Update from Applied Mathematics program

in Advanced Scientific Computing Research (ASCR)

Presented to ASCR Advisory Committee (ASCAC)

March 27, 2019

Steven Lee, PhD
Office of Advanced Scientific Computing Research
Program Manager for Applied Mathematics
Steven.Lee@science.doe.gov

DOE Applied Mathematics program - Snapshot

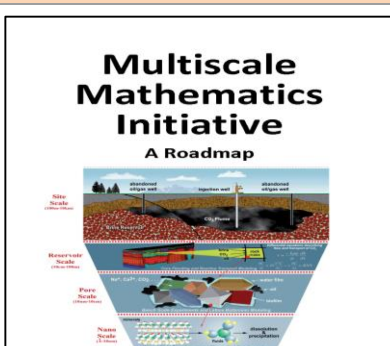
Office of Advanced Scientific Computing Research (ASCR)

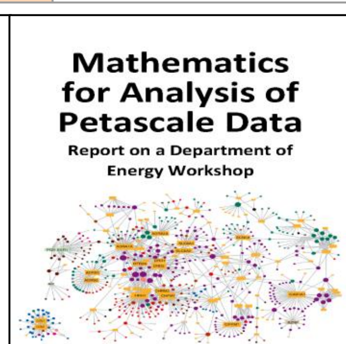
Applied Math program develops the mathematical & scientific computing foundations to accelerate the pace of scientific discoveries

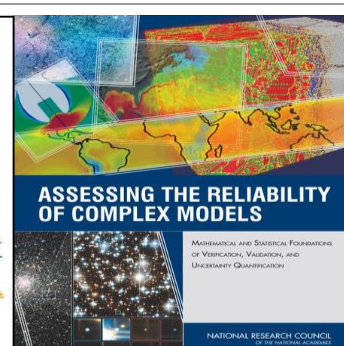
Portfolio in FY19: $30M/year for ~50 projects at Labs, universities, non-profits

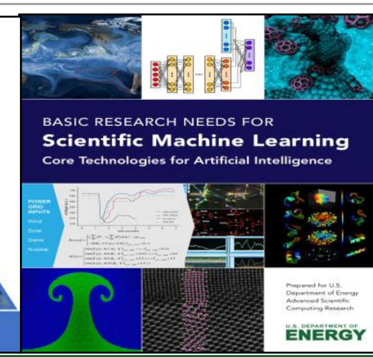
Scientific enabling technologies are being built on ASCR developments in:

Core Applied Math	Optimization, Linear algebra, Uncertainty Quantification (UQ), Differential equations, Machine Learning (ML), Meshes, Multigrid, Reduced order models
Scientific Software	High performance software codes (PETSc, Trilinos, SUNDIALS), Automatic differentiation, Parallel-in-time integrators, Meshes, Tensors, & more
Math Centers	Science at user facilities, Power Grid, Additive Manufacturing, Materials Design
Core Workforce	**Lab Base program, DOE Early Career, Lab Fellowships, University grants**
Workshops	Multiscale Math, Petascale Data, UQ, Extreme Heterogeneity, Scientific ML









Applied Mathematics Highlight – Congratulations!

SIAM/ACM Prize in Computational Science and Engineering

Jack Dongarra of the University of Tennessee.

Jack Dongarra of the University of Tennessee (UT) will receive the SIAM/ACM Prize in Computational Science and Engineering. SIAM awards this prize jointly with the Association for Computing Machinery (ACM) every two years for outstanding contributions to the development and use of mathematical and computational tools and methods for the solution of science and engineering problems. With this award, SIAM and the ACM recognize Dongarra for his key role in the development of software, software standards, software repositories, and performance and benchmarking software, as well as his community efforts to prepare for the challenges of exascale computing, especially in the adaptation of linear algebra infrastructure to emerging architectures.

Dongarra holds appointments as University Distinguished Professor of Computer Science in the Electrical Engineering and Computer Science Department, and director of both the Innovative Computing Laboratory and the Center for

Citation: SIAM and the ACM recognize Dongarra for his key role in the development of software, software standards, software repositories, and performance and benchmarking software, as well as his community efforts to prepare for the challenges of exascale computing …"

ASCR Applied Mathematics Principal Investigators Meeting January 29 & 30, 2019 at Rockville Hilton

PI Meeting Website: http://www.orau.gov/ascr-appliedmath-pi2019

- **Co-Chairs: Tamara Kolda (SNL), Aydin Buluc (LBNL)**
- **Keynote speakers: G. Strang (MIT), S. Cholia / F. Perez (LBNL, UCB)**
- **~130 Attendees PIs & teams, DOE HQ, lab management, stakeholders**
- **50 Project blitzes for 60-second summaries by each project PI**
- **~65 Project posters from project team members**
- **2 Overviews & Joint Panel on 2018 Basic Research Needs workshops**
- **Math Center Plenaries 3 MMICC projects and CAMERA**
- **Working Lunches Special interest talks: T. Kolda (SNL), A. Edelman (MIT)**

DOE Applied Math PI meeting – Jan 29, 2019

Time	Event	Location
9:00 am	**Welcome & Announcements**	Plaza I & II
9:30 am	**2018 Basic Research Needs Workshop Overviews:** **I. Scientific Machine Learning – N. Baker (PNNL)** **II. Extreme Heterogeneity – J. Vetter (ORNL)**	Plaza I & II
11:00 am	**“Functions of Deep Learning” – G. Strang (MIT)**	Plaza I & II
11:45 am	**Project Blitz – Group A**	Plaza I & II
12:30 pm	**Lunch: “Math of Data Science” – T. Kolda (SNL)**	Plaza III
1:30 pm	**Project Blitz – Group B**	Plaza I & II
2:15 pm	**Project Poster Sessions (A, B)**	Atrium
4:00 pm	**Workshop Brainstorming sessions (I, II)**	Plaza I, II, & III
5:15 pm	**End of Day 1**	

DOE Applied Math PI meeting – Jan 30, 2019

Time	Event	Location
9:00 am	**Welcome & Announcements**	Plaza I & II
9:15 am	**AEOLUS – O. Ghattas / K. Willcox**	Plaza I & II
10:00 am	**PhILMs – G. Karniadakis**	Plaza I & II
11:00 am	**MACSER – M. Anitescu / E. Constantinescu**	Plaza I & II
11:45 am	**CAMERA – J. Sethian & team**	Plaza I & II
12:30 pm	**Lunch: "Julia – Programming Language for Scientific Computing " – A. Edelman (MIT)**	Plaza III
1:45 pm	**Project Poster Session - continued**	Atrium
2:30 pm	**"Jupyter – Interactive Platform for Scientific Computing" – S. Cholia / F. Perez (LBNL, UC Berkeley)**	Plaza I & II
3:15 pm	**Wrap-Up & Adjourn – S. Lee (ASCR)**	

