Agentic Scientific Computing and Scientific Machine Learning
Poster Session Abstracts
May 9, 2026
3:15-4:00pm

Agent-Driven Simulation Data Generation for Foundation Models of Fluid Dynamics

Yajie Ji, Yale University

Scientific machine learning for fluid dynamics is currently limited by a basic data bottleneck: there is no large-scale and diverse simulation corpus comparable to the internet-scale datasets that enabled foundation models in natural language processing and computer vision. Existing CFD benchmarks cover only limited ranges of geometries and flow conditions, since each new case usually requires manual mesh generation, solver setup, and validation. To address this gap, we propose a unified framework for autonomous fluid data synthesis and foundation-model training. Our approach uses LLM-based agents to generate simulation code, execute solvers, validate physical consistency, and iteratively refine failed cases, forming a Git-based simulation community that can continuously expand the coverage and diversity of CFD data. Based on this evolving dataset, we will train a foundation neural operator for general fluid simulation and study whether scaling behavior similar to that observed in large language models also appears in fluid learning. We further propose a hybrid inference framework that couples the learned operator with classical solvers to enable zero-shot acceleration on unseen flow configurations. The project relies on tightly coupled GPU-CPU execution on local hardware, and the full pipeline, synthesized dataset, and pretrained models will be released publicly.

Physics-Based Uncertainty-Aware Machine Learning in Materials Analysis and Design

Jie Chen, Virgina Tech

Machine learning (ML) is accelerating scientific discovery by advancing our understanding of the processing–structure–property relationships that govern materials behavior. This poster presents recent progress in ML research for materials analysis and design conducted by the SEAD Lab directed, by Dr. Jie Chen at Virginia Tech. The work focuses on physics-based, uncertainty-aware ML for (1) microscopy image analysis and (2) efficient metamaterials design.

AutoNumerics: An Autonomous, PDE-Agnostic Multi-Agent Pipeline for Scientific Computing

Youran Sun, University of Maryland, College Park

PDEs are central to scientific and engineering modeling, yet designing accurate numerical solvers typically requires substantial mathematical expertise and manual tuning. Recent neural network-based approaches improve flexibility but often demand high computational cost and suffer from limited interpretability. We introduce AutoNumerics, a multi-agent framework that autonomously designs, implements, debugs, and verifies numerical solvers for general PDEs directly from natural language descriptions. Unlike black-box neural solvers, our framework generates transparent solvers grounded in classical numerical analysis. We introduce a coarse-to-fine execution strategy and a residual-based self-verification mechanism. Experiments on 24 canonical and real-world PDE problems demonstrate that AutoNumerics achieves

competitive or superior accuracy compared to existing neural and LLM-based baselines, and correctly selects numerical schemes based on PDE structural properties, suggesting its viability as an accessible paradigm for automated PDE solving.

Hierarchical Mamba Meets Hyperbolic Geometry: A New Paradigm for Structured Language Embeddings
Sarang Patil, New Jersey Institute of Technology

Selective state-space models excel at long-sequence modeling, but their capacity for language representation -- in complex hierarchical reasoning -- remains underexplored. Most large language models rely on \textit{flat} Euclidean embeddings, limiting their ability to capture latent hierarchies. To address this, we propose {\it Hierarchical Mamba (HiM)}, integrating efficient Mamba2 with hyperbolic geometry to learn hierarchy-aware language embeddings for deeper linguistic understanding. Mamba2-processed sequences are projected to the Poincaré ball or Lorentzian manifold with ``learnable'' curvature, optimized with a hyperbolic loss. Our HiM model facilitates the capture of relational distances across varying hierarchical levels, enabling effective long-range reasoning for tasks like mixed-hop prediction and multi-hop inference in hierarchical classification. Experimental results show both HiM variants effectively capture hierarchical relationships across four linguistic and medical datasets, surpassing Euclidean baselines, with HiM-Poincaré providing fine-grained distinctions with higher h-norms, while HiM-Lorentz offers more stable, compact, and hierarchy-preserving embeddings-favoring robustness.

