Domain Overview: Astronomical Sciences

10 minutes presentation + 5 minutes Q&A

Yuan-Sen Ting, The Ohio State University

AI + Astronomical Sciences (AST)

- **Inverse Problems with Simulation Based Inference:** Enables intractable computational tasks through robust inference **without relying on simplistic summary statistics**.

- **Representation Identification:** Accelerates identification of **representations** across domains - from galaxy morphology, gravitational lens to stellar parameter estimation.

- **Multimodality Integration and World Model:** **Combine heterogeneous data types** (imaging, spectroscopy, time series) to construct more complete physical models.

- **Anomaly Detection:** **Deep generative models** improve discovery of novel astronomical phenomena and outliers in large datasets.

AI in Astronomical Science (Inverse Problem)

AI techniques extract maximum physical information from observations through efficient simulation-based inference, enabling parameter estimation from complex data.

- **Cosmology:** Field-level inference methods going beyond summary statistics to constraint cosmology parameters.
- **Gravitational wave astronomy:** Modeling **complex inverse problems** like spin precession, dramatically reduced computational costs for real-time alerts.
- **Exoplanet research:** Efficient exploration of **multimodal posterior distributions**. Integration of multiple observation modes for parameter estimation.
- **Stellar astrophysics:** Advancing spectral analysis. Analyzing stellar oscillation data to determine interior properties **beyond human heuristics**

AI in Astronomical Science (Anomaly Detection)

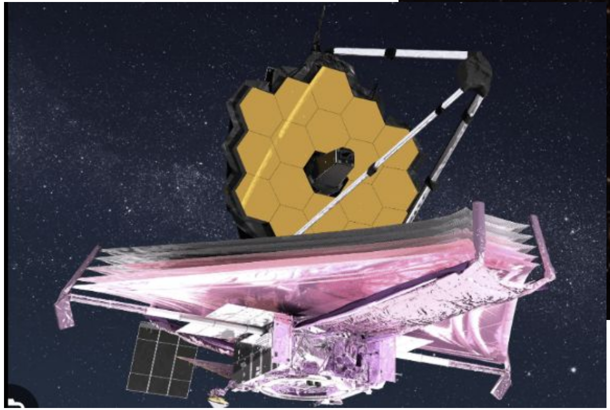
AI enables detection of novel phenomena in datasets without theoretical biases

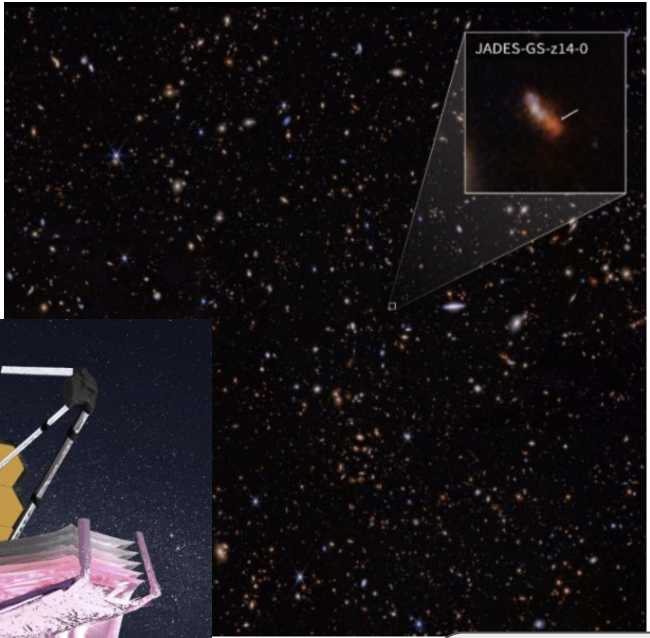
- **Time-domain astronomy:** Handling of **irregular sampling and heterogeneous noise** in light curves enables early detection of transients.
- **Galaxy evolution studies:** Understanding evolution through **latent space representations.** Creation of synthetic populations for training **data augmentation**.
- **Solar and heliophysics:** Identification of features like sunspot groups and space weather events. Neural fields for **improved reconstruction** of solar surface features.
- **Stellar astrophysics:** Identification of rare stellar types. Automated pipelines eliminate false positives in searches for stellar-mass black holes and eclipsing binaries.