Recent Awards – DOE Early Career Research Program
ASCR Applied Mathematics

Name	Institution	Title	FY	Topic Area
Lin Lin	University of California, Berkeley	**Green's function methods for multiphysics simulations**	FY 2017	Multiscale Mathematics
James Ostrowski	University of Tennessee, Knoxville	**Symmetric convex sets: Theory, algorithms, and application**	FY 2017	Algorithms, Solvers, and Optimization
Siqian Shen	University of Michigan	**Extreme-scale stochastic optimization and simulation via learning-enhanced decomposition & parallelization**	FY 2017	Algorithms, Solvers, and Optimization
Timothy Wildey	Sandia National Laboratories	**Enabling beyond forward simulation for predictive multiscale modeling**	FY 2017	Multiscale Mathematics
Paris Perdikaris	University of Pennsylvania	**Probabilistic data fusion and physics-informed machine learning**	FY 2018	Scalable Scientific Data Analysis
Omer San	Oklahoma State University	**Physics-reinforced machine learning algorithms for multiscale closure model discovery**	FY 2018	Multiscale Mathematics
Eric Cyr	Sandia National Laboratories	**Parallel-in-layer methods for extreme-scale machine learning**	FY 2018	Algorithms, Solvers, & Optimization
Benjamin Peherstorfer	New York University	**Operator inference on manifolds for learning physically consistent models from data**	FY 2018	Scalable Scientific Data Analysis

Extreme-Scale Stochastic Optimization and Simulation via Learning-Enhanced Decomposition and Parallelization – Siqian Shen (U. Michigan)

Extreme-Scale Stochastic Optimization and Simulation via Learning-Enhanced Decomposition and Parallelization

Siqian Shen (PI), Department of Industrial and Operations Engineering, University of Michigan

UNIVERSITY OF MICHIGAN

I: Learning-enhanced Benders Decomposition

- Consider Benders Decomposition (BD) for solving two-stage stochastic integer programs:

$$\min_{x\in X} \quad c^T x + \sum_{\omega\in\Omega} p_\omega Q_\omega(x) \quad \text{where} \quad Q_\omega(x) \overset{\text{def}}{=} \min_{y} \quad q_\omega^T y \quad \text{s.t.} \quad W_\omega y = h_\omega - T_\omega x.$$

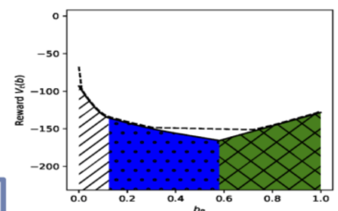
- We develop LearnBD algorithm using support vector machines (SVMs) to learn and only generate "valuable cuts" to relaxed master problems.
  - Phase 1: Cut Sampling
  - Phase 2: SVM Classifier Construction
  - Phase 3: Cut Selection
- Preliminary Results:

Method	100 Scenarios Traditional *BD*	100 Scenarios *LearnBD*	1000 Scenarios Traditional *BD*	1000 Scenarios *LearnBD*
Ratio of Bind cuts to reach 0.1%	33.8%	**62.3%**	18.6%	**56.2%**
Ratio of Bind cuts to reach 10%	20.0%	**70.3%**	23.1%	**73.5%**

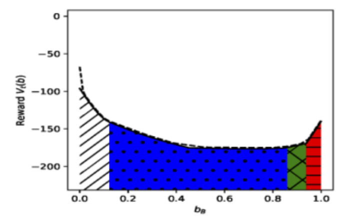
Method	Number of Iterations	Number of Added Cuts	Number of Tight Cuts	Ratio of Tight Cuts in the End	Time for Solving RMP
Traditional BD to 10%	4	4000	924	23.1%	0.665
LearnBD to 10%	**3**	3735	**924**	**24.7%**	**0.601**
Traditional BD to 0.1%	6	5616	1046	18.6%	4.642
LearnBD to 0.1%	**5**	5735	**1210**	**21.1%**	**4.264**

II: Distributionally Robust POMDP

- Consider POMDP with ambiguous transition-observation probabilities.
- Develop Bellman equations for DR-POMDP and show that the value function is convex with respect to belief states, which then can be solved by a heuristic search value iteration (HSVI) method.
- Preliminary Results (Machine Repair and ROCKSAMPLE problems):