Catalyst: A Closed-Loop Agentic Pipeline for Automated Multi-Physics Simulation Code Generation
Harshwardhan Praveen, Atomic Machines

We present Catalyst, an agentic code generation pipeline that automatically produces validated Multiphysics simulation code from high-level physics specifications. Catalyst works by a step-and-check approach: A directed acyclic graph of LLM agents writes code section-by-section, validating each section against a live compiler. When validation fails, a dedicated analysis agent diagnoses the error(s) and produces fix-instructions in a structured format, and regenerates the failed section, thus defining a closed-loop feedback for generating correct code. The design decomposes code generation into a hybrid strategy: Deterministic programmatic generators handle structurally predictable sections, while LLM agents tackle under-specified physics and study configurations guided by expert reference code and structured prompts. We evaluate the system on a suite of multi-physics benchmarks spanning solid mechanics, coupled thermomechanical analysis, and electromagnetics. We assess output correctness via an LLM judge using domain-expert rubrics. Using COMSOL, we demonstrate that structured agentic decomposition with formal validation can produce complex and reliable scientific simulation code.

Monitoring Convexity in PINNs for the Monge–Ampère Equation on the Disk
Edison Vazquez, Center for Research and Advanced Studies of the National Polytechnic Institute (CINVESTAV).

Physics-Informed Neural Networks (PINNs) provide a flexible framework for approximating solutions of nonlinear partial differential equations. For the Monge–Ampère equation, however, small residual error alone is not enough: admissible solutions must also be convex. In this work, we study PINN approximations of the Monge–Ampère equation on the unit disk, focusing on how convexity is preserved, lost, or recovered during training. In addition to standard PDE and boundary losses, we monitor the smallest eigenvalue of the Hessian across the domain and analyze one-dimensional restrictions of the learned solution along selected diameters of the disk. These diagnostics reveal local and transient violations of convexity that may remain hidden even when the residual is small. This perspective provides a geometric lens for understanding neural PDE solvers and suggests practical directions for convexity-aware regularization in fully nonlinear problems.

Hamiltonian bridge: A physics-driven generative framework for targeted pattern control
Sumit Sinha, Harvard Unviersity

Patterns arise spontaneously in a range of systems spanning the sciences, and their study typically focuses on mechanisms to understand their evolution in space-time. Increasingly, there has been a transition towards controlling these patterns in various functional settings, with implications for engineering. Here, we combine our knowledge of a general class of dynamical laws for pattern formation in non-equilibrium systems, and the power of stochastic optimal control approaches to present a framework that allows us to control patterns at multiple scales, which we dub the "Hamiltonian bridge". We use a mapping between stochastic many-body Lagrangian physics and deterministic Eulerian pattern forming PDEs to leverage our recent approach utilizing the Feynman-Kac-based adjoint path integral formulation for the control of interacting particles and generalize this to the active control of patterning fields. We demonstrate the applicability of our computational framework via numerical experiments on the control of phase separation with and without a conserved order parameter, self-assembly of fluid droplets, coupled reaction-diffusion equations and finally a phenomenological model for spatio-temporal tissue differentiation. We interpret our numerical experiments in terms of a theoretical understanding of how the underlying physics shapes the geometry of the pattern manifold, altering the transport paths of patterns and the nature of pattern interpolation. We finally conclude by showing how optimal control can be utilized to generate complex patterns via an iterative control protocol over pattern forming pdes which can be casted as gradient flows. All together, our study shows how we can systematically build in physical priors into a generative framework for pattern control in non-equilibrium systems across multiple length and time scales.