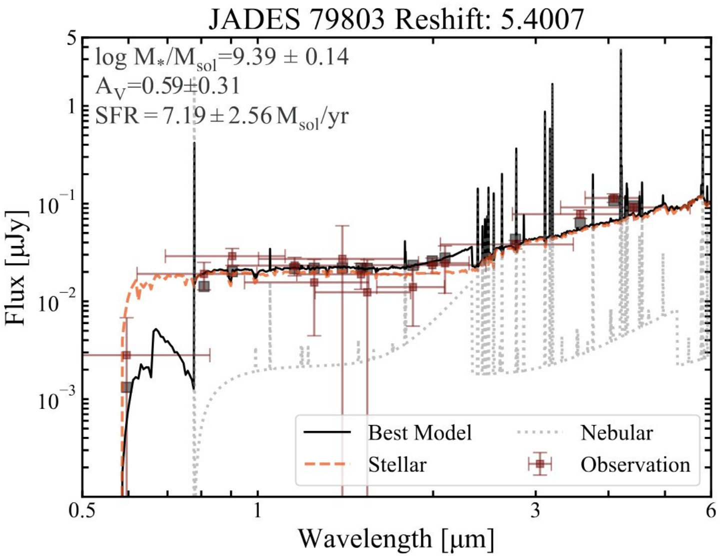
Yuan-Sen Ting, The Ohio State University (AST)

- **Astro Statistician** working mostly at tackling inverse problems (stellar astrophysics and cosmology).
- Care deeply about **robust uncertainty quantification** through proper Bayesian treatment (boo, AI). Responsible AI
- Train **domain specific foundational models** (spectral analysis) with robust domain transfer ability. (yay! AI)
- Advancing **LLM agents** for autonomous astronomical research (led the AstroMLab team - with people from Oak Ridge / Argonne)
- Training specialized LLMs (AstroSage / AstroLLaMA), creating **agent benchmarking**, and applying agents to accelerate discovery.

Yuan-Sen Ting, Astro-Statistician goes Rogue







How to reason about ALL James Webb's observations autonomously?

Yuan-Sen Ting, Educator goes Rogue

Computational Astronomy Tutor

ASTRON 5550 Advanced Astronomical Data Analysis

Tips for Getting Started 🚀

I'm here to help you succeed in your computational astronomy course! I can assist with:

- Course concepts and theoretical explanations
- Data analysis and coding support
- Paper recommendations for assignments and projects

nature astronomy

Explore content ⌄ About the journal ⌄ Publish with us ⌄ Subscribe

nature › nature astronomy › world view › article

World View | Published: 28 February 2025

Artificial intelligence compels the astronomy community to rethink research identity and redefine excellence

Yuan-Sen Ting

Teaching my astro-statistics course with LLM agents

Future Opportunities: AI in Astronomy

- **Astronomical Foundation Models:** Developing models trained across multiple wavelengths and phenomena to enable robust transfer learning.
- **Multi-Agent Research Systems:** Integrated workflows from literature review to theoretical interpretation, accelerating the observation-to-discovery cycle.
- **Responsible AI Adoption:** Establishing astronomical benchmarks and evaluation protocols to maintain scientific rigor.
- **Physics-Informed AI:** Integrating physical principles, symmetries, and conservation laws directly into model architectures to ensure results respect underlying physical reality.
- **Next-Gen Observatory Support:** AI will be crucial for handling petabyte/exabyte-scale datasets from Rubin Observatory, Roman Space Telescope, and Square Kilometre Array.

AI + AST Questions to Consider

- **Besides** making things "faster" / better inverse problem solving, can AI make physically motivated / interesting discoveries?

- Can AI make interesting discoveries **by itself**? If not, to what extent can it contribute?

- What is the utility of AI in u**nderexplored subdomains** of astrophysics?

- What are the key challenges in creating a **healthy ecosystem** of AI × Astronomy?

Driving Question: How can the MPS domains best capitalize on, and contribute to, the future of AI?

Domain Overview: Chemistry

10 minutes presentation + 5 minutes Q&A

Pratyush Tiwary, University of Maryland

AI + Chemistry (CHE)

AI’s impact on Chemistry

- **Accelerated Molecular Discovery:** AI-driven models can help design molecules for drug discovery, catalysis, and materials science.
- **Smarter Reaction Prediction:** Machine learning improves reaction outcome forecasting, retrosynthesis, and arguably mechanistic understanding.
- **Faster & More Accurate Simulations:** AI enhances quantum chemistry, molecular dynamics, and force field accuracy, reducing computational costs.
- **Automated Chemical Experimentation:** AI-guided lab automation optimizes synthesis, high-throughput screening, and data extraction.

These highlights are summarized from responses to the survey–more to be discussed in the breakout groups!