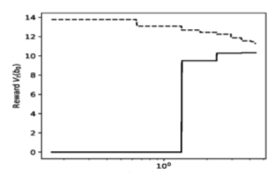


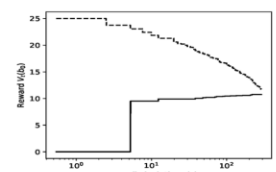
POMDP



DR-POMDP

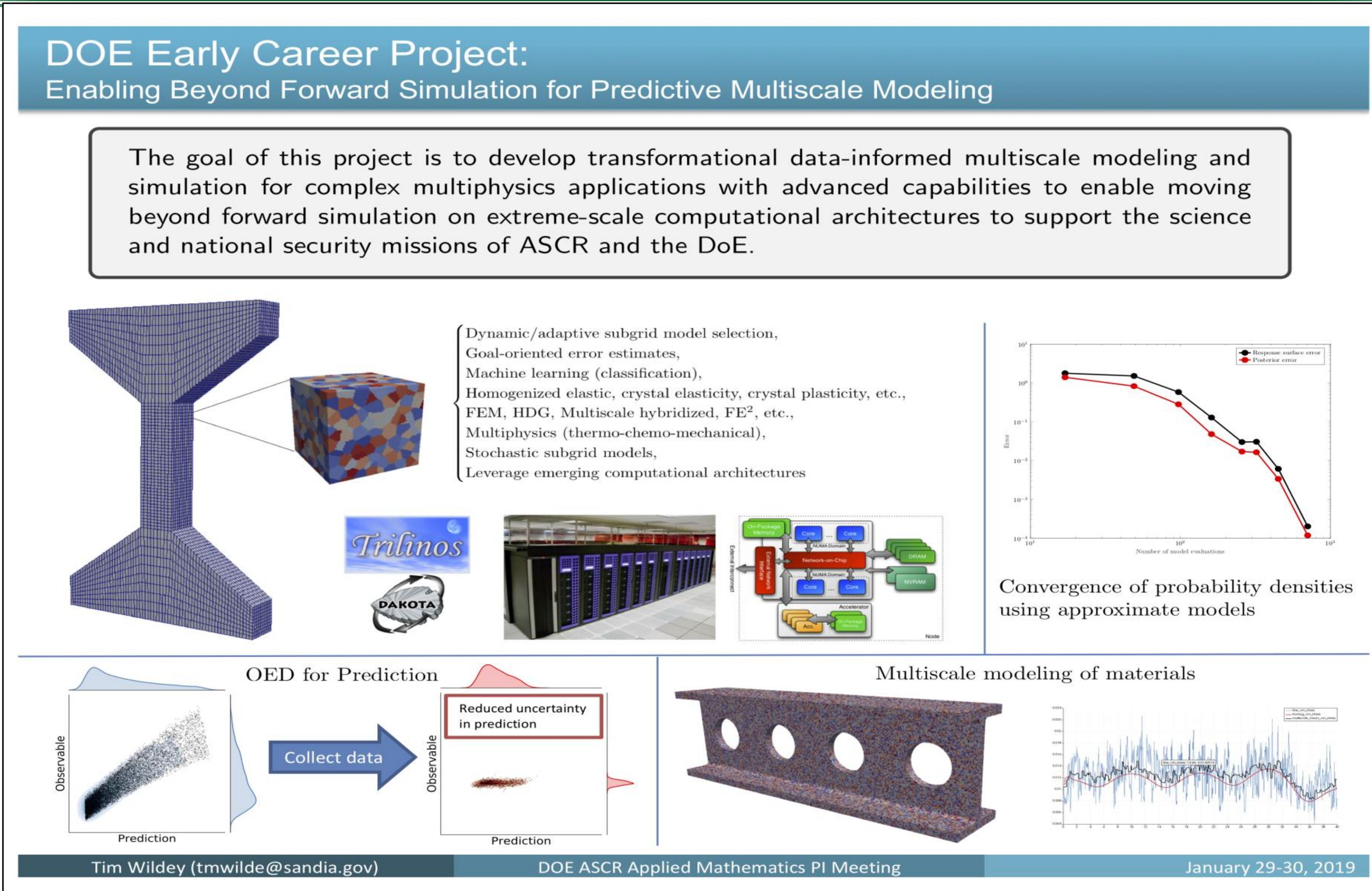
Value Function





CPU time Upper & Lower Bounds

Enabling Beyond Forward Simulation for Predictive Multiscale Modeling – Timothy Wildey (Sandia National Laboratories)



Probabilistic Data Fusion and Physics-informed Machine Learning – Paris Perdikaris (University of Pennsylvania)

Probabilistic data fusion and physics-informed machine learning

PI: Paris Perdikaris (pgp@seas.upenn.edu)
Department of Mechanical Engineering and Applied Mechanics
University of Pennsylvania

This project is focused on the development of mathematical methods and algorithms for:

1. Probabilistic data fusion and multi-fidelity modeling:
for seamlessly synthesizing heterogeneous high-dimensional data of variable fidelity.

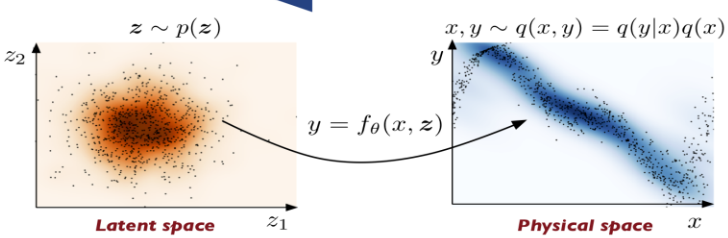
2. Probabilistic physics-informed machine learning:
for discovering and predicting complex dynamics from incomplete models and incomplete data.

3. Active learning and Bayesian optimization with deep generative models:
for effectively exploring high-dimensional spaces in pursuit of feasible and/or optimal configurations.

4. Asynchronous optimization of machine learning algorithms:
for tackling realistic applications and enhancing performance in exa-scale computing environments.

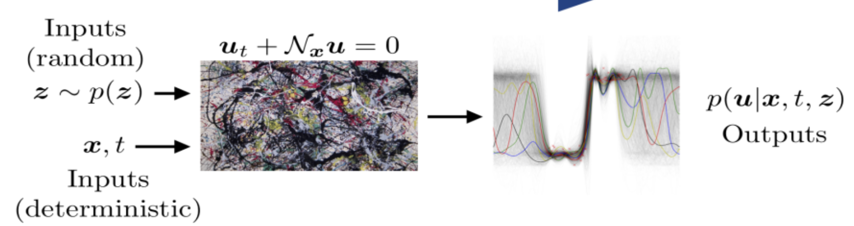
5. Applications (in collaboration with PNNL):
(i) Accelerating the discovery of new materials, (ii) stochastic modeling of sub-surface transport processes.

Data ⟷ Models



$$y = f_\theta(x, z),\ z \sim p(z) \ \Leftrightarrow\ y \sim p_\theta(y|x, z)$$

Poster #1: Conditional deep surrogate models for stochastic, high-dimensional, and multi-fidelity systems

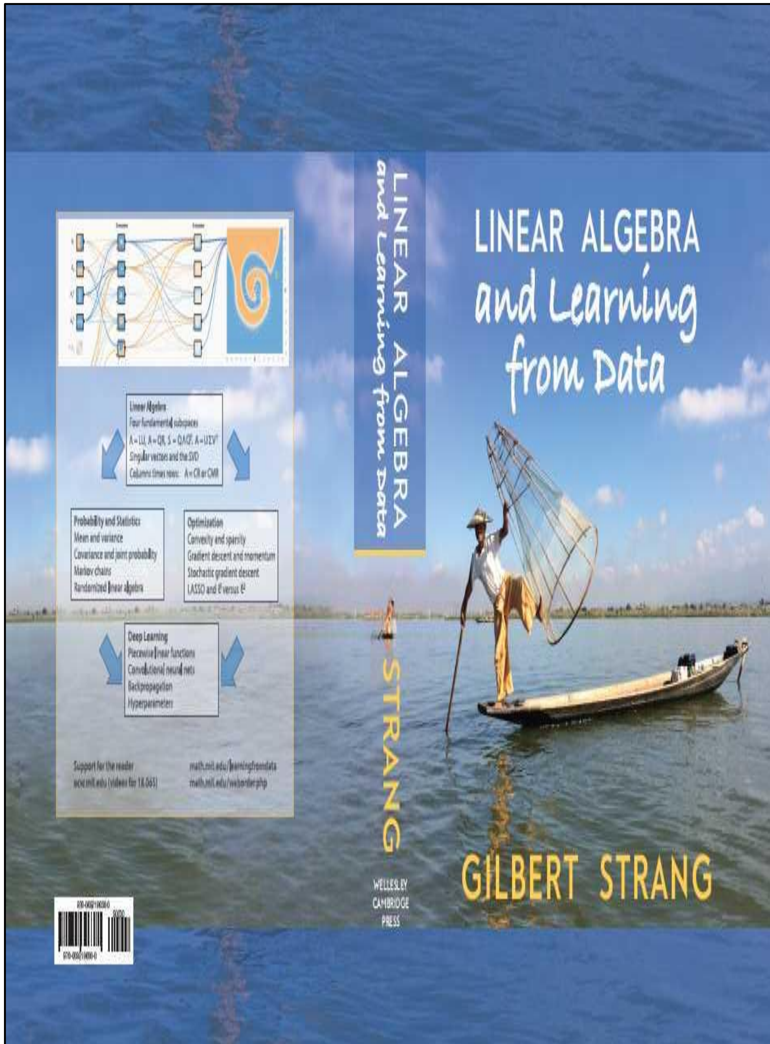


$f_\theta(x, t)$: Neural network
$r_\theta(x, t)$: Physics-informed neural network

