps. Accurately solving GIPs within finite data and computational budgets requires balancing the costs and accuracies of both the inverse problem and downstream tasks. However, the complexity of optimally allocating resources poses significant challenges, especially when managing the feedback loop between the downstream task and the inverse problem formulation. These difficulties are further compounded by the absence of robust metrics for complex downstream tasks and the inherent competition between achieving accurate inversion and optimizing downstream objectives. This inability to jointly optimize these objectives results in significant inefficiencies, wasted resources, and suboptimal solutions, ultimately hindering the ability to solve GIPs within the actionable timeframes required for effective decision-making and autonomous systems. For example, data collection methods such as OED may minimize uncertainty in parameter posteriors but fail to reduce uncertainty in downstream tasks [225]. Similarly, surrogates for observational and QoI maps are typically constructed independently due to the lack of theoretical and practical methods for estimating their contribution to downstream tasks [226]. Errors arising from approximations, low-fidelity models, or model inadequacy are also rarely balanced, despite evidence that multifidelity methods can significantly reduce computational costs in fields adjacent to inverse problem-solving [227–229]. Moreover, real-world GIP applications must navigate dynamic and uncertain conditions, such as intermittent sensor availability, variable observation quality, and rapid system changes that render preplanned resource allocation ineffective and demand adaptive strategies that surpass the sequential, nonadaptive GIP approaches currently available [221, 230]. These challenges become even more pronounced when computing risk-aware solutions or when downstream tasks introduce conflicting priorities.

Why Now? While data measurements, workflows, and compute architectures are becoming more powerful, their rising costs highlight the need for methods that balance measurement and computational fidelity to maximize the speed and accuracy of solving GIPs. Recent advances in computational science, AI, and uncertainty quantification offer unprecedented opportunities to jointly optimize data collection and modeling fidelity for GIPs. Developments in multifidelity modeling, which integrates multiple simulation models, and active learning [231], which identifies optimal strategies for enriching data-driven models, have demonstrated significant potential to allocate resources more effectively for uncertainty quantification [232], surrogate modeling [233], and design [234]. However, these approaches must be reimagined to reduce the cost of solving GIPs. Similarly, recent advances in amortized genAI methods, such as diffusion models [147] and normalizing flows [148], offer the ability to reduce the cost of repeatedly solving inverse problems by learning direct mappings from observations to predictions. However, their accuracy and the data sources used to train them have not yet been tailored to balance the competing needs of GIPs nor to exploit experimental and observation data of varying quality and cost. Reinforcement learning [235] and Bayesian optimization [236] have advanced sequential decision-making under uncertainty by dynamically balancing exploration and exploitation for experimental resource allocation, such as in OED [236], and computational resource allocation, such as in optimal design [237]. However, even greater potential remains untapped due to the absence of approaches that are specifically tailored to GIPs and encode risk preferences.

Research Opportunities. Solving GIPs on actionable timescales requires a fundamental rethinking of how computational models, data, and downstream tasks are integrated. Novel methods are needed to significantly improve computational and data efficiency, enabling real-time decision-making and automation, particularly for GIPs involving bidirectional coupling between models and data, as exemplified by digital twins and autonomous discovery labs. As computational models grow increasingly complex and demanding, approaches must be developed to balance the costs of physics modeling, data collection, surrogate modeling, and generative modeling used to solve the inverse problem. These approaches must also account for the costs of similar processes required for multiple downstream tasks, such as designing and executing control policies. Additionally, they must ensure resource efficiency without compromising accuracy. These efforts could be advanced through new methods that exploit strategies such as multifidelity modeling, active learning, and reinforcement learning, which balance exploration and exploitation for GIPs. In this context, exploration refers to the process of acquiring new information to improve models, such as refining meshes, enriching embedded corrections, or collecting additional training data, while exploitation focuses on leveraging existing knowledge to solve inverse problems and address downstream tasks efficiently. For instance, these methods could guide the refinement of physics-based models for downstream tasks, such as refining meshes and embedding corrections for model misspecification, while simultaneously managing the cost of collecting training data and building generative models to solve the inverse problem, ensuring that the cost of improving these models is carefully balanced against the benefits they provide. Similarly, GIP strategies that determine the amount and type of data to collect based on downstream task requirements—such as enabling confident decision-making between two options under uncertainty—could significantly reduce the financial cost of data collection. Efficiencies could be further enhanced by developing risk-aware frameworks that prioritize computational and data resources to target the estimation of stakeholder-informed risk measures, ensuring solutions align with both technical and decision-making priorities.

3. Conclusion

The Workshop on Basic Research Needs for Inverse Problems for Complex Systems under Uncertainty, convened by the DOE's ASCR program in June 2025, aimed to address the growing complexity and rapidly evolving demands of inverse problems across the DOE's mission space. Inverse problems, which are central to understanding and optimizing complex systems, are increasingly characterized by multiphysics and multiscale models, multimodal data, and hybrid AI-aided simulations. Their cross-cutting nature and escalating computational challenges strongly motivated the formulation of a prioritized research agenda to maximize their capabilities and impact.

Bringing together experts across disciplines, the workshop sought to identify grand challenges and major opportunities in the field. Through collaborative discussions, participants defined transformative research directions to address the mathematical, statistical, and computational challenges posed by inverse problems under uncertainty. The workshop emphasized the importance of foundational mathematical and computational science research in optimization, probabilistic inference, hybrid modeling, and multifidelity techniques, while exploring emerging opportunities in AI and advanced computing architectures. These innovations will be essential for extracting insights from noisy, incomplete, and multimodal data, overcoming model limitations and misspecifications, integrating physics-based and data-driven models, and delivering actionable insights for high-consequence decisions.

The four PRDs outlined in this report provide a cohesive vision for advancing the science of inverse problems under uncertainty. Together, they address the critical challenges of discovering, exploiting, and preserving physical and problem structure; overcoming model limitations; integrating disparate multimodal and/or dynamic data; and tailoring solutions to downstream tasks. While each PRD focuses on a distinct aspect of inverse-problem research, their interconnected nature highlights the importance of a holistic approach that leverages progress across all areas to achieve transformative solutions.

The significance of this research extends far beyond theoretical advancements. Inverse problems underpin a wide range of DOE applications, including, but not limited to, reconstructing 3D structure and dynamics from X-ray, electron, and neutron data; optimizing fusion reactors; enhancing underground energy exploration and production; inferring new fundamental particle properties from particle colliders; designing new materials with targeted properties; understanding nuclear structure; improving wildfire predictions; accelerating quantum information sciences; and enabling autonomous systems and digital twins. Solving these problems with enhanced accuracy, scalability, and uncertainty-awareness while ensuring strong theoretical guarantees will accelerate advancements in scientific discovery, energy innovation, and national security.

Appendix A: Workshop Charge

Dear John, Michael, and Jeffrey,

Thank you for agreeing to be the chairs for the ASCR Applied Mathematics workshop focused on the Basic Research Needs for **Inverse Methods for Complex Systems under Uncertainty**. This email confirms ASCR's invitation for you to lead this important ASCR activity.

The workshop will follow a model based on that used by SC's Basic Energy Sciences program for their Basic Research Needs (BRN) workshops. As you know, critical to ASCR's success are the meeting of a broad group of participants from the community, in-person where possible, and the development of a report that outlines the priority research directions, as identified by the participants, capturing the opportunities to further advance applied mathematics research and impact in this space. This BRN workshop is scheduled for 2.5 days. On the afternoon of the second day, the breakout leads present the priority research directions identified during their breakout to the entire group. The morning of the third day is reserved for writing by the chairs, breakout leads, and other writers who may have been selected by the group.

The purpose of this workshop is to identify the ASCR priority research directions regarding inverse methods for complex systems under uncertainty. Inverse problems underlie many DOE-relevant applications that are becoming increasingly complex, involving multiple interconnected models, physics, and scales with observational data characterized by multiple modalities and fidelities. Both model and data uncertainties need to be accounted for in the estimations, especially for hybrid simulations that incorporate scientific machine learning and for digital twins.

The workshop participants will first consider the status, recent trends, and challenges facing DOE's science and technology missions to which advances might be relevant. The workshop participants will then examine the opportunities, barriers, and potential for high scientific impact through fundamental advances in the underlying mathematical, statistical, and computational research foundations. The resulting priority research directions should span several major algorithmic categories and state-of-the-art modeling and algorithm research from a variety of approaches; and cover different classes of techniques. Themes relevant for the workshop include inverse problems, parameter calibration, uncertainty quantification, numerical optimization under uncertainty, Bayesian approaches, multimodal data fusion, derivatives/sensitivities and adjoint methods, and probabilistic programming with applications to data analysis from experimental facilities, hybrid modeling and simulation, digital twins for control and decision support, and efficient AI/ML training.

The workshop and subsequent report should define the basic research needs and opportunities in applied and computational mathematics that can bring advancements in the creation and application of complex physical systems to many areas of relevance to the DOE mission.

The chairs are responsible for leading the entire workshop planning process. We will schedule regular conference calls among the chairs and DOE to start the planning process beginning next week. The overall tasks are listed below in approximate chronological order:

- Develop the high-level workshop structure, including deciding on the number and focus of the panels. Based on the meeting venue, we can have up to 3 panels.
- Identify possible plenary topics and speakers based on the panel topics.
- Identify panel leads, and then work with the panel leads to identify the workshop participants, including a plan to engage a broad range of DOE Lab personnel, academics and industry representatives. Ideally, this plan will provide for including people who have not participated in ASCR

workshops before. This is a time-consuming process that we should begin as soon as possible to get the meeting on people's calendars.

- As soon as possible, coordinate preparation of a background document on the status of the field that would be distributed to participants ahead of the workshop. ASCR program managers will participate in preparing this document.
- During the workshop, synthesize the panels' ideas, guide the identification and definition of priority research directions, and coordinate an oral report to the full workshop at the closing session.
- Coordinate and integrate the topical narratives provided by the panel leads and other identified writers into a final report. As much of the writing as possible is to be completed during the workshop, but follow-up writing is almost always required.

The goal is to have a final report within three months after the workshop to maximize the report's impact on programmatic planning.

We greatly appreciate your willingness to lead this essential planning activity for ASCR.

Steve Lee, Bill Spotz, and Todd Munson (cc'd) will provide you with additional guidance and coordinate with you going forward, including by communicating to you when the workshop can be discussed publicly.