AI in Chemistry Subdomains (Examples)

- **Drug Discovery:** AI can accelerate virtual screening, molecular docking, and de novo drug design, identifying potential therapeutics faster.
- **Catalysis & Reaction Engineering:** ML can predict catalyst performance and reaction mechanisms, accelerating sustainable chemistry innovations.
- **Materials Chemistry:** AI discovers new polymers, batteries, and semiconductors, optimizing properties with minimal experiments.
- **Quantum Chemistry & Molecular Simulations:** AI-driven force fields can enable long-timescale simulations with near-quantum accuracy; Enhanced sampling methods can generate rare event statistics at timescales of minutes and slower
- **Cheminformatics & Structural Analysis:** AI enhances spectral interpretation, molecular fingerprinting, and 3D structure prediction (e.g., AlphaFold).

Pratyush Tiwary, U Maryland (CHE)

I merge AI with statistical physics to simulate protein, crystals & RNA across otherwise unreachable timescales and with limited training data.

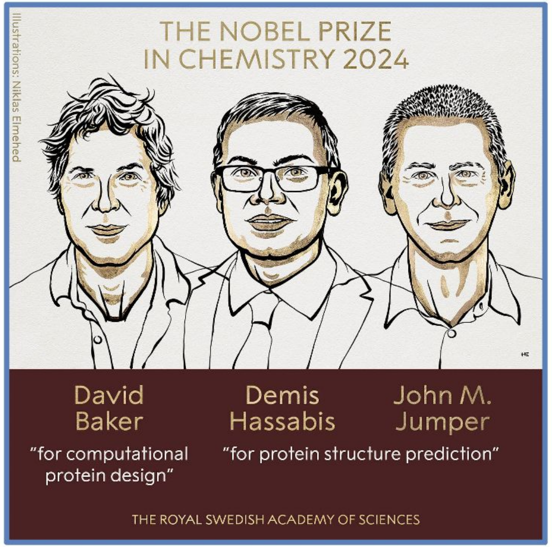
Key Issue for Your Domain

- In computational chemistry, we often lack high-quality data at the timescales and system sizes required to capture rare events (e.g., slow protein conformations or RNA folding). This shortage of "right data" undermines many AI-driven methods, which then struggle with out-of-distribution generalization or extrapolation beyond the training set.
    - How do we ensure AI predictions remain physically valid in data-sparse settings?
    - What methods promote robust out-of-distribution performance and true extrapolation?
    - Chemistry needs precise environment. Which strategies best incorporate environmental factors (e.g., temperature, pH) into AI models for more realistic outcomes?

Domain-Specific Case Study/Example

- A recent success (Herron et al., PNAS 2024) predicted phase transitions in Ising and RNA systems without sampling near the critical point. By embedding thermodynamics into AI, we inferred critical exponents from sparse data—showcasing statistical physics-informed AI for emergent phenomena.

Exemplar 1: GPU-enabled success in structure prediction and protein design



Structure prediction and protein design

During gold rush, buy shovels

Exemplar 2: Generative AI *might* transform molecular dynamics and ensemble prediction

The New York Times

› OpenAI's 'Reckless' Culture | The New ChatGPT | Replacing the C.E.O. | Chatbots and Disinformation

Google Unveils A.I. for Predicting Behavior of Human Molecules

The system, AlphaFold3, could accelerate efforts to understand the human body and fight disease.

nature machine intelligence

Article https://doi.org/10.1038/s42256-024-00837-3

Predicting equilibrium distributions for molecular systems with deep learning

Received: 2 August 2023
Accepted: 10 April 2024
Published online: 8 May 2024
Check for updates

Shuxin Zheng[1,5], Jiyan He[1,2,5], Chang Liu[1,5], Yu Shi[1,5], Ziheng Lu[1,5], Weitao Feng[1,2], Fusong Ju[1], Jiaxi Wang[1], Jianwei Zhu[1], Yaosen Min[1], He Zhang[1], Shidi Tang[1], Hongxia Hao[1], Peiran Jin[1], Chi Chen[3], Frank Noé[4], Haiguang Liu[1] & Tie-Yan Liu[1]

MIT News
ON CAMPUS AND AROUND THE WORLD

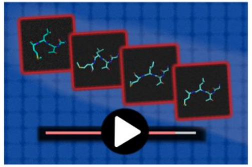
SUBSCRIBE | BROWSE | SEARCH NEWS

Toward video generative models of the molecular world

Starting with a single frame in a simulation, a new system uses generative AI to emulate the dynamics of molecules, connecting static molecular structures and developing blurry pictures into videos.