$$\begin{cases} [x, t] \xrightarrow{f_\theta} u(x, t) \\ [x, t] \xrightarrow{r_\theta} \frac{\partial}{\partial t} f_\theta(x, t) + \mathcal{N}_x f_\theta(x, t) \end{cases}$$

Poster #2: Physics-informed deep generative models

“The Functions of Deep Learning”
Gil Strang (MIT)



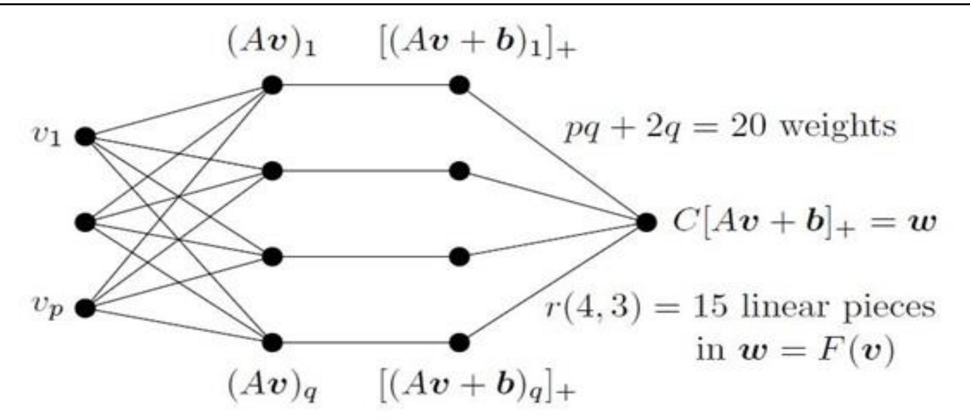
Construction of Deep Neural Networks

1	Key operation	Composition $F = F_3(F_2(F_1(\mathbf{x}, \mathbf{v}_0)))$
2	Key rule	Chain rule for x-derivatives of F
3	Key algorithm	Stochastic gradient descent to find $\mathbf{x}$
4	Key subroutine	Backpropagation to compute grad F
5	Key nonlinearity	ReLU $(y) = \max(y, 0) =$ ramp function

$$\textbf{Layer k} \quad \mathbf{v_k} = \mathbf{F_k}(\mathbf{v_{k-1}}) = \mathbf{ReLU}(\mathbf{A_k v_{k-1}} + \mathbf{b_k})$$

Weights $\mathbf{x}$ for layer k $\mathbf{A_k}$ = matrix and $\mathbf{b_k}$ = offset vector

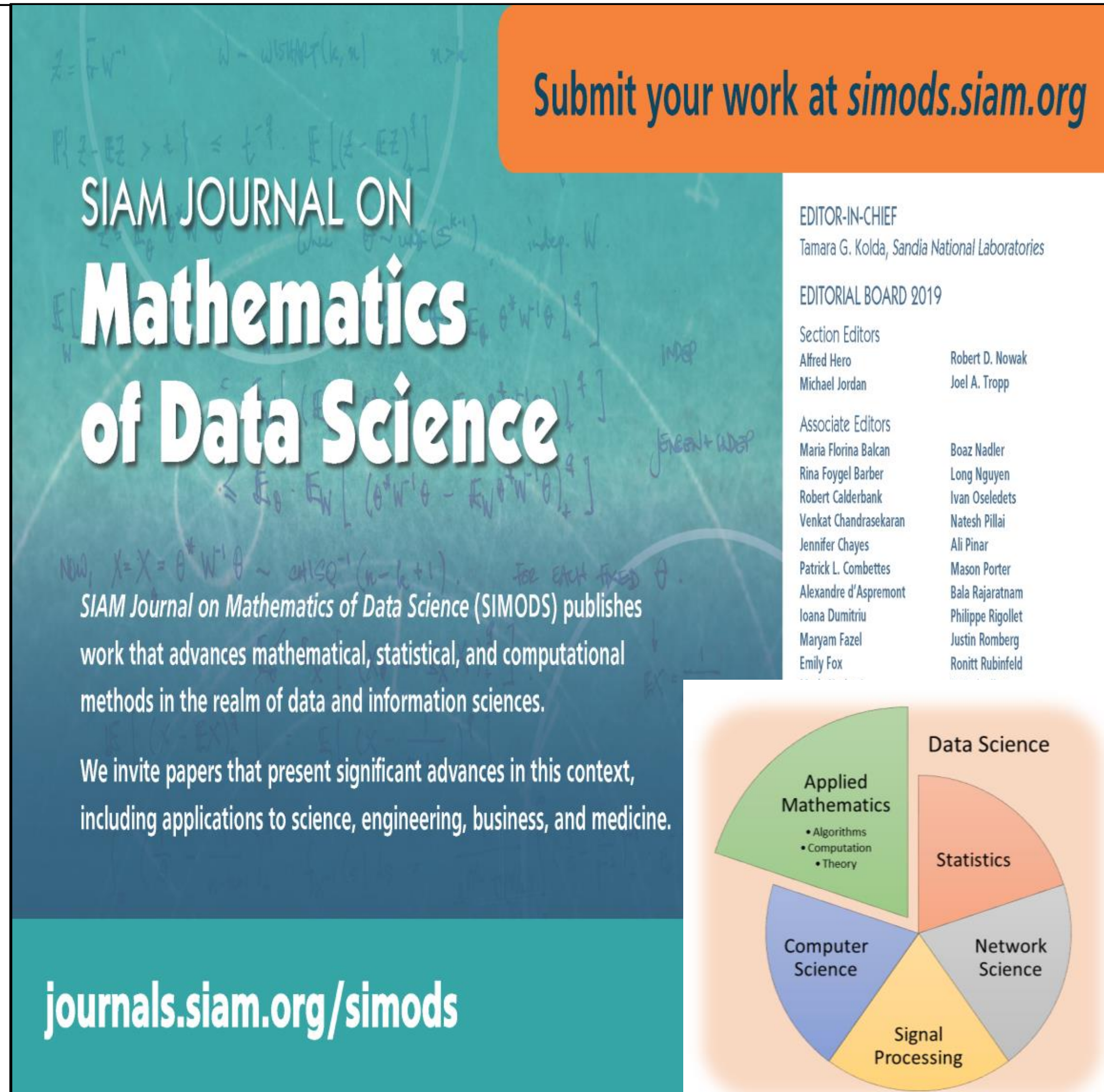
$\mathbf{v}_0$ = training data / $\mathbf{v}_1, \ldots, \mathbf{v}_{\ell-1}$ hidden layers / $\mathbf{v}_\ell$ = output