Sincerely,
Hal
--
Hal Finkel
Director, Computational Science Research and Partnerships Division
Advanced Scientific Computing Research
Office of Science
U.S. Department of Energy
hal.finkel@science.doe.gov

Appendix B: Workshop Agenda

June 10, 2025

8:00 AM	Breakfast
9:00 AM	DOE ASCR and Co-Chairs Welcome
9:20 AM	Plenary Talks
	Drew Kouri, Sandia National Laboratories
	Youssef Marzouk, Massachusetts Institute of Technology
	Jeffrey Fessler, University of Michigan
10:35 AM	Break
11:00 AM	Parallel Breakout Discussion & Flash Talks
	Topic 1: Optimization Algorithms
	Topic 2: Probabilistic Approaches
	Topic 3: Limited/Noisy/Multimodal Data
12:00 PM	Lunch
1:00 PM	Parallel Breakout Discussion
	Topic 1: Optimization Algorithms
	Topic 2: Probabilistic Approaches
	Topic 3: Limited/Noisy/Multimodal Data
2:30 PM	Group Photo
2:40 PM	Break
3:00 PM	Parallel Breakout Writing Sessions / Report Out Preparation
	Topic 1: Optimization Algorithms
	Topic 2: Probabilistic Approaches
	Topic 3: Limited/Noisy/Multimodal Data
4:00 PM	Report Out and Discussion
5:00 PM	Adjourn

June 11, 2025

8:00 AM	Breakfast
9:00 AM	Welcome
9:05 AM	Plenary Talks
	Wei Cai, Southern Methodist University
	Omar Ghattas, University of Texas at Austin
	Noemi Petra, University of California, Merced
10:20 AM	Industry Panel
	Abhishek Chopra, BosonQ Psi Corp (BQP)
	Kevin Daly, ExxonMobil
	Genghis Khan, GE Aerospace
	Mark Kostuk, General Atomics
10:50 AM	Break
11:10 AM	Parallel Breakout Discussion & Flash Talks
	Topic 4: Uncertainty-Aware Modeling
	Topic 5: Goal-Oriented Problems

	Topic 6: Scalable Algorithms
12:10 PM	Lunch
1:10 PM	Parallel Breakout Discussion
	Topic 4: Uncertainty-Aware Modeling
	Topic 5: Goal-Oriented Problems
	Topic 6: Scalable Algorithms
2:40 PM	Break
3:00 PM	Parallel Breakout Writing Sessions / Report Out Preparation
	Topic 4: Uncertainty-Aware Modeling
	Topic 5: Goal-Oriented Problems
	Topic 6: Scalable Algorithms
4:00 PM	Report Out and Discussion
5:00 PM	Adjourn

June 12, 2025

8:00 AM	Breakfast
9:00 AM	Writing Session
10:30 AM	Break
11:00 AM	Writing Session
12:00 PM	Adjourn

Appendix C: Workshop Participants

Name	Institution
Nick Alger	The University of Texas at Austin
Rick Archibald	Oak Ridge National Laboratory
Tushar Athawale	Oak Ridge National Laboratory
Masoud Barati	University of Pittsburgh
Pouria Behnoudfar	University of Wisconsin-Madison
Julie Bessac	National Renewable Energy Laboratory
Debsindhu Bhowmik	Oak Ridge National Laboratory
Tommie Catanach	Sandia National Laboratories
Peng Chen	Georgia Institute of Technology
Youngsoo Choi	Lawrence Livermore National Laboratory
Abhishek Chopra	BosonQ Psi Corp
Emil Constantinescu	Argonne National Laboratory
Kevin Daly	ExxonMobil
Agnimitra Dasgupta	University of Southern California
Anabel del Val	University of Minnesota
Somayajulu Dhulipala	Idaho National Laboratory
Zichao Di	Argonne National Laboratory
Jeffrey Donatelli	Lawrence Berkeley National Laboratory
Kwassi Joseph Dzahini	Argonne National Laboratory
Hillary Fairbanks	Lawrence Livermore National Laboratory
Jeffrey Fessler	University of Michigan
Nando Fioretto	University of Virginia
Mark Fornace	Lawrence Berkeley National Laboratory
Anne Gelb	Dartmouth College
Omar Ghattas	The University of Texas at Austin
Seyede Fatemeh Ghoreishi	Northeastern University
Jinwoo Go	Brookhaven National Laboratory
Jonathan Gorard	Princeton Plasma Physics Laboratory
Salman Habib	Argonne National Laboratory
Thomas Hagstrom	Southern Methodist University
Malik Hassanaly	National Renewable Energy Laboratory
Felix Herrmann	Georgia Institute of Technology
Bamdad Hosseini	University of Washington
Zixi Hu	Lawrence Berkeley National Laboratory
Natalie Isenberg	Pacific Northwest National Laboratory
John Jakeman	Sandia National Laboratories
Sanket Jantre	Brookhaven National Laboratory
Ruiwei Jiang	University of Michigan
Ulugbek Kamilov	Washington University in St. Louis
Conlain Kelly	National Renewable Energy Laboratory
Genghis Khan	GE Aerospace
Mark Kostuk	General Atomics

Continued on next page

Name	Institution
Drew Kouri	Sandia National Laboratories
Dinesh Kumar	Lawrence Berkeley National Laboratory
Jeffrey Larson	Argonne National Laboratory
Qi Lei	New York University
Sven Leyffer	Argonne National Laboratory
Jiaming Liang	University of Rochester
Yen Ting Lin	Los Alamos National Laboratory
Youzuo Lin	University of North Carolina at Chapel Hill
Frank Liu	Old Dominion University
Fernando Llorente	Brookhaven National Laboratory
Andreas Mang	University of Houston
Youssef Marzouk	Massachusetts Institute of Technology
Roopesh Mathur	BosonQ Psi Corp
Michael McCann	Los Alamos National Laboratory
Matt Menickelly	Argonne National Laboratory
Susan Minkoff	Brookhaven National Laboratory
Sifat Afroj Moon	Oak Ridge National Laboratory
Juliane Mueller	National Renewable Energy Laboratory
Rie Nakata	Lawrence Berkeley National Laboratory
Assad Oberai	University of Southern California
Kanupriya Pande	Lawrence Berkeley National Laboratory
Lekha Patel	Sandia National Laboratories
Benjamin Peherstorfer	New York University
Cosmin Petra	Lawrence Livermore National Laboratory
Noemi Petra	University of California, Merced
Teresa Portone	Sandia National Laboratories
Hong Qin	Princeton Plasma Physics Laboratory
Juan Restrepo	Oak Ridge National Laboratory
Kristofer Reyes	Brookhaven National Laboratory
Pieterjan Robbe	Sandia National Laboratories
Johann Rudi	Virginia Tech
Steffen Schotthoefer	Oak Ridge National Laboratory
Siqian Shen	University of Michigan - Ann Arbor
Michael Shields	Johns Hopkins University
Guohui Song	Old Dominion University
Miroslav Stoyanov	Oak Ridge National Laboratory
Yu Sun	Johns Hopkins University
Kowshik Thopalli	Lawrence Livermore National Laboratory
Adam Thorpe	University of Texas at Austin
Bart van Bloemen Waanders	Sandia National Laboratories
Deepanshu Verma	Clemson University
Rebekah White	Sandia National Laboratories
Tim Wildey	Sandia National Laboratories
Brendt Wohlberg	Los Alamos National Laboratory
Pengcheng Xie	Lawrence Berkeley National Laboratory
Dongbin Xiu	The Ohio State University

Continued on next page

Name	Institution
Yanwen Xu	University of Texas at Dallas
Peng Xu	Ames National Laboratory
Yunan Yang	Cornell University
Zhi (Jackie) Yao	Lawrence Berkeley National Laboratory
Guannan Zhang	Oak Ridge National Laboratory

References

[1] J. J. Donatelli, J. A. Sethian, and P. H. Zwart. "Reconstruction from limited single-particle diffraction data via simultaneous determination of state, orientation, intensity, and phase". In: *Proceedings of the National Academy of Sciences* 114.28, 2017, pp. 7222–7227. DOI: 10.1073/pnas.1708217114.

[2] W. L. Miller, D. Bard, A. Boehnlein, K. Fagnan, C. Guok, E. Lançon, S. Ramprakash, M. Shankar, N. Schwarz, and B. L. Brown. *Integrated Research Infrastructure Architecture Blueprint Activity (Final Report 2023)*. Tech. rep. US Department of Energy (USDOE), Washington, DC (United States). Office of Science; Lawrence Berkeley National Laboratory (LBNL), Berkeley, CA (United States), 2023. DOI: 10.2172/1984466.

[3] US Department of Energy (USDOE), Washington, DC (United States). Office of Science. *High Performance Data Facility Project Status and Plan*. https://hpdf.science/sites/default/files/HPDFWebinars/20240617_HPDFStatusandPlan_final_0.pdf. Accessed: 2025-09-08. 2024.

[4] J. Degrave et al. "Magnetic control of tokamak plasmas through deep reinforcement learning". In: *Nature* 602.7897, 2022, pp. 414–419. DOI: 10.1038/s41586-021-04301-9.

[5] S. M. Benson and J. Deutch. "Advancing enhanced oil recovery as a sequestration asset". In: *Joule* 2.8, 2018, pp. 1386–1389. DOI: 10.1016/j.joule.2018.07.026.

[6] A. P. Gahlot, R. Orozco, Z. Yin, G. Bruer, and F. J. Herrmann. "An uncertainty-aware digital shadow for underground multimodal CO_2 storage monitoring". In: *Geophysical Journal International* 242.1, 2025. DOI: 10.1093/gji/ggaf176.

[7] A. Shmakov, K. Greif, M. Fenton, A. Ghosh, P. Baldi, and D. Whiteson. "End-To-End Latent Variational Diffusion Models for Inverse Problems in High Energy Physics". In: *Advances in Neural Information Processing Systems*. Ed. by A. Oh, T. Naumann, A. Globerson, K. Saenko, M. Hardt, and S. Levine. Vol. 36. Curran Associates, Inc., 2023, pp. 65102–65127. URL: https://proceedings.neurips.cc/paper_files/paper/2023/file/cd830afc6208a346e4ec5caf1b08b4b4-Paper-Conference.pdf.

[8] L. Grbčić, M. Park, M. Elzouka, R. Prasher, J. Müller, C. P. Grigoropoulos, S. D. Lubner, V. Zorba, and W. A. de Jong. "Inverse design of photonic surfaces via multi fidelity ensemble framework and femtosecond laser processing". In: *npj Computational Materials* 11.1, 2025, p. 35. DOI: 10.1038/s41524-025-01518-4.

[9] T. Gahlmann and P. Tassin. "Deep neural networks for the prediction of the optical properties and the free-form inverse design of metamaterials". In: *Phys. Rev. B* 106 8,2022, p. 085408. DOI: 10.1103/PhysRevB.106.085408.

[10] C. S. Ha et al. "Rapid inverse design of metamaterials based on prescribed mechanical behavior through machine learning". In: *Nature Communications* 14.1, 2023, p. 5765. DOI: 10.1038/s41467-023-40854-1.

[11] C. Xiao et al. "Ultrabroadband and band-selective thermal meta-emitters by machine learning". In: *Nature* 643.8070, 2025, pp. 80–88. DOI: 10.1038/s41586-025-09102-y.

[12] J. Zenihiro et al. "Neutron density distributions of ^{204,206,208}Pb deduced via proton elastic scattering at $E_p = 295$ MeV". In: *Phys. Rev. C* 82 4,2010, p. 044611. DOI: 10.1103/PhysRevC.82.044611.

[13] S. Bogner et al. “Computational nuclear quantum many-body problem: The UNEDF project”. In: *Computer Physics Communications* 184.10, 2013, pp. 2235–2250. DOI: 10.1016/j.cpc.2013.05.020.