Alex Shipps | MIT CSAIL
January 23, 2025

PRESS INQUIRIES



By hitting the "play button" on molecules, MDGen could potentially help chemists design new molecules and closely study how well their drug prototypes for cancer and other diseases would interact with the molecular structure it intends to impact.
Image: Alex Shipps/MIT CSAIL

As the capabilities of generative AI models have grown, you've probably seen how they can transform simple text prompts into hyperrealistic images and even extended video clips.

SHARE

nature machine intelligence

Article https://doi.org/10.1038/s42256-024-00792-z

State-specific protein–ligand complex structure prediction with a multiscale deep generative model

Received: 2 June 2023
Accepted: 9 January 2024
Published online: 12 February 2024
Check for updates

Zhuoran Qiao[1,2], Weili Nie[3], Arash Vahdat[3], Thomas F. Miller III[2] & Animashree Anandkumar[3,4]

The binding complexes formed by proteins and small molecule ligands are ubiquitous and critical to life. Despite recent advancements in protein structure prediction, existing algorithms are so far unable to systematically

KI NEWS24

Home | AI Agents | Business | Science | Tools | KINEWS24

Microsoft BioEmu-1: Proteine sind die Grundpfeiler des Lebens – sie bilden Muskeln, schützen vor Krankheiten und steuern nahezu alle biologischen Prozesse. Doch ihre Strukturen und dynamischen Veränderungen zu entschlüsseln, bleibt eine der größten Herausforderungen der modernen Wissenschaft. Traditionelle Methoden wie Molekulardynamik-Simulationen (MD) sind zwar leistungsfähig, aber extrem zeitaufwendig und rechenintensiv – oft dauert es Jahre, um die Struktur eines einzigen Proteins zu simulieren. Genau hier setzt BioEmu-1 an: ein bahnbrechendes Deep-Learning-Modell, das die Proteinstrukturforschung revolutioniert. Entwickelt von Microsoft Research in Zusammenarbeit mit führenden Wissenschaftlern, kann BioEmu-1 Tausende von Proteinstrukturen pro Stunde generieren, und das mit einer Genauigkeit, die experimentelle Methoden herausfordert.

AI + CHE Priorities (preliminary)

- **Protein and Materials Design:** Develop models that can design proteins, small molecules, and materials with desired functions that can be tuned for different environments (temperature, pH etc).
- **Simulations:** Leverage AI to accelerate computational chemistry simulations, improve the experimental accuracy of condensed phase simulations, advance enzyme designs, and create self-driving labs.
- **Research Workflow Enhancement:** Rethink the workflows and structure of chemistry research itself, including emphasizing open-ended hypothesis generation and transitioning to an idea-limited, versus resource-limited, mindset.
- **Chemistry-aware ML:** Train foundation models on higher-level "languages" and incorporate chemical and physical knowledge into the models. Develop a foundation model that integrates universal interatomic potentials, generative sampling, and property predictions.
- **AI Standardization:** Incorporate AI into standard procedures for chemical synthesis and predictive chemistry.
- **Data and Publication Infrastructure:** Work together as a community to build an infrastructure for obtaining large and high-quality simulation data from wave function theory and for standardizing the publication of AI methods.

AI + CHE Questions to Consider (1)

- **Data Scarcity:** How do we train AI when chemical data is limited or biased?
- **Dynamics and environment :** Chemistry is not just “what (i.e. one structure) ” but “what all (i.e. ensemble)”, “when” and “under what conditions”. How do we predict these?
- **Physical Validity:** How do we ensure AI-generated molecules and reactions obey real-world chemistry?
- **Generalization & Extrapolation:** Can AI predict *new* chemical matter and behavior beyond training data?

Driving Question: How can the MPS domains best capitalize on, and contribute to, the future of AI?

AI + CHE Questions to Consider (2)

- What are flagship examples of AI+CHE that will help demonstrate where the field is now?
- Are the following subdomains the best way to categorize Chemistry?
  - Drug Discovery
  - Catalysis and Reaction Engineering
  - Materials Chemistry
  - Quantum Chemistry and Molecular Simulations
  - Cheminformatics & Structural Analysis
- What are the most important priorities and opportunities in AI+CHE in the next 5 years?
- How do we train new chemists to be competitive in the AI driven economy?

Driving Question: How can the MPS domains best capitalize on, and contribute to, the future of AI?