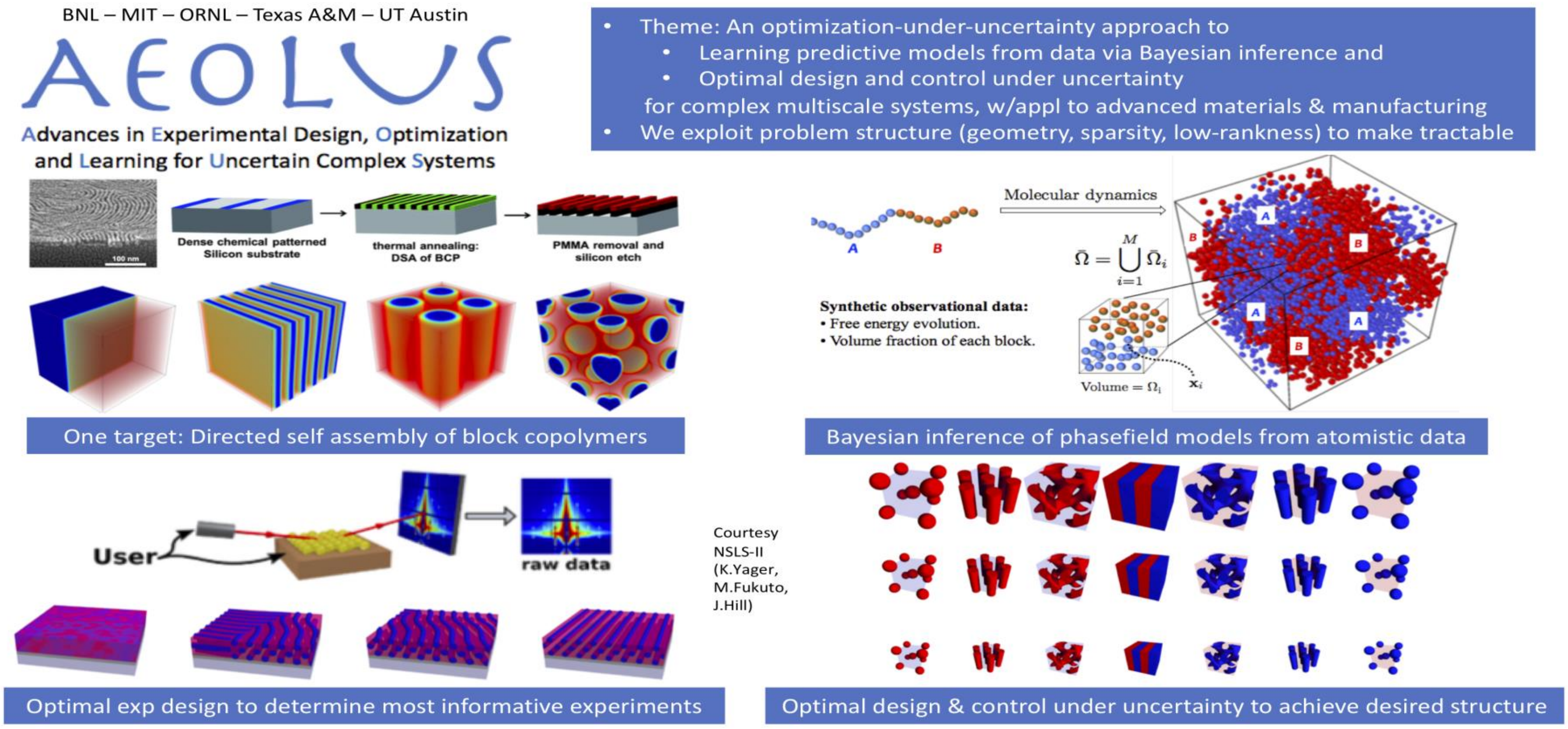
https://sinews.siam.org/Details-Page/the-functions-of-deep-learning

New SIAM Journal on Mathematics of Data Science (SIMODS) Highlights Role of Math in Advancing Data Science

- **Curated articles on key role of mathematics in advancing data science**
- **Rapid developments in hardware & software architectures for data management (peta-, zettabytes)**
- **Algorithm advances in massive data analysis require mathematical innovations**
- **Applying machine learning models is not the same as advancing the mathematics of data science**
- **Submissions exceed those of any prior SIAM launch (first 6 months)**
- **SIMODS Editor-In-Chief:**
  Tamara G. Kolda (Sandia)
- **SIAM Conference on Mathematics of Data Science: May 2020**



AEOLUS – Advances in Experimental Design, Optimization and Learning for Uncertain Complex Systems: Ghattas / Willcox (UT Austin)

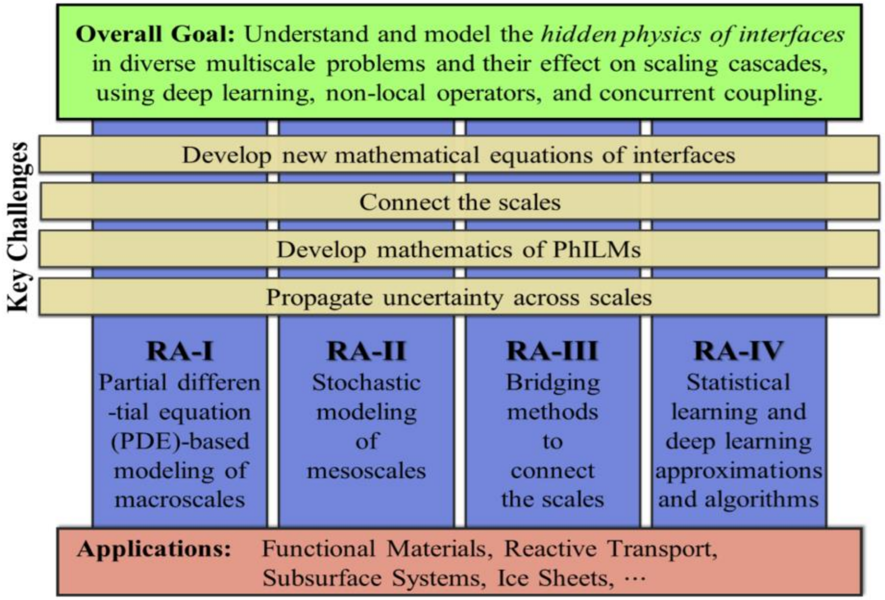


PhILMs – Physics-Informed Learning Machines for Multiscale and Multiphysics Systems: G. Karniadakis (PNNL)