[14] B. Shaddy, D. Ray, A. Farguell, V. Calaza, J. Mandel, J. Haley, K. Hilburn, D. V. Mallia, A. Kochanski, and A. Oberai. “Generative algorithms for fusion of physics-based wildfire spread models with satellite data for initializing wildfire forecasts”. In: *Artificial Intelligence for the Earth Systems* 3.3, 2024, e230087. DOI: 10.1175/AIES-D-23-0087.1.

[15] R. Clemens. *An Evaluation of the Varied Economic Impacts of the January 2025 California Wildfires*. Wesleyan Business Review. 2025. URL: https://www.wesleyanbusinessreview.com/issue-x-environmental-economics/blog-post-title-three-bnmr5.

[16] A. Farguell, J. Mandel, J. Haley, D. V. Mallia, A. Kochanski, and K. Hilburn. “Machine learning estimation of fire arrival time from level-2 active fires satellite data”. In: *Remote Sensing* 13.11, 2021, p. 2203. DOI: 10.3390/rs13112203.

[17] F. Belliardo, F. Zoratti, F. Marquardt, and V. Giovannetti. “Model-aware reinforcement learning for high-performance Bayesian experimental design in quantum metrology”. In: *Quantum* 8, 2024, p. 1555. DOI: 10.22331/q-2024-12-10-1555.

[18] C. L. Degen, F. Reinhard, and P. Cappellaro. “Quantum sensing”. In: *Reviews of Modern Physics* 89.3, 2017. DOI: 10.1103/revmodphys.89.035002.

[19] P. Lougovski et al. *Report for the ASCR Workshop on Basic Research Needs in Quantum Computing and Networking*. 2023. DOI: 10.2172/2001045.

[20] F. Alexander, M. Anitescu, J. Bell, D. Brown, M. Ferris, M. Luskin, S. Mehrotra, B. Moser, A. Pinar, A. Tartakovsky, et al. *A Multifaceted Mathematical Approach for Complex Systems*. Tech. rep. Lawrence Berkeley National Laboratory (LBNL), Berkeley, CA (United States), 2012. DOI: 10.2172/1093585.

[21] A. Buluc, T. G. Kolda, S. M. Wild, M. Anitescu, A. Degennaro, J. D. Jakeman, C. Kamath, R. R. Kannan, M. E. Lopes, P.-G. Martinsson, et al. “Randomized Algorithms for Scientific Computing (RASC)”. In: Oak Ridge National Laboratory (ORNL), Oak Ridge, TN (United States). 2021. DOI: 10.2172/1807223.

[22] S. Klasky, J. Thayer, and H. Najm. *Data Reduction for Science: Brochure from the Advanced Scientific Computing Research Workshop*. Tech. rep. Oak Ridge National Laboratory (ORNL), Oak Ridge, TN (United States); SLAC National Accelerator Laboratory (SLAC), Menlo Park, CA (United States); Sandia National Laboratories (SNL-CA), Livermore, CA (United States); Sandia National Laboratories (SNL-NM), Albuquerque, NM (United States), 2021. DOI: 10.2172/1770192.

[23] N. Baker, F. Alexander, T. Bremer, A. Hagberg, Y. Kevrekidis, H. Najm, M. Parashar, A. Patra, J. Sethian, S. Wild, et al. *Workshop Report on Basic Research Needs for Scientific Machine Learning: Core Technologies for Artificial Intelligence*. Tech. rep. USDOE Office of Science (SC), Washington, D.C. (United States), 2019. DOI: 10.2172/1478744.

[24] R. Stevens, V. Taylor, J. Nichols, A. B. Maccabe, K. Yelick, and D. Brown. *AI for Science: Report on the Department of Energy (DOE) Town Halls on Artificial Intelligence (AI) for Science*. Tech. rep. Argonne National Laboratory (ANL), Argonne, IL (United States), 2020. DOI: 10.2172/1604756.

[25] J. Dongarra, J. Hittinger, J. Bell, L. Chacon, R. Falgout, M. Heroux, P. Hovland, E. Ng, C. Webster, and S. Wild. *Applied Mathematics Research for Exascale Computing*. Tech. rep. Lawrence Livermore National Laboratory (LLNL), Livermore, CA, 2014. DOI: 10.2172/1149042.

[26] A. Almgren, P. DeMar, J. Vetter, K. Riley, K. Antypas, D. Bard, R. Coffey, E. Dart, S. Dosanjh, R. Gerber, et al. *Advanced Scientific Computing Research Exascale Requirements Review. An Office of Science review sponsored by Advanced Scientific Computing Research, September 27-29, 2016, Rockville, Maryland*. Tech. rep. Argonne National Lab. (ANL), Argonne, IL (United States). Argonne Leadership Computing Facility, 2017. DOI: 10.2172/1375638.

[27] R. Ross, L. Ward, P. Carns, G. Grider, S. Klasky, Q. Koziol, G. K. Lockwood, K. Mohror, B. Settlemyer, and M. Wolf. *Storage Systems and Input/Output: Organizing, Storing, and Accessing Data for Scientific Discovery. Report for the DOE ASCR Workshop on Storage Systems and I/O. [Full Workshop Report]*. Tech. rep. USDOE Office of Science (SC) (United States), 2018. DOI: 10.2172/1491994.

[28] T. Peterka, D. Bard, J. Bennett, E. W. Bethel, R. Oldfield, L. Pouchard, C. Sweeney, and M. Wolf. *ASCR Workshop on In Situ Data Management: Enabling Scientific Discovery from Diverse Data Sources*. Tech. rep. US Department of Energy (USDOE), Washington, DC (United States). Office of Science, 2019. DOI: 10.2172/1493245.

[29] S. Byna, S. Idreos, T. Jones, K. Mohror, R. Ross, and F. Rusu. *Report for the ASCR Workshop on the Management and Storage of Scientific Data*. Tech. rep. US Department of Energy (USDOE), Washington, DC (United States). Office of Science, 2021. DOI: 10.2172/1845707.

[30] J. Ang, A. A. Chien, S. D. Hammond, A. Hoisie, I. Karlin, S. Pakin, J. Shalf, and J. S. Vetter. *Reimagining Codesign for Advanced Scientific Computing: Report for the ASCR Workshop on Reimagining Codesign*. Tech. rep. USDOE Office of Science (SC) (United States), 2022. DOI: 10.2172/1822199.

[31] N. Alger, T. Hartland, N. Petra, and O. Ghattas. "Point spread function approximation of high-rank Hessians with locally supported nonnegative integral kernels". In: *SIAM Journal on Scientific Computing* 46.3, 2024, A1658–A1689. DOI: 10.1137/23M1584745.

[32] H. Kersting, N. Krämer, M. Schiegg, C. Daniel, M. Tiemann, and P. Hennig. "Differentiable likelihoods for fast inversion of 'likelihood-free' dynamical systems". In: *International Conference on Machine Learning*. 2020, pp. 5198–5208. URL: https://proceedings.mlr.press/v119/kersting20a.html.

[33] K. Carlberg, Y. Choi, and S. Sargsyan. "Conservative model reduction for finite-volume models". In: *Journal of Computational Physics* 371, 2018, pp. 280–314. DOI: 10.1016/j.jcp.2018.05.019.

[34] A. M. DeGennaro, N. M. Urban, B. T. Nadiga, and T. Haut. "Model structural inference using local dynamic operators". In: *International Journal for Uncertainty Quantification* 9.1, 2019, pp. 59–83. DOI: 10.1615/Int.J.UncertaintyQuantification.2019025828.

[35] Z. Zhang and J. Zhao. "General numerical framework to derive structure preserving reduced order models for thermodynamically consistent reversible-irreversible PDEs". In: *Journal of Computational Physics* 521, 2025, p. 113562. DOI: 10.1016/j.jcp.2024.113562.

[36] S. Jantre, D. Akhare, Z. Wang, X. Qian, and N. M. Urban. *Data-augmented few-shot neural emulator for computer-model system identification*. 2025. DOI: 10.48550/arXiv.2508.19441.

[37] H. Dou, Z. Li, J. Du, L. Yang, W. Yao, and Y. Deng. “Hybrid regularization improves diffusion-based inverse problem solving”. In: *The Thirteenth International Conference on Learning Representations*. 2025. URL: https://openreview.net/forum?id=d7pr2doXn3.

[38] A. Aali, M. Arvinte, S. Kumar, and J. I. Tamir. “Solving inverse problems with score-based generative priors learned from noisy data”. In: *2023 57th Asilomar Conference on Signals, Systems, and Computers*. IEEE. 2023, pp. 837–843. DOI: 10.1109/IEEECONF59524.2023.10477042.

[39] T. Schneider, A. M. Stuart, and J.-L. Wu. “Ensemble Kalman inversion for sparse learning of dynamical systems from time-averaged data”. In: *Journal of Computational Physics* 470, 2022, p. 111559. DOI: 10.1016/j.jcp.2022.111559.

[40] M. Garcia-Gasulla and M. J. Mantsinen. “Challenges and opportunities in exascale fusion simulations”. In: *Nature Reviews Physics*, 2025, pp. 1–10. DOI: 10.1038/s42254-025-00830-8.

[41] O. Shahid, S. Pouriyeh, R. M. Parizi, Q. Z. Sheng, G. Srivastava, and L. Zhao. *Communication efficiency in federated learning: Achievements and challenges*. 2021. DOI: 10.48550/arXiv.2107.10996.

[42] K. J. Rahman, F. Ahmed, N. Akhter, M. Hasan, R. Amin, K. E. Aziz, A. M. Islam, M. S. H. Mukta, and A. N. Islam. “Challenges, applications and design aspects of federated learning: A survey”. In: *IEEe Access* 9, 2021, pp. 124682–124700. DOI: 10.1109/ACCESS.2021.3111118.

[43] G. Carvalho, B. Cabral, V. Pereira, and J. Bernardino. “Edge computing: Current trends, research challenges and future directions”. In: *Computing* 103.5, 2021, pp. 993–1023. DOI: 10.1007/s00607-020-00896-5.

[44] B. Peherstorfer, K. Willcox, and M. D. Gunzburger. “Survey of multifidelity methods in uncertainty propagation, inference, and optimization”. In: *SIAM Review* 60.3, 2018, pp. 550–591. DOI: 10.1137/16M1082469.

[45] Z. Azarhoosh and M. I. Ghazaan. “A review of recent advances in surrogate models for uncertainty quantification of high-dimensional engineering applications”. In: *Computer Methods in Applied Mechanics and Engineering* 433, 2025, p. 117508. DOI: 10.1016/j.cma.2024.117508.

[46] Z. Sun, J. Fan, H. Li, and H. Jiang. “Current status of single particle imaging with X-ray lasers”. In: *Applied Sciences* 8.1, 2018, p. 132. DOI: 10.3390/app8010132.

[47] W. Du, F. Iacoviello, M. Mirza, S. Zhou, J. Bu, S. Feng, P. S. Grant, R. Jervis, D. J. Brett, and P. R. Shearing. “X-ray computed laminography: A brief review of mechanisms, reconstruction, applications and perspectives”. In: *Materials Today*, 2025. DOI: 10.1016/j.mattod.2025.02.023.