Domain Overview: Materials Research

10 minutes presentation + 5 minutes Q&A

Andrew Ferguson, University of Chicago

AI + Materials Research (DMR)

- AI presents powerful tools and concepts to **enable new paradigms in materials and molecular modeling**, provide **deeper understanding of structure-processing-property relations**, and **accelerate molecular and materials discovery**.

- AI is used across DMR in a number of applications, including:
    - Prediction, characterization, and modeling of molecule interactions,
    - Acceleration of existing workflows
    - Hypothesis generation
    - Advanced characterization
    - Multi-modal data analysis
    - More efficient explorations over high dimensional spaces using surrogate models
    - Uncertainty modeling

- Discovering **optimal combinations of composition and processing** for new polymeric formulations

- Powering **self-driving laboratories** integrating AI with automated robotics

- **Incorporating strong physical priors** into learning models that are critical for physical systems

These highlights are summarized from responses to the survey–more to be discussed in the breakout groups!

AI in Materials Research Subdomains (Examples)

- **Materials chemistry:** Machine learning interatomic potentials (ML-IAPs) have contributed to significant advancements in modeling, including enhanced model accuracy, robustness, development automation, capabilities for uncertainty quantification, and transferability.
- **Small molecules:** Huge strides have been made to use AI to predict properties (given dataset prerequisites) generate novel/valid molecules, and identify synthetic pathways.
- **Molecule and materials discovery:** These discoveries have been enabled by leveraging faster, low fidelity predictions to accompany medium fidelity simulations and high fidelity experiments.
- **Molecule and materials design:** Researchers are developing ML techniques with a focus on energy harvesting and storing and electronics applications
- **Quantum, photonic, and optical materials:** AI is revolutionizing the design, analysis, and optimization of complex optical systems, which has helped advance the development of engineered photonic materials, such as metasurfaces, photonic crystals, and novel metamaterials.

Andrew Ferguson, UChicago (DMR + CHE)

AI/ML for dimensionality reduction, enhanced sampling methods, collective variable discovery, active learning, rare event dynamics, active learning materials and molecular discovery, deep generative protein design

Key Issue for Your Domain

- General strategies for incorporating hard constraints or informative physical priors into deep generative models
- Sustainability and energy cost of model training, efficient learning strategies (e.g., adapters, LoRA)

Domain-Specific Case Study/Example

- Multi-modal learning of a joint latent space of protein annotations and sequences to enable a “ChatGPT” for proteins with experimental wet lab validations
  “Natural Language Prompts Guide the Design of Novel Functional Protein Sequences”
  https://doi.org/10.1101/2024.11.11.622734
- Democratization of deep generative protein design, compositionality of different known facets protein function, scope for supernatural function via active learning?

Exemplar: Lila Sciences

FIERCE Biotech

Biotech | Research | Medtech | CRO | Special Reports | Trending Topics | Fierce 50

BIOTECH

With $200M in seed funding, Flagship-backed Lila Sciences touts ambitious AI vision

By Darren Incorvaia · Mar 10, 2025 1:47pm

Biotech incubator Flagship Pioneering has uncorked its latest company. Lila Sciences is looking to use $200 million in seed funding to develop new advanced artificial intelligence that can power fully autonomous research labs, according to a March 10 press release.

In addition to Flagship, the financing comes from General Catalyst, March Capital, the ARK Venture Fund, Altitude Life Science Ventures, Blue Horizon Advisors, the State of Michigan Retirement System, Modi Ventures and a wholly owned subsidiary of the Abu Dhabi Investment Authority, according to the release.

Lila's mission is to achieve "scientific superintelligence" that is able to help scientists generate ideas and hypotheses and then design and conduct experiments to test those hypotheses, CEO Geoffrey von Maltzahn, Ph.D., said in the release.

"To achieve this, we must solve the hard problems to allow AI to autonomously and in a scalable manner run each step—from AI models generating an idea to reducing it to practice with robotics and automation," von Maltzahn said.

Since Lila's 2023 founding, Flagship claims the company's platform has already achieved groundbreaking results in various scientific areas. This includes large language models with state-of-the-art scientific reasoning abilities, the generation of genetic medicine constructs that perform better than commercially available therapeutics, and the discovery and validation of hundreds of new antibodies, peptides and binders for a broad range of therapeutic targets, according to the release.