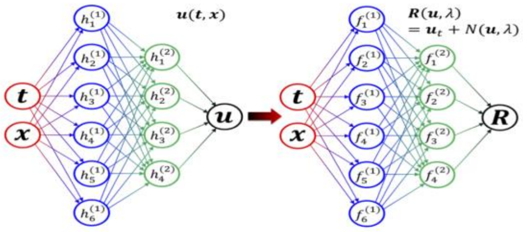
PhILMs: Collaboratory on Mathematics and Physics-Informed Learning Machines for Multiscale and Multiphysics Problems

PI: G. Karniadakis (Brown), co-PIs: A. Tartakovsky (PNNL), M. Parks (SNL), M. Ainsworth (Brown), E. Darve (Stanford), and P. Atzberger (UCSB)

Pacific Northwest NATIONAL LABORATORY



Approach: Develop physics-informed learning machines by encoding conservation laws and prior physical knowledge into Bayesian deep learning networks and analyzing their mathematical properties.



1

MACSER – Multifaceted Mathematics for Rare, High Impact Events in Complex Energy and Environment Systems: M. Anitescu (ANL)

Multifaceted Mathematics for Rare, High Impact Events in Complex Energy and Environment Systems(MACSER)

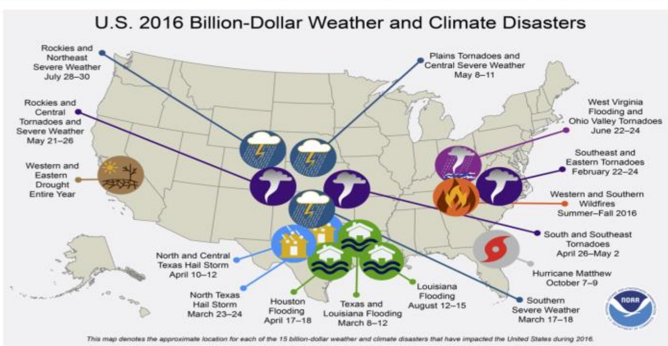
Project Director: Mihai Anitescu, Argonne National Lab

Goals:

- By taking a holistic view, quantify the occurrence and features of rare, high-impact events and design and optimize energy systems that withstand such events and recover from them.
- Address the mathematical and computational complexities of extreme space-time statistics of environmental events and its impact on analyzing, planning, and operating the energy infrastructure

Integrated Novel Mathematics Research:

- *Statistics of rare space-time events* that characterize distributions of extremes and efficiently sample from them.
- *Novel formulations for optimization under uncertainty* that best balance worst-case and probabilistic energy systems requirements.
- *Advanced optimization algorithms* that employ model reduction and decomposition to address the feature and rare event complexities.



Location of the 15 weather events in 2016 with more than $1B in damages. A large part due to energy infrastructure damage (NOAA)

Long-Term DOE Impact:

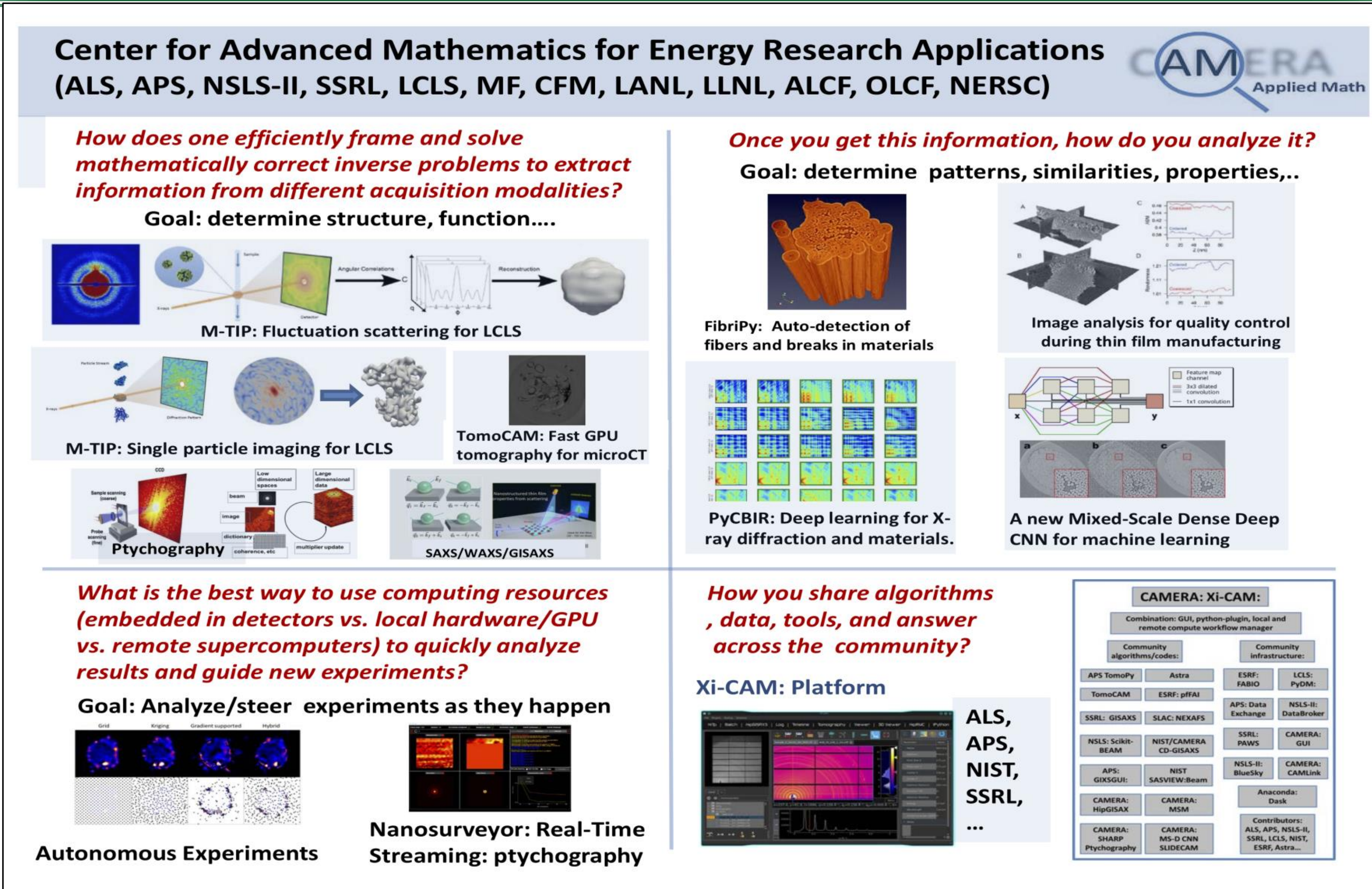
- Development of new mathematics at the intersection of multiple mathematical sub-domains.
- Addresses a broad class of applications for complex energy systems, such as :
  - Limiting the probability of cascading failures such as the 2003 blackout.
  - Planning an energy infrastructure that is robust and resilient to rare weather events.