[48] B. Toader, F. J. Sigworth, and R. R. Lederman. “Methods for cryo-EM single particle reconstruction of macromolecules having continuous heterogeneity”. In: *Journal of Molecular Biology* 435.9, 2023, p. 168020. DOI: 10.1016/j.jmb.2023.168020.

[49] M. A. Marcus. “Information content of and the ability to reconstruct dichroic X-ray tomography and laminography”. In: *Optics Express* 30.22, 2022, pp. 39445–39465. DOI: 10.1364/OE.462410.

[50] Z. Hu and J. J. Donatelli. "Multitiered computational methodology for extracting three-dimensional rotational diffusion coefficients from X-ray photon correlation spectroscopy data without structural information". In: *Physical Review B* 110.21, 2024, p. 214305. DOI: 10.1103/PhysRevB.110.214305.

[51] H. Goel, O. Chubar, R. Li, L. Wiegart, M. Rakitin, and A. Fluerasu. "Efficient end-to-end simulation of time-dependent coherent X-ray scattering experiments". In: *Journal of Synchrotron Radiation* 31.3, 2024. DOI: 10.1107/S1600577524001267.

[52] R. A. Khalek, A. Accardi, J. Adam, D. Adamiak, W. Akers, M. Albaladejo, A. Al-Bataineh, M. Alexeev, F. Ameli, P. Antonioli, et al. "Science requirements and detector concepts for the electron-ion collider: EIC yellow report". In: *Nuclear Physics A* 1026, 2022, p. 122447. DOI: 10.1016/j.nuclphysa.2022.122447.

[53] S. Gori, N. Tran, K. DiPetrillo, B. Echenard, J. Eldred, R. Harnik, P. Machado, M. Toups, I. P. Fernando, R. Bernstein, et al. "ACE Science Workshop Report". In: Fermi National Accelerator Laboratory (FNAL), Batavia, IL (United States). 2024. DOI: 10.48550/arXiv.2403.02422.

[54] X. Zhang and J. Almer. *A new capability facilitating nuclear materials research: the Activated Materials Laboratory at the Advanced Photon Source*. Tech. rep. Argonne National Laboratory (ANL), 2025. DOI: 10.2172/2572622.

[55] T. Boltz, J. L. Martinez, C. Xu, K. R. Baker, Z. Zhu, J. Morgan, R. Roussel, D. Ratner, B. Mustapha, and A. L. Edelen. "Leveraging prior mean models for faster Bayesian optimization of particle accelerators". In: *Scientific Reports* 15.1, 2025, p. 12232. DOI: 10.1038/s41598-025-95297-z.

[56] M. Chu, Z. Jiang, M. Wojcik, T. Sun, M. Sprung, and J. Wang. "Probing three-dimensional mesoscopic interfacial structures in a single view using multibeam X-ray coherent surface scattering and holography imaging". In: *Nature Communications* 14.1, 2023, p. 5795. DOI: 10.1038/s41467-023-39984-3.

[57] S. Kevan. "ALS-U: Solving scientific challenges with coherent soft X-Rays". In: *Workshop Report on Early Science Enabled by the Advanced Light Source Upgrade, ALS, LBNL, Berkeley, CA*. 2017. URL: https://als.lbl.gov/wp-content/uploads/2017/08/ALS-U-Early-Science-Workshop-Report-Abbreviated.pdf.

[58] S. M. Kaye, D. Battaglia, D. Baver, E. Belova, J. Berkery, V. N. Duarte, N. Ferraro, E. Fredrickson, N. Gorelenkov, W. Guttenfelder, et al. "NSTX/NSTX-U theory, modeling and analysis results". In: *Nuclear Fusion* 59.11, 2019, p. 112007. DOI: 10.1088/1741-4326/ab023a.

[59] J. Arrington, M. Battaglieri, A. Boehnlein, S. A. Bogacz, W. Brooks, E. Chudakov, I. Cloet, R. Ent, H. Gao, J. Grames, et al. "Physics with CEBAF at 12 GeV and future opportunities". In: *Progress in Particle and Nuclear Physics* 127, 2022, p. 103985. DOI: 10.1016/j.ppnp.2022.103985.

[60] E. Ye and N. F. Loureiro. "Quantum-inspired method for solving the Vlasov-Poisson equations". In: *Physical Review E* 106.3, 2022, p. 035208. DOI: 10.1103/PhysRevE.106.035208.

[61] R. D. Peddinti, S. Pisoni, A. Marini, P. Lott, H. Argentieri, E. Tiunov, and L. Aolita. "Quantum-inspired framework for computational fluid dynamics". In: *Communications Physics* 7.1, 2024, p. 135. DOI: 10.1038/s42005-024-01623-8.

[62] Y. Núñez Fernández, M. K. Ritter, M. Jeannin, J.-W. Li, T. Kloss, T. Louvet, S. Terasaki, O. Parcollet, J. von Delft, H. Shinaoka, et al. "Learning tensor networks with tensor cross interpolation: New algorithms and libraries". In: *SciPost Physics* 18.3, 2025, p. 104. DOI: 10.21468/SciPostPhys.18.3.104.

[63] L. Hölscher, P. Rao, L. Müller, J. Klepsch, A. Luckow, T. Stollenwerk, and F. K. Wilhelm. "Quantum-inspired fluid simulation of two-dimensional turbulence with GPU acceleration". In: *Physical Review Research* 7.1, 2025, p. 013112. DOI: 10.1103/PhysRevResearch.7.013112.

[64] H. Shinaoka, M. Wallerberger, Y. Murakami, K. Nogaki, R. Sakurai, P. Werner, and A. Kauch. "Multiscale space-time ansatz for correlation functions of quantum systems based on quantics tensor trains". In: *Physical Review X* 13.2, 2023, p. 021015. DOI: 10.1103/PhysRevX.13.021015.

[65] N. Gourianov, M. Lubasch, S. Dolgov, Q. Y. van den Berg, H. Babaee, P. Givi, M. Kiffner, and D. Jaksch. "A quantum-inspired approach to exploit turbulence structures". In: *Nature Computational Science* 2.1, 2022, pp. 30–37. DOI: 10.1038/s43588-021-00181-1.

[66] R. Abadía-Heredia, M. López-Martín, B. Carro, J. I. Arribas, J. M. Pérez, and S. Le Clainche. "A predictive hybrid reduced order model based on proper orthogonal decomposition combined with deep learning architectures". In: *Expert Systems with Applications* 187, 2022, p. 115910. DOI: 10.1016/j.eswa.2021.115910.

[67] S. Brivio, S. Fresca, N. R. Franco, and A. Manzoni. "Error estimates for POD-DL-ROMs: a deep learning framework for reduced order modeling of nonlinear parametrized PDEs enhanced by proper orthogonal decomposition". In: *Advances in Computational Mathematics* 50.3, 2024, p. 33. DOI: 10.1007/s10444-024-10110-1.

[68] Z. Li, S. Patil, F. Ogoke, D. Shu, W. Zhen, M. Schneier, J. R. Buchanan Jr, and A. B. Farimani. "Latent neural PDE solver: A reduced-order modeling framework for partial differential equations". In: *Journal of Computational Physics* 524, 2025, p. 113705. DOI: 10.1016/j.jcp.2024.113705.

[69] A. Solera-Rico, C. Sanmiguel Vila, M. Gómez-López, Y. Wang, A. Almashjary, S. T. Dawson, and R. Vinuesa. "β-variational autoencoders and transformers for reduced-order modelling of fluid flows". In: *Nature Communications* 15.1, 2024, p. 1361. DOI: 10.1038/s41467-024-45578-4.

[70] Z. Li, D. Shu, and A. Barati Farimani. "Scalable transformer for PDE surrogate modeling". In: *Advances in Neural Information Processing Systems* 36, 2023, pp. 28010–28039. DOI: 10.48550/arXiv.2305.17560.

[71] L. Serrano, T. X. Wang, E. Le Naour, J.-N. Vittaut, and P. Gallinari. "AROMA: Preserving spatial structure for latent PDE modeling with local neural fields". In: *Advances in Neural Information Processing Systems* 37, 2024, pp. 13489–13521. DOI: 10.48550/arXiv.2406.02176.

[72] F. Pichi, B. Moya, and J. S. Hesthaven. "A graph convolutional autoencoder approach to model order reduction for parametrized PDEs". In: *Journal of Computational Physics* 501, 2024, p. 112762. DOI: 10.1016/j.jcp.2024.112762.

[73] P. Buchfink, S. Glas, and B. Haasdonk. "Symplectic model reduction of Hamiltonian systems on nonlinear manifolds and approximation with weakly symplectic autoencoder". In: *SIAM Journal on Scientific Computing* 45.2, 2023, A289–A311. DOI: 10.1137/21M1466657.

[74] L.-H. Yang, X.-L. Luo, Z.-B. Yang, C.-F. Nan, X.-F. Chen, and Y. Sun. “FE reduced-order model-informed neural operator for structural dynamic response prediction”. In: *Neural Networks* 188, 2025, p. 107437. DOI: 10.1016/j.neunet.2025.107437.

[75] A. Gruber and I. Tezaur. “Canonical and noncanonical Hamiltonian operator inference”. In: *Computer Methods in Applied Mechanics and Engineering* 416, 2023, p. 116334. DOI: 10.1016/j.cma.2023.116334.

[76] P.-G. Martinsson and J. A. Tropp. “Randomized numerical linear algebra: Foundations and algorithms”. In: *Acta Numerica* 29, 2020, pp. 403–572. DOI: 10.1017/S0962492920000021.

[77] T. Le, A. Myers, T. Bui-Thanh, and Q. Nguyen. *A data-scalable randomized misfit approach for solving large-scale PDE-constrained inverse problems*. 2016. DOI: 10.48550/arXiv.1603.01562.

[78] E. Herman, A. Alexanderian, and A. K. Saibaba. “Randomization and Reweighted ℓ_1-Minimization for A-Optimal Design of Linear Inverse Problems”. In: *SIAM Journal on Scientific Computing* 42.3, 2020, A1714–A1740. DOI: 10.1137/19M1267362.

[79] J. Wittmer, C. Krishnanunni, H. V. Nguyen, and T. Bui-Thanh. “On unifying randomized methods for inverse problems”. In: *Inverse Problems* 39.1, 2023, pp. 211–229. DOI: 10.1088/1361-6420/acd36e.

[80] K. Chen, Q. Li, K. Newton, and S. J. Wright. “Structured random sketching for PDE inverse problems”. In: *SIAM Journal on Matrix Analysis and Applications* 41.4, 2020, pp. 1742–1770. DOI: 10.1137/20M1310497.

[81] N. Boumal. *An Introduction to Optimization on Smooth Manifolds*. Cambridge University Press, 2023. DOI: 10.1017/9781009166164.

[82] J. Wang and C. G. Petra. “A sequential quadratic programming algorithm for nonsmooth problems with upper-$\mathcal{C}^2$ objective”. In: *SIAM Journal on Optimization* 33.3, 2023, pp. 2379–2405. DOI: 10.1137/22M1490995.