Lila's chief executive, Geoffrey von Maltzahn, left, with the founder and chief executive of Flagship Pioneering, Noubar Afeyan.

https://www.fiercebiotech.com/biotech/new-flagship-backed-ai-firm-aims-build-scientific-superintelligence
https://www.nytimes.com/2025/03/10/technology/ai-science-lab-lila.html

Exemplar: Lila Sciences

The Lila team has completed five projects to demonstrate the abilities of its A.I., a powerful version of one of a growing number of sophisticated assistants known as agents. In each case, scientists — who typically had no specialty in the subject matter — typed in a request for what they wanted the A.I. program to accomplish. After refining the request, the scientists, working with A.I. as a partner, ran experiments and tested the results — again and again, steadily homing in on the desired target.

One of those projects found a new catalyst for green hydrogen production, which involves using electricity to split water into hydrogen and oxygen. The A.I. was instructed that the catalyst had to be abundant or easy to produce, unlike iridium, the current commercial standard. With A.I.'s help, the two scientists found a novel catalyst in four months — a process that more typically might take years.

That success helped persuade John Gregoire, a prominent researcher in new materials for clean energy, to leave the California Institute of Technology last year to join Lila as head of physical sciences research.

George Church, a Harvard geneticist known for his pioneering research in genome sequencing and DNA synthesis who has co-founded dozens of companies, also joined recently as Lila's chief scientist.

"More power to them, if they can do it," said David Baker, a biochemist and director of the Institute for Protein Design at the University of Washington. "It seems beyond anything I'm familiar with in scientific discovery."

https://www.fiercebiotech.com/biotech/new-flagship-backed-ai-firm-aims-build-scientific-superintelligence
https://www.nytimes.com/2025/03/10/technology/ai-science-lab-lila.html

AI + DMR Priorities (preliminary)

- **Database Development:** Develop better community-supported databases for soft materials, like those that exist for hard materials
- **Benchmarking and Standardization:** Define a suite of standard and challenging benchmarks to compare new algorithms in enhanced sampling and molecular and materials design applications
- **Robust AI:** Develop interpretable domain-specific AI models, physics-informed neural networks, and uncertainty quantification techniques to advance our understanding of how to manipulate chemical and material systems
- **Enhance Techniques and Lab Design:** Combine molecular simulation methods with deep learning (DL) to develop neural network based potentials, enhance sampling techniques, and facilitate lab design and optimization. Use AI techniques such as generative AI and reinforcement learning to efficiently navigate high-dimensional design spaces and achieve unprecedented performance metrics
- **Materials Discovery:** Use AI for inverse design, enabling the rapid discovery of metasurfaces, photonic crystals, and other nanostructures with tailored optical properties. Explore chemical space through generative AI models and examine structure-process-property relationships to help identify impacts of disorder/defects.
- **Materials Research Impact on AI:** Fostering a two-way relationship where materials challenges impact AI innovation

AI + DMR Questions to Consider

- What are flagship examples of AI+DMR that will help demonstrate where the field is now?
- What key subdomains are poised to make best use of AI in Materials Research?
    - Soft Materials – polymers, gels, rubbers, networks
    - Biomolecular Materials – proteins, peptides, nucleic acids
    - Hard Materials – catalysts, electronic and optical materials
    - Biomedical Materials – small molecule ligands and drugs, adjuvants, vaccines
    - Quantum Materials – substrates for qubits, quantum sensing and information
    - Automated Robotics & Self-Driving Labs
- What are the most important priorities and opportunities in AI+DMR in the next 5 years?
- What are the key challenges to address in AI+DMR?
- What does DMR have in common with the other MPS disciplines when it comes to AI?

Driving Question: How can the MPS domains best capitalize on, and contribute to, the future of AI?

Domain Overview: Mathematical Sciences

10 minutes presentation + 5 minutes Q&A

Soledad Villar, Johns Hopkins University

AI + Mathematical Sciences (DMS)

MATH for AI

- **Theoretical understanding of AI:** Math can be used to understand the properties and behavior of existing AI models (training optimization, generalization, finetuning, scaling laws, emergent behavior).
- **AI Innovation:** Math can inform the design of new machine learning models with useful properties, such as symmetries, conservation laws, fairness constraints, and multi-scale behavior. Other examples are continuous-time models like neural ODEs or diffusion models and invariant and equivariant architectures that capture known physics. This is specially important for AI for science.

AI for MATH

- **Proof Generation/Validation**: AI tools could potentially be used as “theorem prover copilots,” where the mathematician has a theorem or proof in mind, but has not worked out the details and asks the AI for assistance on modules. It can also helps formalize and discover mathematical proofs.
- **Computational Mathematics:** AI can help tackle previously intractable problems such as high-dimensional differential equations, statistics, control theory, and surrogate modeling.