An Agentic AI Science Community for Automated Neural Operator Discovery
Luis Loo, Texas A&M University

We present an agentic implementation of the AI Science Community (AI-SC) framework (Braga-Neto, 2026) for the automated discovery of neural operators. A swarm of N virtual laboratories collaborates under a

citation-based mechanism of influence. Each lab is structured as a team of five agents: a Planner that proposes the experimental configuration, a Worker that executes the experiment via its own genetic algorithm over structural hyperparameters, an Evaluator that scores results under a multi-objective fitness criterion, a Reviewer that participates in peer review across the community, and a shared Coordinator (using a community-wide particle swarm optimization) that evolves lab strategies based on accumulated citations. The architecture space is composed of five operator-family blocks, Fourier spectral, attention, wavelet multi-scale, residual convolution, and DeepONet-style branch-trunk, together with hybrid recombinations. Cited labs gain compute budget and breeding priority; uncited labs are deactivated and replaced, realizing natural selection over research directions. Agents are model-agnostic and can be instantiated with small or large language models. Applied to a piecewise regression benchmark, 1D linear advection, 1D Burgers equation, and the 2D Navier-Stokes vorticity equation under Kolmogorov forcing, the swarm autonomously identifies which operator family best fits each regime, without human priors on architecture. The discovered hybrids beat or tie existing neural operators on the 1D problems and match a heavily tuned Fourier Neural Operator on Navier-Stokes 2D at roughly three times fewer parameters. We discuss how multi-agent coordination, citation dynamics, and operator diversity accelerate the discovery of neural PDE solvers.

Toward Autonomous CAE: Geometric Reasoning for Simulation Workflows
Stefan Gavranovic, Siemenes Industry Software GmbH.

Recent advances in machine learning are enabling new workflows that can accelerate one of the most time-consuming stages of computer engineering analysis (CEA): simulation setup. While fast meshless solvers can now deliver near-instantaneous results, the specification of boundary conditions, load locations, and other simulation inputs remains a significant bottleneck that typically requires expert knowledge. A key step toward fully autonomous end-to-end simulation is therefore geometric understanding, i.e. identifying which regions of a CAD model are functionally intended to carry loads, supports, or fixation constraints. We present a system for autonomous structural analysis that integrates a CAD environment, a meshless simulation tool, a large language model-based copilot, and a geometry segmentation model. Starting from a minimal technical design specification, the system interprets the design intent, segments the CAD geometry into functionally meaningful regions, configures the simulation model, executes the structural analysis, and automatically generates a results report. The core technical contribution is a geometry segmentation approach for native CAD representation (BReps) based on self-supervised learning and fine-tuned on a very limited set of historical simulation examples. Our approach is motivated by the central challenge of data scarcity in CAD learning. Labeled datasets for tasks such as mechanical component segmentation are typically very small, or don't contain simulation relevant segmentations, making fully supervised learning impractical. To address this challenge, we pretrain a model on unlabeled CAD geometries using self-supervised learning and fine-tune the resulting representations to identify functionally relevant surfaces, such as load-bearing and fixed regions. This direction aligns with recent research showing that combining geometric pretraining with semantically informed supervision can improve data-efficient segmentation for engineering tasks. In addition, we demonstrate the same autonomous simulation workflow coupled with generative model capabilities. Beginning with a hand-drawn 2D sketch, a diffusion-based model trained on native 3D voxel representations generates candidate part geometries, which are then passed directly into

the autonomous simulation pipeline for automatic inference of loads and boundary conditions and subsequent analysis. Together, these results illustrate a practical path toward fully autonomous design-to-simulation workflows driven by geometric deep learning.