[83] J. Larson, M. Menickelly, and B. Zhou. “Manifold sampling for optimizing nonsmooth nonconvex compositions”. In: *SIAM Journal on Optimization* 31.4, 2021, pp. 2638–2664. DOI: 10.1137/20M1378089.

[84] J. Larson, M. Menickelly, and S. M. Wild. “Manifold sampling for ℓ_1 nonconvex optimization”. In: *SIAM Journal on Optimization* 26.4, 2016, pp. 2540–2563. DOI: 10.1137/15M1042097.

[85] K. A. Khan, J. Larson, and S. M. Wild. “Manifold sampling for optimization of nonconvex functions that are piecewise linear compositions of smooth components”. In: *SIAM Journal on Optimization* 28.4, 2018, pp. 3001–3024. DOI: 10.1137/17m114741x.

[86] J. Liu, Y. Cui, J.-S. Pang, and S. Sen. “Two-stage stochastic programming with linearly bi-parameterized quadratic recourse”. In: *SIAM Journal on Optimization* 30.3, 2020, pp. 2530–2558. DOI: 10.1137/19M1276819.

[87] F. E. Curtis and M. L. Overton. “A sequential quadratic programming algorithm for nonconvex, nonsmooth constrained optimization”. In: *SIAM Journal on Optimization* 22.2, 2012, pp. 474–500. DOI: 10.1137/090780201.

[88] S. Arridge, P. Maass, O. Öktem, and C.-B. Schönlieb. “Solving inverse problems using data-driven models”. In: *Acta Numerica* 28, 2019, pp. 1–174. DOI: 10.1017/S0962492919000059.

[89] T. Zhou, P. Krahenbuhl, and A. A. Efros. “Learning data-driven reflectance priors for intrinsic image decomposition”. In: *Proceedings of the IEEE International Conference on Computer Vision*. 2015. DOI: 10.1109/ICCV.2015.396.

[90] G. A. Kaissis, M. R. Makowski, D. Rückert, and R. F. Braren. "Secure, privacy-preserving and federated machine learning in medical imaging". In: *Nature Machine Intelligence* 2.6, 2020, pp. 305–311. DOI: 10.1038/s42256-020-0186-1.

[91] F. Knoll, K. Hammernik, C. Zhang, S. Moeller, T. Pock, D. K. Sodickson, and M. Akcakaya. "Deep-learning methods for parallel magnetic resonance imaging reconstruction: A survey of the current approaches, trends, and issues". In: *IEEE Signal Processing Magazine* 37.1, 2020, pp. 128–140. DOI: 10.1109/msp.2019.2950640.

[92] J. Yuan, X. Luo, Y. Qin, Y. Zhao, W. Ju, and M. Zhang. "Learning on graphs under label noise". In: *IEEE International Conference on Acoustics, Speech and Signal Processing*. 2023, pp. 1–5. DOI: 10.1109/ICASSP49357.2023.10096088.

[93] J. Xie, L. Sun, and Y. F. Zhao. "On the data quality and imbalance in machine learning-based design and manufacturing—A systematic review". In: *Engineering* 45, 2025, pp. 105–131. DOI: 10.1016/j.eng.2024.04.024.

[94] A. Tran, X. He, D. A. Messenger, Y. Choi, and D. M. Bortz. "Weak-form latent space dynamics identification". In: *Computer Methods in Applied Mechanics and Engineering* 427, 2024, p. 116998. DOI: 10.1016/j.cma.2024.116998.

[95] M. Chung, D. Verma, M. Collins, A. N. Subrahmanya, V. K. Sastry, and V. Rao. *Latent twins*. 2025. DOI: 10.48550/arXiv.2509.20615.

[96] X. He, Y. Choi, W. D. Fries, J. L. Belof, and J.-S. Chen. "gLaSDI: Parametric physics-informed greedy latent space dynamics identification". In: *Journal of Computational Physics* 489, 2023, p. 112267. DOI: 10.1016/j.jcp.2023.112267.

[97] A. Bora, A. Jalal, E. Price, and A. G. Dimakis. "Compressed sensing using generative models". In: *Proceedings of the 34th International Conference on Machine Learning*. Ed. by D. Precup and Y. W. Teh. Vol. 70. Proceedings of Machine Learning Research. PMLR, 2017, pp. 537–546. URL: https://proceedings.mlr.press/v70/bora17a.html.

[98] N. Kovachki, Z. Li, B. Liu, K. Azizzadenesheli, K. Bhattacharya, A. Stuart, and A. Anandkumar. "Neural operator: Learning maps between function spaces with applications to PDEs". In: *Journal of Machine Learning Research* 24.89, 2023, pp. 1–97. URL: https://jmlr.org/papers/v24/21-1524.html.

[99] B. Ghojogh, M. Crowley, F. Karray, and A. Ghodsi. *Elements of Dimensionality Reduction and Manifold Learning*. Springer, 2023. DOI: 10.1007/978-3-031-10602-6.

[100] O. Ghattas and K. Willcox. "Learning physics-based models from data: perspectives from inverse problems and model reduction". In: *Acta Numerica* 30, 2021, pp. 445–554. DOI: 10.1017/S0962492921000064.

[101] K. Kontolati, S. Goswami, G. Em Karniadakis, and M. D. Shields. "Learning nonlinear operators in latent spaces for real-time predictions of complex dynamics in physical systems". In: *Nature Communications* 15.1, 2024, p. 5101. DOI: 10.1038/s41467-024-49411-w.

[102] N. B. Kovachki, S. Lanthaler, and A. M. Stuart. "Operator learning: Algorithms and analysis". In: *Handbook of Numerical Analysis* 25, 2024, pp. 419–467. DOI: 10.1016/bs.hna.2024.05.009.

[103] A. Cozad, N. V. Sahinidis, and D. C. Miller. "Learning surrogate models for simulation-based optimization". In: *AIChE Journal* 60.6, 2014, pp. 2211–2227. DOI: 10.1002/aic.14418.

[104] S. H. Kim and F. Boukouvala. "Machine learning-based surrogate modeling for data-driven optimization: A comparison of subset selection for regression techniques". In: *Optimization Letters* 14.4, 2020, pp. 989–1010. DOI: 10.1007/s11590-019-01428-7.

[105] T. O'Leary-Roseberry, U. Villa, P. Chen, and O. Ghattas. "Derivative-informed projected neural networks for high-dimensional parametric maps governed by PDEs". In: *Computer Methods in Applied Mechanics and Engineering* 388, 2022, p. 114199. DOI: 10.1016/j.cma.2021.114199.

[106] D. Luo, T. O'Leary-Roseberry, P. Chen, and O. Ghattas. "Efficient PDE-constrained optimization under high-dimensional uncertainty using derivative-informed neural operators". In: *SIAM Journal on Scientific Computing* 47.4, 2025, pp. C899–C931. DOI: 10.1137/23M157956X.

[107] G. Daras, H. Chung, C.-H. Lai, Y. Mitsufuji, J. C. Ye, P. Milanfar, A. G. Dimakis, and M. Delbracio. *A survey on diffusion models for inverse problems*. 2024. DOI: 10.48550/arXiv.2410.00083.

[108] M. Elad, B. Kawar, and G. Vaksman. "Image denoising: The deep learning revolution and beyond—a survey paper". In: *SIAM Journal on Imaging Sciences* 16.3, 2023, pp. 1594–1654. DOI: 10.1137/23M1545859.

[109] U. Kamilov, C. A. Bouman, G. T. Buzzard, and B. Wohlberg. "Plug-and-Play Methods for Integrating Physical and Learned Models in Computational Imaging". In: *IEEE Signal Processing Magazine* 40.1, 2023, pp. 85–97. DOI: 10.1109/MSP.2022.3199595.

[110] J. Hu, B. Song, X. Xu, L. Shen, and J. A. Fessler. "Learning image priors through patch-based diffusion models for solving inverse problems". In: *Advances in Neural Information Processing Systems* 37, 2024, pp. 1625–1660. URL: https://proceedings.neurips.cc/paper_files/paper/2024/file/03261886741f1f21f52f2a2d570616a2-Paper-Conference.pdf.

[111] K. Cranmer, J. Brehmer, and G. Louppe. "The frontier of simulation-based inference". In: *Proceedings of the National Academy of Sciences* 117.48, 2020, pp. 30055–30062. DOI: 10.1073/pnas.1912789117.

[112] M. C. Kennedy and A. O'Hagan. "Bayesian calibration of computer models". In: *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 63.3, 2001, pp. 425–464. DOI: 10.1111/1467-9868.00294.

[113] T. A. Oliver, G. Terejanu, C. S. Simmons, and R. D. Moser. "Validating predictions of unobserved quantities". In: *Computer Methods in Applied Mechanics and Engineering* 283, 2015, pp. 1310–1335. DOI: 10.1016/j.cma.2014.08.023.

[114] K. Sargsyan, X. Huan, and H. N. Najm. "Embedded model error representation for Bayesian model calibration". In: *International Journal for Uncertainty Quantification* 9.4, 2019, pp. 365–394. DOI: 10.1615/Int.J.UncertaintyQuantification.2019027384.

[115] R. E. Morrison, T. A. Oliver, and R. D. Moser. "Representing Model Inadequacy: A Stochastic Operator Approach". In: *SIAM/ASA Journal on Uncertainty Quantification* 6.2, 2018. Publisher: Society for Industrial and Applied Mathematics, pp. 457–496. DOI: 10.1137/16M1106419.

[116] J. Brynjarsdottir and A. O'Hagan. "Learning about physical parameters: the importance of model discrepancy". In: *Inverse Problems* 30.11, 2014, p. 114007. DOI: 10.1088/0266-5611/30/11/114007.

[117] Y. Ling, J. Mullins, and S. Mahadevan. "Selection of model discrepancy priors in Bayesian calibration". In: *Journal of Computational Physics* 276, 2014, pp. 665–680. DOI: 10.1016/j.jcp.2014.08.005.

[118] J.-L. Wu, M. E. Levine, T. Schneider, and A. Stuart. "Learning about structural errors in models of complex dynamical systems". In: *Journal of Computational Physics* 513, 2024, p. 113157. DOI: 10.1016/j.jcp.2024.113157.

[119] K. Pande, J. J. Donatelli, E. Malmerberg, L. Foucar, C. Bostedt, I. Schlichting, and P. H. Zwart. "Ab initio structure determination from experimental fluctuation X-ray scattering data". In: *Proceedings of the National Academy of Sciences* 115.46, 2018, pp. 11772–11777. DOI: 10.1073/pnas.1812064115.

[120] F. Dallari, M. Reiser, I. Lokteva, A. Jain, J. Möller, M. Scholz, A. Madsen, G. Grübel, F. Perakis, and F. Lehmkühler. "Analysis strategies for MHz XPCS at the European XFEL". In: *Applied Sciences* 11.17, 2021, p. 8037. DOI: 10.3390/app11178037.