U.S. DEPARTMENT OF ENERGY | Office of Science

Team: Argonne National Lab (Lead), Pacific Northwest NL, Lawrence Livermore NL University of Wisconsin, University of Chicago, Ohio State University

1

CAMERA – Center for Advanced Mathematics for Energy Research Applications: J. Sethian (LBNL)



"Julia - A Programming Language for Scientific Computing"
Alan Edelman (MIT)

JuliaCon 2019 Call for Proposals has been issued and Early Bird Tickets are available

julia Download Documentation Blog Community Learning Research GSoC Donate

The Julia Programming Language

Download v1.1 Documentation

Star 21,308

Julia in a Nutshell

Julia is fast!

Julia was designed from the beginning for high performance. Julia programs compile to efficient native code for multiple platforms via LLVM.

Dynamic

Julia is dynamically-typed, feels like a scripting language, and has good support for interactive use.

Optionally typed

Julia has a rich language of descriptive datatypes, and type declarations can be used to clarify and solidify programs.

General

Julia uses multiple dispatch as a paradigm, making it easy to express many object-oriented and functional programming patterns. The standard library provides asynchronous I/O, process control, logging, profiling, a package manager, and more.

Easy to use

Julia has high level syntax, making it an accessible language for programmers from any background or experience level.

Open source

Julia is free for everyone to use, and all source code is publicly viewable on GitHub.

SIAM – 2019 James H. Wilkinson Prize for Numerical Software

MIT News Browse or Search

Three co-creators of the MIT-incubated Julia programming language, (l-r) Stefan Karpinski, Viral Shah, and Jeff Bezanson, will receive the 2019 James H. Wilkinson Prize for Numerical Software. MIT Julia Lab Leader and Professor Alan Edelman is pictured at right.

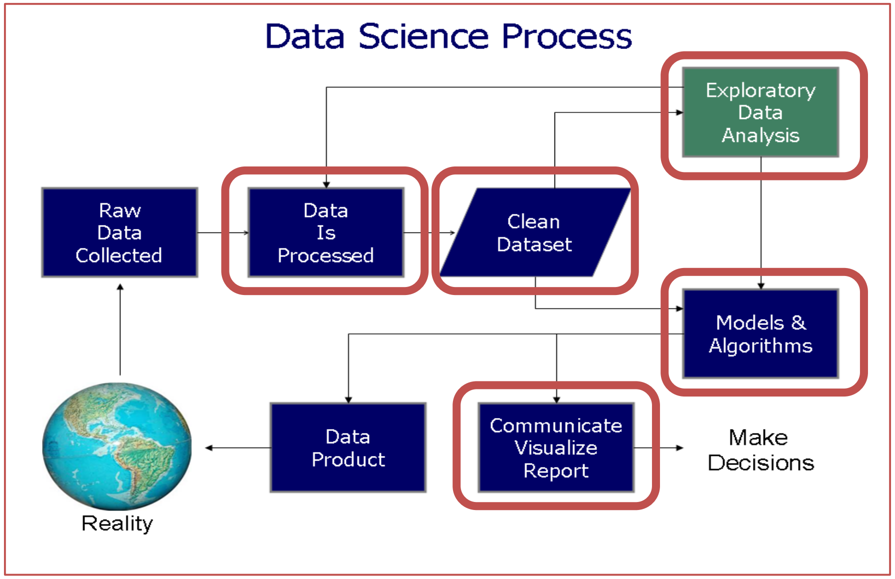
Photo courtesy of the prize winners.

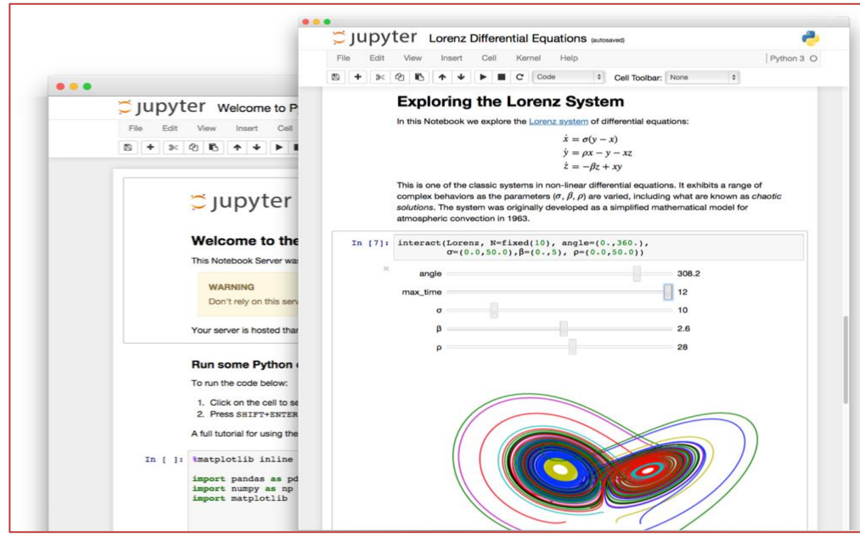
Julia language co-creators win James H. Wilkinson Prize for Numerical Software

Awarded every four years, the Wilkinson Prize last came to MIT in 1999.

"Jupyter – An Interactive Platform for Scientific Computing"
Shreyas Cholia / Fernando Perez (LBNL, UC Berkeley)

What is Jupyter? An open-source web application that allows you to create and share documents that contain live code, equations, visualizations, and narrative text. Uses include: data cleaning and transformation, numerical simulation, statistical modeling, data visualization, machine learning, and more.