LEADS: LEarning-Accelerated Domain Science, A Scientific Discovery through Advanced Computing Institute

Panos Stinis, Pacific Northwest National Laboratory

Scientific machine learning (SciML) aims to integrate the capabilities of artificial intelligence and machine learning into scientific computing. SciML leverages the often impressive ability of modern ML tools to represent complex mappings to uncover and quantify phenomena that have proven formidable for non-ML approaches. Thus, SciML is poised to significantly accelerate scientific discovery—the primary objective of the Scientific Discovery through Advanced Computing (SciDAC) program. The newly awarded LEADS SciDAC Institute's overarching goal is to advance the frontiers of domain science through applied mathematics research in SciML. LEADS will introduce a paradigm shift by integrating SciML directly into domain-specific challenges. It will employ a protocol to gracefully scale the SciML approach from simple examples to the full complexity of a DS problem. The Institute will develop innovative SciML techniques for complex physical phenomena, to deliver scalable SciML software optimized for DOE's computational infrastructure, and to provide customized SciML solutions addressing high-priority DS problems identified in collaboration with Department of Energy (DOE) stakeholders. Innovation, scalability, and customization will be the three pillars that LEADS builds on to create a SciDAC Institute that can benefit all aspects of domain science in DOE's purview.

AdamFLIP: Adaptive Momentum Feedback Linearization Optimization for Hard Constrained PINN Training

Binghang Lu, Purdue University

Physics-informed neural networks (PINNs) provide a flexible machine learning framework for solving forward and inverse problems governed by partial differen- tial equations (PDEs). However, standard PINN training relies on a soft-penalty for- mulation that incorporates PDE residuals and initial/boundary conditions (ICs/BCs) as weighted loss terms, often leading to ill-conditioning, sensitivity to penalty co-efficients, and lack of feasibility guarantees. In this work, we reformulate PINN training as an equality-constrained optimization problem and propose a novel constrained optimizer based on feedback linearization (FL). Our approach designs the multiplier/feedback term to shape the dynamics of constraint violations into a stable linear system, yielding a closed-form, first-order update that explicitly drives residuals toward feasibility. Building on this principle, we introduce AdamFLIP, an adaptive variant that integrates FL with Adam-style moment estimation. AdamFLIP preserves the efficiency and robustness of adaptive gradient methods while incorporating principled constraint handling. Across a range of benchmark forward and inverse PDE problems, AdamFLIP consistently achieves superior constraint satisfaction and solution accuracy compared with both standard soft-penalty PINN training using Adam and representative constrained PINN optimizers. These results highlight the effectiveness of importing control-

theoretic feedback-linearization ideas into scientific machine learning and establish FL-based optimization as a simple and powerful alternative for stable and reliable PINN training.

FEM-Bench: A Structured Scientific Reasoning Benchmark for Evaluating Code-Generating LLMs
Erfan Hamdi, Boston University

As LLMs advance their reasoning capabilities about the physical world, the absence of rigorous benchmarks for evaluating their ability to generate scientifically valid physical models has become a critical gap. Computational mechanics, which develops and applies mathematical models and numerical methods to predict the behavior of physical systems under forces, deformation, and constraints, provides an ideal foundation for structured scientific reasoning evaluation. Problems follow clear mathematical structure, enforce strict physical and numerical constraints, and support objective verification. The discipline requires constructing explicit models of physical systems and reasoning about geometry, spatial relationships, and material behavior, connecting directly to emerging AI goals in physical reasoning and world modeling. We introduce FEM-Bench, a computational mechanics benchmark designed to evaluate the ability of LLMs to generate correct finite element method (FEM) and related code. FEM-Bench 2025 contains a suite of introductory but nontrivial tasks aligned with material from a first graduate course on computational mechanics. These tasks capture essential numerical and physical modeling challenges while representing only a small fraction of the complexity present in the discipline. Despite their simplicity, state-of-the-art LLMs do not reliably solve all of them. In a five attempt run, the best performing model at function writing, Gemini 3 Pro, completed 30/33 tasks at least once and 26/33 tasks all five times. The best performing model at unit test writing, GPT-5, had an Average Joint Success Rate of 73.8%. Other popular models showed broad performance variation. FEM-Bench establishes a structured foundation for evaluating AI-generated scientific code, and future iterations will incorporate increasingly sophisticated tasks to track progress as models evolve.