These highlights are summarized from responses to the survey–more to be discussed in the breakout groups!

AI in Mathematical Sciences Subdomains (Examples)

- **Mathematical Foundations:** Models like DeepMind's AlphaProof and AlphaGeometry are promising, as well as theorem prover co-pilots like Lean and SAT solvers.
- **Combinatorics:** Transformer models have been used for finding geometric combinatorial objects, and reinforcement learning is being used to discover counterexamples in combinatorics.
- **Applied Mathematics:** Neural networks are being used for solving PDEs. Generative models have been used for imaging and inverse problems.
- **New mathematics inspired by AI problems.** Mathematical problems inspired by AI produce new mathematics in different areas. For example high dimensional probability, invariant theory, optimization.
- We'll hear more examples from you during our breakout session

Soledad Villar, Johns Hopkins (DMS, computational math, modl)

Equivariant machine learning, graph neural networks, mathematical theory of deep learning

Key Issue for Your Domain

- Mathematical research in AI:
  - Explain the behavior of AI models (e.g. training dynamics, generalization properties of trained models, sample complexity, etc).
  - Use mathematical principles to design machine learning models (e.g. group equivariant models implemented via representation theory or invariant theory).
  - Use of AI to prove theorems, verify proofs, or make conjectures.

Domain-Specific Case Study/Example

- My work uses techniques from algebra (invariant theory, galois theory, representation theory) to design machine learning models that are (approximately) invariant/equivariant with respect to group actions.
- Examples: machine learning on point clouds (applications to cosmology/computer vision), graph neural networks, equivariant self-supervised learning.
- Also work in mathematical theory of deep learning (eg. generalization guarantees).

AI + DMS Priorities (preliminary from questionnaire)

- **Mathematical theory for AI:** Mathematical theory to explain learning. Generalization, optimization, transfer learning, finetuning, scaling laws, emergent behavior, generative AI.
- **Mathematics Verification and Discovery:** Use symbolic AI for verifying and discovering mathematics, including generating proofs.
- **Problem Solving Agents:** Develop domain-specific AI agents that assist in solving problems across areas of formalization of mathematics, especially those that are beyond reach of current techniques. Example: RL for combinatorics.
- **Mathematical foundations of AI+Science:** Mathematical principles to design AI tools to be used in other domains such as physics, chemistry, and biology.
- **Robust and Trustworthy AI:** Build AI tools that emphasize safety, security, and privacy. Can we have reliable uncertainty quantification of AI models? Neurosymbolic AI was mentioned in the questionnaire as a possible direction.
- **Interpretable/explainable AI:** For some applications, like medical applications it is important to be able to explain the decisions that AI models make.

Examples of math-focused NSF programs

NSF Tripods (Transdisciplinary Research In Principles Of Data Science)

Artificial Intelligence, Formal Methods, and Mathematical Reasoning (AIMing)

Mathematical Foundations of AI (MFAI)

NSF-Simons Mathematical theory of deep learning (MoDL)

Driving Question: As researchers in the mathematical sciences, what programs do you want to see put forward by NSF MPS?

AI + DMS Questions to Consider

- What are flagship examples of AI+DMS that will help demonstrate where the field is now?
- What is the best way to categorize Mathematics Research and AI sub domains?
- What are the most important priorities and opportunities in AI+DMS in the next 5 years?
- What are the key challenges to address in AI+DMS?
- What does DMS have in common with the other MPS disciplines when it comes to AI?

Driving Question: How can the MPS domains best capitalize on, and contribute to, the future of AI?

Domain Overview: Physics

10 minutes presentation + 5 minutes Q&A

Jesse Thaler, MIT

AI + Physics (PHY)

- **Theoretical Calculations:** AI is being used to accelerate quantum field theory calculations and discover new structures in mathematical data
- **Data Analysis:** Deep learning has enabled high-dimensional statistical analyses, enabling the use of much more information than ever before without having to resort to summary statistics. Increasingly, AI is being used as a primary analysis tool, including to search for anomalous features in datasets.
- **Noise Mitigation:** AI has been used to improve the detection and removal of noise artifacts in data.
- **Simulations:** Across physics, AI has been used to revolutionize the use (and reuse) of simulations.
- **Experimental Design and Optimization**: Neural network-based computational methods are impacting computation design, including protocol design and sampling. AI is also helping to advance multiparameter optimization and has provided essential control functions for experimental protocols.
- **Robust and Interpretable AI:** Physicists are developing methods to understand and improve AI, including the necessary precision for scientific applications.