2017 ACM Software System Award: "… ***a de facto standard for data analysis in research, education, journalism and industry.*** Jupyter has broad impact across domains and use cases. Today more than ***2,000,000 Jupyter notebooks are on GitHub***, each a distinct instance of a Jupyter application—covering a range of uses from technical documentation to course materials, books and academic publications."

Applied Math program has laid the groundwork to harness Machine Learning and Artificial Intelligence for scientific purposes

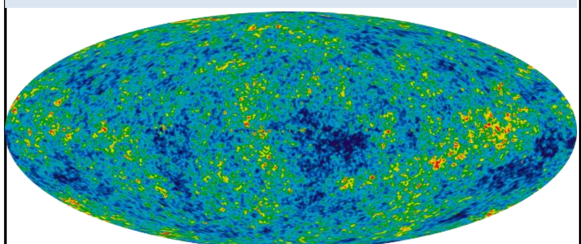
Scientific progress will be driven by

- **Massive Data: sensors, simulations, networks**
- **Predictive Models & Adaptive Algorithms**
- **Heterogeneous High-Performance Computing**
- ➢ **Scientific Machine Learning & AI**

Trend: Human-AI collaborations will transform the way science is done.

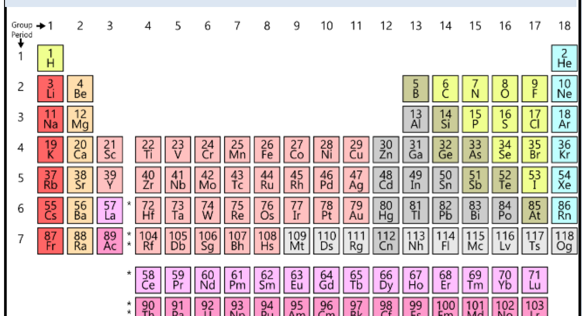
Exemplars of Scientific Achievement

Cosmic Microwave Background

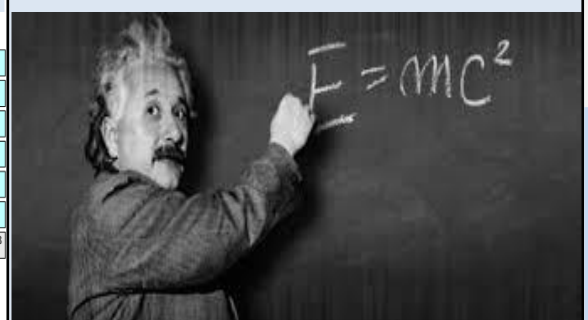


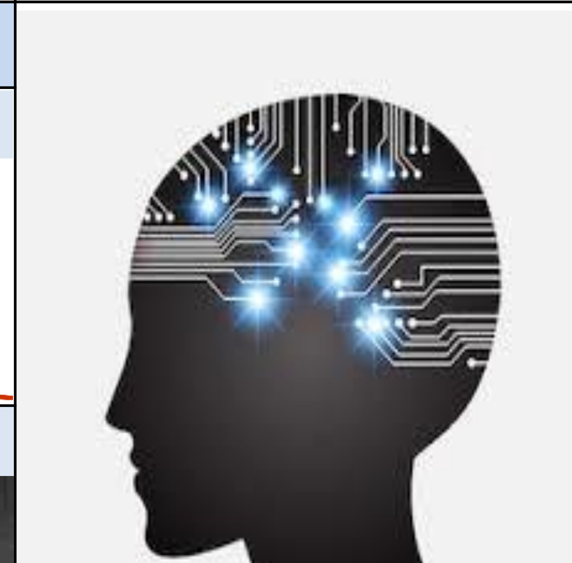
DNA Structure

Periodic Table of the Elements



Special Relativity





Human-AI insights enabled via scientific method, experimentation, & AI reinforcement learning.

DOE Applied Mathematics Research Program
➢ Scientific Machine Learning Workshop (January 2018)
SciML Workshop Report: https://www.osti.gov/biblio/1478744

DOE Applied Mathematics: Preparing for the Future

Areas of Interest through Lens of Applied Math & Scientific Computing

I. Scientific Machine Learning and AI: Foundations & Use Cases

Data-Intensive science; Adaptive modeling & simulation; Intelligent automation & decision support

II. Post-Moore Algorithms & Programming

Same algorithms & ways of programming for increasingly heterogeneous computing systems?

III. Enabling Technologies for Complex Systems Research

How to troubleshoot & improve the operational capabilities of scientific facilities, materials synthesis, power grid, additive manufacturing, or key components of complex processes and systems?

Foundations & Capabilities	History of Funding Opportunity Announcements
Massive Scientific Data Analysis & Solvers	2009 – 2012: Mathematics for Analysis of Petascale Data 2009 – 2012: Joint Mathematics Computer Science Institute 2012 – 2015: Resilient Extreme-Scale Solvers 2013 – **2016**: DOE Data-Centric Science at Scale
Predictive Algorithms, Modeling & Simulation	2005 – 2008: Multiscale Mathematics Research and Education 2008 – 2011: Multiscale Mathematics for Complex Systems 2013 – **2016**: Uncertainty Quantification for Extreme-Scale Science
Complex Systems & Processes	2009 – 2012: Mathematics for Complex, Interconnected Systems 2010 – 2013: Uncertainty Quantification for Complex Systems 2012 – 2017: Mathematical Multifaceted Integrated Capability Centers I **2017 – 2022**: Mathematical Multifaceted Integrated Capability Centers II

Long-Term Scientific Computing Research
DOE Applied Math Themes

ASCR Criteria for Themes I - III	I. Adaptive High-Performance Algorithms & Solvers
Basic research will develop & sustain 1. Computational Leadership 2. Discovery-Enabling Technologies 3. S/W Tools, Prototypes, Ecosystems 4. High-Tech DOE Workforce in advanced scientific computing.	Adaptive HPC solvers & Scalable data analysis. Forward, Inverse, & Optimization problems. Post-Moore computational motifs/patterns. * Massively parallel, Asynchronous, Ensembles * Preconditioners, Statistics, Learning from data * Randomization, Graphs, Black/Gray boxes * Poly-algorithms, Adaptive precision
II. Predictive Multifaceted Modeling & Simulation	**III. Integrated Capabilities**
Predictive scientific computing. Hierarchical, coupled, & hybrid simulations. Novel formulation, meshing, & interfaces. * Smart machine learning-enhanced models * Data-driven models, Surrogate sub-models * Uncertainty & Error propagation, Validation * Forward, Adjoint, & Parameter sensitivities	Intelligent automation & decision-support. Complex systems, processes, & infrastructure. Effective human-AI collaboration. * AI-enhanced scientific method & discoveries * Optimal experimental design, Outer loop * Systems resilience, reliability, & control * Automated performance optimization