[121] Y. Kim, J. Bielecki, M. Sikorski, R. de Wijn, C. Fortmann-Grote, J. Sztuk-Dambietz, J. Koliyadu, R. Letrun, H. Kirkwood, T. Sato, et al. "Expected resolution limits of X-ray free-electron laser single-particle imaging for realistic source and detector properties". In: *Structural Dynamics* 9.6, 2022. DOI: 10.1063/4.0000169.

[122] T. You, J. Bielecki, and F. R. Maia. "Impact of gas background on XFEL single-particle imaging". In: *Scientific Reports* 15.1, 2025, p. 29559. DOI: 10.1038/s41598-025-15092-8.

[123] W. Vågberg, J. C. Larsson, and H. M. Hertz. "Removal of ring artifacts in microtomography by characterization of scintillator variations". In: *Optics Express* 25.19, 2017, pp. 23191–23198. DOI: 10.1364/OE.25.023191.

[124] Z. Chen, X. Xi, Y. Han, S. Tan, C. Wang, Y. Sun, L. Li, and B. Yan. "Research status and prospect of ring artifact correction method for X-ray CT". In: *International Conference on Optoelectronic Information and Functional Materials (OIFM 2024)*. Vol. 13183. SPIE. 2024, pp. 245–252. DOI: 10.1117/12.3033856.

[125] T. N. Palmer. "Stochastic weather and climate models". In: *Nature Reviews Physics* 1.7, 2019, pp. 463–471. DOI: 10.1038/s42254-019-0062-2.

[126] S. Lunz, A. Hauptmann, T. Tarvainen, C.-B. Schönlieb, and S. Arridge. "On learned operator correction in inverse problems". In: *SIAM Journal on Imaging Sciences* 14.1, 2021, pp. 92–127. DOI: 10.1137/20m1338460.

[127] S. L. Brunton, J. L. Proctor, and J. N. Kutz. "Discovering governing equations from data by sparse identification of nonlinear dynamical systems". In: *Proceedings of the National Academy of Sciences* 113.15, 2016, p. 3932. DOI: 10.1073/pnas.1517384113.

[128] S.-M. Udrescu and M. Tegmark. *AI Feynman: A physics-inspired method for symbolic regression*. 2020. DOI: 10.48550/arXiv.1905.11481.

[129] Z. Li, N. Kovachki, K. Azizzadenesheli, B. Liu, K. Bhattacharya, A. Stuart, and A. Anandkumar. *Fourier neural operator for parametric partial differential equations*. 2021. DOI: 10.48550/arXiv.2010.08895.

[130] L. Lu, P. Jin, G. Pang, Z. Zhang, and G. E. Karniadakis. "Learning nonlinear operators via DeepONet based on the universal approximation theorem of operators". In: *Nature Machine Intelligence* 3.3, 2021, pp. 218–229. DOI: 10.1038/s42256-021-00302-5.

[131] M. Xie, J. Liu, Y. Sun, W. Gan, B. Wohlberg, and U. S. Kamilov. "Joint reconstruction and calibration using regularization by denoising with application to computed tomography". In: *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV) Workshops*. 2021, pp. 4028–4037. DOI: 10.1109/ICCVW54120.2021.00448.

[132] S. Gupta, K. Kothari, V. Debarnot, and I. Dokmanić. "Differentiable uncalibrated imaging". In: *IEEE Transactions on Computational Imaging* 10, 2024, pp. 1–16. DOI: 10.1109/tci.2023.3346294.

[133] J. J. Donatelli and J. C. Spence. "Inversion of many-beam Bragg intensities for phasing by iterated projections: Removal of multiple scattering artifacts from diffraction data". In: *Physical Review Letters* 125.6, 2020, p. 065502. DOI: 10.1103/PhysRevLett.125.065502.

[134] K. C. Zhou and R. Horstmeyer. "Diffraction tomography with a deep image prior". In: *Optics Express* 28.9, 2020, pp. 12872–12896. DOI: 10.1364/OE.379200.

[135] J. Lee, H. Seo, W. Lee, and H. Park. "Unsupervised motion artifact correction of turbo spin-echo MRI using deep image prior". In: *Magnetic Resonance in Medicine* 92.1, 2024, pp. 28–42. DOI: 10.1002/mrm.30026.

[136] M. U. Ghani and W. C. Karl. "Deep learning based sinogram correction for metal artifact reduction". In: *Electronic Imaging* 30, 2018, pp. 4721–4728. DOI: 10.2352/ISSN.2470-1173.2018.15.COIMG-472.

[137] H. Yanxon, J. Weng, H. Parraga, W. Xu, U. Ruett, and N. Schwarz. "Artifact identification in X-ray diffraction data using machine learning methods". In: *Synchrotron Radiation* 30.1, 2023, pp. 137–146. DOI: 10.1107/S1600577522011274.

[138] H.-Q. Mu, X.-X. Liang, J.-H. Shen, and F.-L. Zhang. "Analysis of structural health monitoring data with correlated measurement error by Bayesian system identification: Theory and application". In: *Sensors* 22.20, 2022. DOI: 10.3390/s22207981.

[139] J. Ghorbanian, N. Casaprima, and A. Olivier. "Empowering approximate Bayesian neural networks with functional priors through anchored ensembling for mechanics surrogate modeling applications". In: *Computer Methods in Applied Mechanics and Engineering* 438, 2025, p. 117645. DOI: 10.1016/j.cma.2024.117645.

[140] T. E. Fricker, J. E. Oakley, and N. M. Urban. "Multivariate Gaussian process emulators with nonseparable covariance structures". In: *Technometrics* 55.1, 2013, pp. 47–56. DOI: 10.1080/00401706.2012.715835.

[141] Z. Yin, R. Orozco, M. Louboutin, and F. J. Herrmann. "Solving multiphysics-based inverse problems with learned surrogates and constraints". In: *Advanced Modeling and Simulation in Engineering Sciences* 10.1, 2023, p. 14. DOI: 10.1186/s40323-023-00252-0.

[142] M. Moscoso, A. Novikov, G. Papanicolaou, and C. Tsogka. "Imaging with highly incomplete and corrupted data". In: *Inverse Problems* 36.3, 2020, p. 035010. DOI: 10.1088/1361-6420/ab5a21.

[143] Z. Zhou, L.-Z. Guo, Z. Cheng, Y.-F. Li, and S. Pu. "STEP: Out-of-distribution detection in the presence of limited in-distribution labeled data". In: *Advances in Neural Information Processing Systems*. Ed. by M. Ranzato, A. Beygelzimer, Y. Dauphin, P. Liang, and J. W. Vaughan. Vol. 34. Curran Associates, Inc., 2021, pp. 29168–29180. URL: https://proceedings.neurips.cc/paper_files/paper/2021/file/f4334c131c781e2a6f0a5e34814c8147-Paper.pdf.

[144] D. Huang, A. Bharti, A. Souza, L. Acerbi, and S. Kaski. "Learning robust statistics for simulation-based inference under model misspecification". In: *Advances in Neural Information Processing Systems* 36, 2023, pp. 7289–7310. URL: https://openreview.net/forum?id=STrXsSIEiq.

[145] L. Cao, T. O'Leary-Roseberry, P. K. Jha, J. T. Oden, and O. Ghattas. "Residual-based error correction for neural operator accelerated infinite-dimensional Bayesian inverse problems". In: *Journal of Computational Physics* 486, 2023, p. 112104. DOI: https://doi.org/10.1016/j.jcp.2023.112104.

[146] B. Calderhead, M. Girolami, and N. Lawrence. "Accelerating Bayesian inference over nonlinear differential equations with Gaussian processes". In: *Advances in Neural Information Processing Systems* 21, 2008. URL: https://proceedings.neurips.cc/paper/2008/file/07563a3fe3bbe7e3ba84431ad9d055af-Paper.pdf.

[147] J. Ho, A. Jain, and P. Abbeel. "Denoising Diffusion Probabilistic Models". In: *Advances in Neural Information Processing Systems*. Ed. by H. Larochelle, M. Ranzato, R. Hadsell, M. Balcan, and H. Lin. Vol. 33. Curran Associates, Inc., 2020, pp. 6840–6851. URL: https://dl.acm.org/doi/abs/10.5555/3495724.3496298.

[148] G. Papamakarios, E. Nalisnick, D. J. Rezende, S. Mohamed, and B. Lakshminarayanan. "Normalizing flows for probabilistic modeling and inference". In: *Journal of Machine Learning Research* 22.57, 2021, pp. 1–64. URL: http://jmlr.org/papers/v22/19-1028.html.

[149] R. P. Kelly, D. J. Warne, D. T. Frazier, D. J. Nott, M. U. Gutmann, and C. Drovandi. *Simulation-based Bayesian inference under model misspecification*. 2025. DOI: 10.48550/arXiv.2503.12315.

[150] M. Schmitt, P.-C. Bürkner, U. Köthe, and S. T. Radev. "Detecting model misspecification in amortized Bayesian inference with neural networks". In: *DAGM German Conference on Pattern Recognition*. Springer. 2023, pp. 541–557. DOI: 10.1007/978-3-031-54605-1_35.

[151] S. Talts, M. Betancourt, D. Simpson, A. Vehtari, and A. Gelman. *Validating Bayesian inference algorithms with simulation-based calibration*. 2018. DOI: 10.48550/arXiv.1804.06788.

[152] T. D. Savitsky and J. Gershunskaya. *Simulation-based calibration of uncertainty intervals under approximate Bayesian estimation*. 2024. DOI: 10.48550/arXiv.2407.04659.

[153] M. Modrák, A. H. Moon, S. Kim, P. Bürkner, N. Huurre, K. Faltejsková, A. Gelman, and A. Vehtari. "Simulation-based calibration checking for Bayesian computation: The choice of test quantities shapes sensitivity". In: *Bayesian Analysis* 20.2, 2025, pp. 461–488. DOI: 10.1214/23-BA1404.

[154] A. Wehenkel, J. L. Gamella, O. Sener, J. Behrmann, G. Sapiro, J.-H. Jacobsen, and M. Cuturi. *Addressing misspecification in simulation-based inference through data-driven calibration*. 2024. DOI: 10.48550/arXiv.2405.08719.

[155] L. Elsemüller, V. Pratz, M. von Krause, A. Voss, P.-C. Bürkner, and S. T. Radev. *Does unsupervised domain adaptation improve the robustness of amortized Bayesian inference? a systematic evaluation*. 2025. DOI: 10.48550/arXiv.2502.04949.

[156] M. Schmitt, D. R. Ivanova, D. Habermann, U. Köthe, P.-C. Bürkner, and S. T. Radev. *Leveraging self-consistency for data-efficient amortized Bayesian inference*. 2023. DOI: 10.48550/arXiv.2310.04395.

[157] P. G. Bissiri, C. C. Holmes, and S. G. Walker. "A general framework for updating belief distributions". In: *Journal of the Royal Statistical Society. Series B (Statistical Methodology)* 78.5, 2016. DOI: 10.1111/rssb.12158.

[158] J. Knoblauch, J. Jewson, and T. Damoulas. *Generalized variational inference: Three arguments for deriving new posteriors*. 2019. DOI: 10.48550/arXiv.1904.02063.