Monotonicity of the First Dirichlet Eigenvalue of Regular Polygons
Joana Pech-Alberich, Brown University

In this poster we show a proof of the Antunes-Freitas conjecture (2006): for regular N-sided polygons of fixed area, the first Dirichlet eigenvalue $\lambda_1(N)$ and the quotients $\lambda_1(N)/\lambda_1(N+1)$ are monotonically decreasing in N. The proof combines two approaches: sharp asymptotic expansions with rigorous error bounds (for $N \geq 64$) and computer-assisted proofs using interval arithmetic (for $N < 64$). This is joint work with Joel Dahne and Javier Gómez-Serrano.

Quantification of modelling uncertainty in physics-informed neural networks by interval arithmetic
Zlatan Dimitrov, GATE Institute, Sofia University "St. Kliment Ohridski"

Engineering design is increasingly leveraging data-driven machine learning (ML) approaches, motivated by their recent growth in popularity and demonstrated performance gains. However, these approaches suffer from an inherent lack of physical knowledge, causing some of the produced results to be non-physical or to

require vast amounts of data to achieve convergence, unavailable in many real-world measurement scenarios. Compounding this is the tendency of some ML models to ascribe complete confidence to their outputs, limiting the ability of engineers to balance performance and safety under limited resources. Physics-informed neural networks (PINN) address the problem of spasre and noisy data, as well as the limited generalization performance of ML models, by incorporating existing physical knowledge in the form of partial differential equations. These equations serve as a regularizer on the models' outputs, enforcing compliance with known physical laws, and enabling physically consistent inference under limited data. Nevertheless, the models remain black boxes with unknown sensitivity to environmental factors, and the epistemic uncertainty stemming from the data or the model itself remains unquantified or quantified in a biased way. This work explores the augmentation of a PINN with interval-valued outputs, as a way to tackle epistemic uncertainty. The interval PINN produces an arithmetic interval within which the solution to the problem is expected to lie, thereby avoiding the assignment of any prior probability distribution to the behavior of the model. Implementation is done using numerical weights, which avoids repeated interval variables. Two parallel layers output values for the interval center and radius at each point in time and space. Validation is done on predicting the epistemic uncertainty in the oscillations of a cantilever beam. During training a physics-informed loss function is used, which enforces the corresponding Euler-Lagrange equation. The final proposal is validated on a synthetic dataset, which is generated through the analytical solution of the aforementioned problem. The proposed architecture is compared to alternative approaches, in particular: a classical physics-informed neural network, Bayesian neural network, and an L2-regularized interval neural network. The Interval PINN models structural problems by leveraging both prior physical knowledge and measurement data, while outputting a bias-free uncertain prediction interval, which enables better reliability than pure data-driven ML models.

Separable Deep Minimizing Movement Approach for solving the Allen-Cahn Equation
Nazmin Atker Mini, Michigan Technological University

In this work, we propose a novel machine learning framework for solving the Allen-Cahn equation by reformulating it as an L2 gradient flow of the Ginzburg-Landau free energy functional. Rather than solving the equation in its strong form using collocation-based approaches, which often suffer from inaccuracies in computing higher-order spatial derivatives, we adopt a minimizing movement scheme that directly evolves the system through energy minimization. This eliminates the need for space-time discretization and improves numerical stability. We introduce a separable neural network architecture that represents the phase field using a low-rank tensor decomposition, significantly reducing the number of trainable parameters while maintaining high accuracy. To further enhance robustness, we evaluate the energy functional using Gauss quadrature instead of collocation and apply a hyperbolic tangent transformation to ensure the phase field remains strictly bounded within physically admissible limits. We also establish theoretical energy stability of the proposed scheme. Our results demonstrate that the Separable Deep Minimizing Movement (SDMM) method achieves superior accuracy and efficiency compared to existing machine learning-based solvers. Notably, we observe an order-of-magnitude speed improvement over conventional finite element methods while preserving the energy-dissipative structure of the gradient flow. These findings highlight SDMM as a stable and scalable approach for phase-field modeling.