[159] Z. Di, S. Leyffer, and S. M. Wild. “Optimization-based approach for joint X-ray fluorescence and transmission tomographic inversion”. In: *SIAM Journal on Imaging Sciences* 9.1, 2016, pp. 1–23. DOI: 10.1137/15M1021404.

[160] Z. Chen et al. “Machine learning on neutron and X-ray scattering and spectroscopies”. In: *Chemical Physics Reviews* 2.3, 2021, p. 031301. DOI: 10.1063/5.0049111.

[161] S. V. Kalinin and M. Ziatdinov. “Machine learning in scanning transmission electron microscopy”. In: *Nature Reviews Methods Primers* 2, 2022, pp. 1–19. DOI: 10.1038/s43586-022-00095-w.

[162] A. Ter-Petrosyan et al. *Revealing the evolution of order in materials microstructures using multimodal computer vision*. 2024. DOI: 10.48550/arXiv.2411.09896.

[163] A. Murari, R. Rossi, T. Craciunescu, J. Vega, M. Gelfusa, and J. Contributors. “A control oriented strategy of disruption prediction to avoid the configuration collapse of tokamak reactors”. In: *Nature Communications* 15.1, 2024, p. 2424. DOI: 10.1038/s41467-024-46242-7.

[164] B. Clavier, D. Zarzoso, D. Del-Castillo-Negrete, and É. Frénod. “A generative artificial intelligence framework for long-time plasma turbulence simulations”. In: *Physics of Plasmas* 32.6, 2025, p. 063905. DOI: 10.1063/5.0255386.

[165] A. K. Ambika, K. Tayal, V. Mishra, and D. Lu. “Novel deep learning transformer model for short to sub-seasonal streamflow forecast”. In: *Geophysical Research Letters* 52.14, 2025, e2025GL116707. DOI: 10.1029/2025GL116707.

[166] Y. Yu, J. Mao, S. D. Wullschleger, and et al. “Machine learning–based observation-constrained projections reveal elevated global socioeconomic risks from wildfire”. In: *Nature Communications* 13.1, 2022, p. 1250. DOI: 10.1038/s41467-022-28853-0.

[167] R. U. Shaik, M. Alipour, E. Rowell, B. Balaji, A. Watts, and E. Taciroglu. “FUELVISION: A multimodal data fusion and multimodel ensemble algorithm for wildfire fuels mapping”. In: *International Journal of Applied Earth Observation and Geoinformation* 138, 2025, p. 104436. DOI: 10.1016/j.jag.2025.104436.

[168] K. J. Hernandez et al. “L-PBF high-throughput data pipeline approach for multi-modal integration”. In: *Integrating Materials and Manufacturing Innovation* 13.3, 2024, pp. 758–772. DOI: 10.1007/s40192-024-00368-0.

[169] S. Cheng, J. Chen, C. Anastasiou, P. Angeli, O. K. Matar, Y.-K. Guo, C. C. Pain, and R. Arcucci. “Generalised latent assimilation in heterogeneous reduced spaces with machine learning surrogate models”. In: *Journal of Scientific Computing* 94.1, 2022, p. 11. DOI: 10.1007/s10915-022-02059-4.

[170] J. Bischof, G. Fletcher, P. Verkade, C. Kuntner, J. Fernandez-Rodriguez, L. Chaabane, L. A. Rose, A. Walter, M. Vandenbosch, M. A. van Zandvoort, et al. “Multimodal bioimaging across disciplines and scales: Challenges, opportunities and breaking down barriers”. In: *npj Imaging* 2.1, 2024, p. 5. DOI: 10.1038/s44303-024-00010-w.

[171] K. Pande, J. Donatelli, D. Parkinson, H. Yan, and J. Sethian. “Joint iterative reconstruction and 3D rigid alignment for X-ray tomography”. In: *Optics Express* 30.6, 2022, pp. 8898–8916. DOI: 10.1364/OE.443248.

[172] Z. Huang, G. Niu, X. Liu, W. Ding, X. Xiao, H. Wu, and X. Peng. “Learning with noisy correspondence for cross-modal matching”. In: *Advances in Neural Information Processing Systems*. Ed. by M. Ranzato, A. Beygelzimer, Y. Dauphin, P. Liang, and J. W. Vaughan. Vol. 34. Curran Associates, Inc., 2021, pp. 29406–29419. URL: https://proceedings.neurips.cc/paper_files/paper/2021/file/f5e62af885293cf4d511ceef31e61c80-Paper.pdf.

[173] H. Anzid, G. Le Goïc, A. Bekkari, A. Mansouri, and D. Mammass. “A new SURF-based algorithm for robust registration of multimodal images data”. In: *The Visual Computer* 39.4, 2023, pp. 1667–1681. DOI: 10.1007/s00371-022-02435-z.

[174] H. O. Velesaca, G. Bastidas, M. Rouhani, and A. D. Sappa. “Multimodal image registration techniques: A comprehensive survey”. In: *Multimedia Tools and Applications* 83.23, 2024, pp. 63919–63947. DOI: 10.1007/s11042-023-17991-2.

[175] X. Peng, Y. Wei, A. Deng, D. Wang, and D. Hu. “Balanced multimodal learning via on-the-fly gradient modulation”. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2022, pp. 8238–8247. DOI: 10.1109/CVPR52688.2022.00806.

[176] D. Lahat, T. Adali, and C. Jutten. “Multimodal data fusion: An overview of methods, challenges, and prospects”. In: *Proceedings of the IEEE* 103.9, 2015, pp. 1449–1477. DOI: 10.1109/JPROC.2015.2460697.

[177] P. P. Liang, A. Zadeh, and L.-P. Morency. “Foundations & trends in multimodal machine learning: Principles, challenges, and open questions”. In: *ACM Comput. Surv.* 56.10, 2024. DOI: 10.1145/3656580.

[178] Q. Zhang et al. “Multimodal fusion on low-quality data: A comprehensive survey”. In: *CoRR* abs/2404.18947, 2024. DOI: 10.48550/arXiv.2404.18947.

[179] D. P. Kingma and M. Welling. *Auto-encoding variational Bayes*. 2013. DOI: 10.48550/arXiv.1312.6114.

[180] D. P. Kingma, M. Welling, et al. “An introduction to variational autoencoders”. In: *Foundations and Trends® in Machine Learning* 12.4, 2019, pp. 307–392. DOI: 10.1561/2200000056.

[181] T. Chen, S. Kornblith, M. Norouzi, and G. Hinton. “A simple framework for contrastive learning of visual representations”. In: *International Conference on Machine Learning*. PmLR. 2020, pp. 1597–1607. URL: https://proceedings.mlr.press/v119/chen20j/chen20j.pdf.

[182] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, et al. *An image is worth 16x16 words: Transformers for image recognition at scale*. 2020. DOI: 10.48550/arXiv.2010.11929.

[183] X. Wang et al. “ORBIT: Oak Ridge base foundation model for earth system predictability”. In: *SC24: International Conference for High Performance Computing, Networking, Storage and Analysis*. 2024, pp. 1–11. DOI: 10.1109/SC41406.2024.00007.

[184] S. Mai, H. Hu, and S. Xing. “Modality to modality translation: An adversarial representation learning and graph fusion network for multimodal fusion”. In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 34. 01. 2020, pp. 164–172. DOI: 10.1609/aaai.v34i01.5347.

[185] W. Guo, J. Wang, and S. Wang. “Deep multimodal representation learning: A survey”. In: *IEEE Access* 7, 2019, pp. 63373–63394. DOI: 10.1109/ACCESS.2019.2916887.

[186] C. Zhang, Z. Yang, X. He, and L. Deng. “Multimodal intelligence: Representation learning, information fusion, and applications”. In: *IEEE Journal of Selected Topics in Signal Processing* 14.3, 2020, pp. 478–493. DOI: 10.1109/JSTSP.2020.2987728.

[187] T. Baltrušaitis, C. Ahuja, and L.-P. Morency. “Multimodal machine learning: A survey and taxonomy”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 41.2, 2018, pp. 423–443. DOI: 10.1109/tpami.2018.2798607.

[188] J. Hu, S. Sun, X. Yang, S. Zhou, X. Wang, Y. Fu, J. Zhou, Y. Yin, K. Cao, Q. Song, et al. “Towards accurate and robust multi-modal medical image registration using contrastive metric learning”. In: *IEEE Access* 7, 2019, pp. 132816–132827. DOI: 10.1109/ACCESS.2019.2938858.

[189] T. Van Leeuwen, S. Maretzke, and K. J. Batenburg. “Automatic alignment for three-dimensional tomographic reconstruction”. In: *Inverse Problems* 34.2, 2018, p. 024004. DOI: 10.1088/1361-6420/aaa0f8.

[190] Z. Di, S. Chen, D. Gursoy, T. Paunesku, S. Leyffer, S. M. Wild, and S. Vogt. “Optimization-based simultaneous alignment and reconstruction in multi-element tomography”. In: *Optics Letters* 44.17, 2019, pp. 4331–4334. DOI: 10.1364/ol.44.004331.

[191] G. Evensen. *Data Assimilation: The Ensemble Kalman Filter*. Springer, 2009. DOI: 10.1007/978-3-642-03711-5.

[192] P. J. van Leeuwen, H. R. Künsch, L. Nerger, R. Potthast, and S. Reich. “Particle filters for high-dimensional geoscience applications: A review”. In: *Quarterly Journal of the Royal Meteorological Society* 145.723, 2019, pp. 2335–2365. DOI: 10.1002/qj.3551.

[193] N. T. Mücke, S. M. Bohté, and C. W. Oosterlee. “The deep latent space particle filter for real-time data assimilation with uncertainty quantification”. In: *Scientific Reports* 14.1, 2024, p. 19447. DOI: 10.1038/s41598-024-69901-7.

[194] C. Snyder, T. Bengtsson, and M. Morzfeld. “Performance bounds for particle filters using the optimal proposal”. In: *Monthly Weather Review* 143.11, 2015, pp. 4750–4761. DOI: 10.1175/MWR-D-15-0144.1.

[195] S. Maskell and S. Julier. “Optimised proposals for improved propagation of multi-modal distributions in particle filters”. In: *Proceedings of the 16th International Conference on Information Fusion*. 2013, pp. 296–303. URL: https://ieeexplore.ieee.org/document/6641292.

[196] G. L. Knapp, T. Mukherjee, J. S. Zuback, H. Wei, T. A. Palmer, A. De, and T. DebRoy. “Building blocks for a digital twin of additive manufacturing”. In: *Acta Materialia* 135, 2017, pp. 390–399. DOI: 10.1016/j.actamat.2017.06.039.

[197] T. Shen and B. Li. “Digital twins in additive manufacturing: a state-of-the-art review”. In: *The International Journal of Advanced Manufacturing Technology* 131.1, 2024, pp. 63–92. DOI: 10.1007/s00170-024-13092-y.

[198] C. Y. Lin, T. Matsuo, J. Y. Liu, C. H. Lin, H. F. Tsai, and E. A. Araujo-Pradere. “Ionospheric assimilation of radio occultation and ground-based GPS data using non-stationary background model error covariance”. In: *Atmospheric Measurement Techniques* 8.1, 2015, pp. 171–182. DOI: 10.5194/amt-8-171-2015.

[199] C. K. R. Lim and D. Mba. “Switching Kalman filter for failure prognostic”. In: *Mechanical Systems and Signal Processing* 52-53, 2015, pp. 426–435. DOI: 10.1016/j.ymssp.2014.08.006.

[200] K. Willcox, D. Bingham, C. Chung, J. Chung, C. Cruz-Neira, C. Grant, J. Kinter, R. Leung, P. Moin, L. Ohno-Machado, et al. *Foundational Research Gaps and Future Directions for Digital Twins*. Tech. rep. National Academies Press, 2024. DOI: 10.17226/26894.

[201] F. Bao, Z. Zhang, and G. Zhang. “A score-based filter for nonlinear data assimilation”. In: *Journal of Computational Physics* 514, 2024, p. 113207. DOI: 10.1016/j.jcp.2024.113207.

[202] J.-H. Bastek and D. M. Kochmann. “Inverse design of nonlinear mechanical metamaterials via video denoising diffusion models”. In: *Nature Machine Intelligence* 5.12, 2023, pp. 1466–1475. DOI: 10.1038/s42256-023-00762-x.

[203] S. Zeng, H. Li, A. P. Gahlot, and F. J. Herrmann. *Well2Flow: Reconstruction of reservoir states from sparse wells using score-based generative models*. 2025. DOI: 10.48550/arXiv.2504.06305.

[204] A. Dasgupta, H. Ramaswamy, J. Murgoitio-Esandi, K. Y. Foo, R. Li, Q. Zhou, B. F. Kennedy, and A. A. Oberai. “Conditional score-based diffusion models for solving inverse elasticity problems”. In: *Computer Methods in Applied Mechanics and Engineering* 433, 2025, p. 117425. DOI: 10.1016/j.cma.2024.117425.

[205] F. Bao, H. Chipilski, S. Liang, G. Zhang, and J. Whitaker. “Nonlinear ensemble filtering with diffusion models: Application to the surface quasi-geostrophic dynamics”. In: *Monthly Weather Review* 153.7, 2025, pp. 1155–1169. DOI: 10.1175/MWR-D-24-0245.1.

[206] A. Spantini, R. Baptista, and Y. Marzouk. “Coupling techniques for nonlinear ensemble filtering”. In: *SIAM Review* 64.4, 2022, pp. 921–953. DOI: 10.1137/20M1312204.

[207] A. Taghvaei and B. Hosseini. “An optimal transport formulation of Bayes’ law for nonlinear filtering algorithms”. In: *2022 IEEE 61st Conference on Decision and Control*. IEEE. 2022, pp. 6608–6613. DOI: 10.1109/CDC51059.2022.9992776.

[208] H. G. Chipilski. “Exact nonlinear state estimation”. In: *Journal of the Atmospheric Sciences* 82.4, 2025, pp. 809–827. DOI: 10.1175/JAS-D-24-0171.s1.

[209] M. Al-Jarrah, N. Jin, B. Hosseini, and A. Taghvaei. *Nonlinear filtering with Brenier optimal transport maps*. 2023. DOI: 10.48550/arXiv.2310.13886.

[210] Y. Shi, V. De Bortoli, G. Deligiannidis, and A. Doucet. “Conditional simulation using diffusion Schrödinger bridges”. In: *Uncertainty in Artificial Intelligence*. PMLR. 2022, pp. 1792–1802. URL: https://proceedings.mlr.press/v180/shi22a.html.

[211] F. Bao, G. Zhang, and Z. Zhang. *A unified filter method for jointly estimating state and parameters of stochastic dynamical systems via the ensemble score filter*. 2023. DOI: 10.48550/arXiv.2312.10503.

[212] E. Bach, R. Baptista, E. Calvello, B. Chen, and A. Stuart. *Learning Enhanced Ensemble Filters*. 2025. DOI: 10.48550/arXiv.2504.17836.

[213] M. McCabe and J. Brown. “Learning to Assimilate in Chaotic Dynamical Systems”. In: *Advances in Neural Information Processing Systems*. Ed. by M. Ranzato, A. Beygelzimer, Y. Dauphin, P. Liang, and J. W. Vaughan. Vol. 34. Curran Associates, Inc., 2021, pp. 12237–12250. URL: https://proceedings.neurips.cc/paper_files/paper/2021/file/65cc2c8205a05d7379fa3a6386f710e1-Paper.pdf.

[214] A. Spantini, T. Cui, K. Willcox, L. Tenorio, and Y. Marzouk. “Goal-oriented optimal approximations of Bayesian linear inverse problems”. In: *SIAM Journal on Scientific Computing* 39.5, 2017, S167–S196. DOI: 10.1137/16M1082123.

[215] A. Attia, A. Alexanderian, and A. K. Saibaba. “Goal-oriented optimal design of experiments for large-scale Bayesian linear inverse problems”. In: *Inverse Problems* 34.9, 2018, p. 095009. DOI: 10.1088/1361-6420/aad210.

[216] S. Zhong, W. Shen, T. Catanach, and X. Huan. *Goal-oriented Bayesian optimal experimental design for nonlinear models using Markov Chain Monte Carlo*. 2025. DOI: 10.48550/arXiv.2403.18072.

[217] T. C. Y. Chan, R. Mahmood, and I. Y. Zhu. “Inverse optimization: Theory and applications”. In: *Operations Research* 73.2, 2025, pp. 1046–1074. DOI: 10.1287/opre.2022.0382.

[218] N. Ye, F. Roosta-Khorasani, and T. Cui. *Optimization Methods for Inverse Problems*. 2017. DOI: 10.48550/arXiv.1712.00154.

[219] D. P. Kouri, J. D. Jakeman, and J. Gabriel Huerta. “Risk-adapted optimal experimental design”. In: *SIAM/ASA Journal on Uncertainty Quantification* 10.2, 2022, pp. 687–716. DOI: 10.1137/20M1357615.

[220] J. Peters, D. Janzing, and B. Schölkopf. *Elements of Causal Inference: Foundations and Learning Algorithms*. Cambridge, MA: MIT Press, 2017. URL: https://mitpress.mit.edu/9780262037310/elements-of-causal-inference/.

[221] S. Henneking, S. Venkat, and O. Ghattas. *Goal-oriented real-time bayesian inference for linear autonomous dynamical systems with application to digital twins for tsunami early warning*. 2025. DOI: 10.48550/arXiv.2501.14911.

[222] X. Tang, X. Li, R. Yu, Y. Wu, J. Ye, F. Tang, and Q. Chen. “Digital-twin-assisted task assignment in multi-UAV systems: A deep reinforcement learning approach”. In: *IEEE Internet of Things Journal* 10.17, 2023, pp. 15362–15375. DOI: 10.1109/JIOT.2023.3263574.

[223] A. Vadeboncoeur, M. Girolami, and A. M. Stuart. *Efficient deconvolution in populational inverse problems*. 2025. DOI: 10.48550/arXiv.2505.19841.

[224] R. D. White, J. D. Jakeman, T. Wildey, and T. Butler. “Building population-informed priors for Bayesian inference using data-consistent stochastic inversion”. In: *SIAM/ASA Journal on Uncertainty Quantification* 13.3, 2025, pp. 1475–1500. DOI: 10.1137/24M1678234.

[225] T. Butler, J. Jakeman, and T. Wildey. “Optimal experimental design for prediction based on push-forward probability measures”. In: *Journal of Computational Physics* 416, 2020, p. 109518. DOI: 10.1016/j.jcp.2020.109518.

[226] S. A. Mattis and B. Wohlmuth. “Goal-oriented adaptive surrogate construction for stochastic inversion”. In: *Computer Methods in Applied Mechanics and Engineering* 339, 2018, pp. 36–60. DOI: 10.1016/j.cma.2018.04.045.

[227] M. B. Giles. “Multilevel Monte Carlo methods”. In: *Acta Numerica* 24, 2015, pp. 259–328. DOI: 10.1017/S096249291500001X.

[228] A. Gorodetsky, G. Geraci, M. Eldred, and J. Jakeman. “A generalized approximate control variate framework for multifidelity uncertainty quantification”. In: *Journal of Computational Physics* 408, 2020, p. 109257. DOI: 10.1016/j.jcp.2020.109257.

[229] M. C. Kennedy and A. O’Hagan. “Predicting the output from a complex computer code when fast approximations are available”. In: *Biometrika* 87.1, 2000, pp. 1–13. URL: http://www.jstor.org/stable/2673557.

[230] J. N. Neuberger, A. Alexanderian, and B. van Bloemen Waanders. *Goal oriented optimal design of infinite-dimensional Bayesian inverse problems using quadratic approximations*. 2024. DOI: 10.48550/arXiv.2411.07532.

[231] P. Ren, Y. Xiao, X. Chang, P.-Y. Huang, Z. Li, B. B. Gupta, X. Chen, and X. Wang. “A survey of deep active learning”. In: *ACM Comput. Surv.* 54.9, 2021. DOI: 10.1145/3472291.

[232] D. Schaden and E. Ullmann. “On multilevel best linear unbiased estimators”. In: *SIAM/ASA Journal on Uncertainty Quantification* 8.2, 2020, pp. 601–635. DOI: 10.1137/19M1263534.

[233] K. Cutajar, M. Pullin, A. Damianou, N. Lawrence, and J. González. *Deep Gaussian processes for multi-fidelity modeling*. 2019. DOI: 10.48550/arXiv.1903.07320.

[234] L. Grbcic, J. Müller, and W. A. de Jong. "Efficient inverse design optimization through multi-fidelity simulations, machine learning, and boundary refinement strategies". In: *Engineering with Computers* 40.6, 2024, pp. 4081–4108. DOI: 10.1007/s00366-024-02053-4.

[235] W. Shen and X. Huan. "Bayesian sequential optimal experimental design for nonlinear models using policy gradient reinforcement learning". In: *Computer Methods in Applied Mechanics and Engineering* 416, 2023, p. 116304. DOI: 10.1016/j.cma.2023.116304.

[236] B. Shahriari, K. Swersky, Z. Wang, R. P. Adams, and N. de Freitas. "Taking the human out of the loop: A review of Bayesian optimization". In: *Proceedings of the IEEE* 104.1, 2016, pp. 148–175. DOI: 10.1109/JPROC.2015.2494218.

[237] Q. Liang et al. "Benchmarking the performance of Bayesian optimization across multiple experimental materials science domains". In: *npj Computational Materials* 7.1, 2021, p. 188. DOI: 10.1038/s41524-021-00656-9.

Acknowledgments

The organizers extend their sincere gratitude to all workshop participants for their valuable ideas and contributions. Special thanks are due to Steven Lee, Margaret Lentz, and William Spotz for their essential roles in planning and organizing the workshop. Appreciation is also extended to the Oak Ridge Institute for Science and Education, including Kutter Craig, Deneise Terry, and David Vick, for managing the logistics of the workshop. Thanks are also owed to Madie Wickstrom for designing the graphical layout of the report and brochure, and to Faith Tinnin for carefully editing the report.