These highlights are summarized from responses to the survey–more to be discussed in the breakout groups!

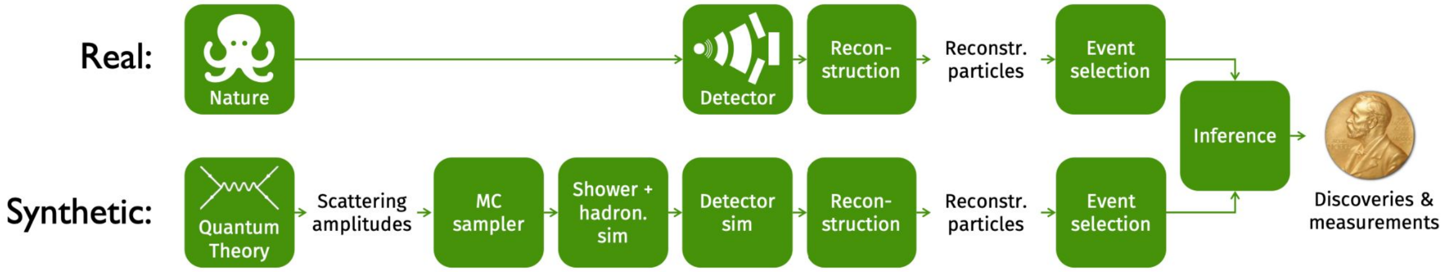
AI in Physics Subdomains (examples)

- **Quantum Physics:** Advanced multiparameter optimization achieved state-of-the-art decoding performance in quantum error correction experiments, which exemplifies how AI can enhance the reliability and effectiveness of quantum computing systems.
- **Atomic, Molecular, and Optical Physics:** Reinforcement learning (RL) has been used to design quantum control protocols for atom interferometry. Unlike optimal control, the training can easily be performed model free and in the lab using experimental observation and feedback to determine the reward values.
- **Particle Physics**: Experimental collaborations have used AI to help process and analyze huge datasets, which allow researchers to accelerate simulations, build powerful architectures for event and jet classification, conduct anomaly detection for new physics searches, and make more sensitive and precise measurements.
- **Gravitational Wave Physics**: AI has been used to remove noise from time-series data and detect signals faster and at lower costs. A promising approach to measuring the stochastic gravitational wave background is simulation-based Inference, which allows for performing inference without an analytical likelihood.
- **Biophysics:** AI is becoming central to organizing and analyzing large datasets and has become a lens through which to view learning in living matter.

Personal Perspective on AI + (Particle) Physics

The Standard Story:

- Shallow machine learning and low-dimensional statistics have a long history in particle physics
- Now, we are revisiting our field through lens of **deep learning and high-dimensional statistics**
- Key paradigm of **simulation-based inference** enabled by vast data sets and trustable simulations



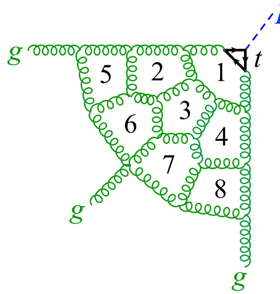
+ AI

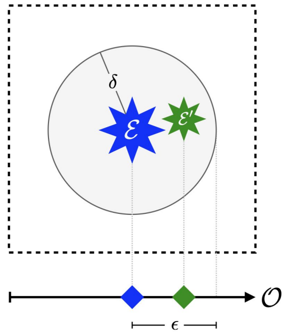
This standard story is true! SBI and its extensions are driving immense progress, and there is much more to come! But is that why I keep working on AI?

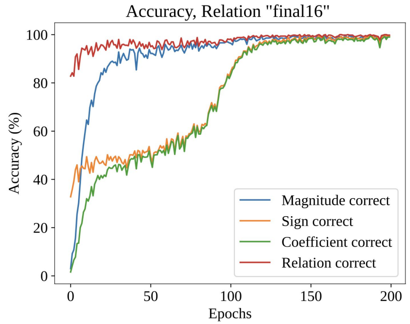
Personal Perspective on AI + (Particle) Physics

Beyond the Standard Story:

- AI/ML (and mathematics, statistics, computation more generally) allows a **reframing of physics-specific questions** in a more universal language
  - E.g. Question (1970s): What calculations are "allowed" in perturbative quantum field theory?
  - Answer (2020s via 1780s): Optimal transport!
- AI/ML allows a shift from solving problems algorithmically to **specifying problems as optimization/search tasks**
  - E.g. "Solving" amplitude bootstrap with transformers





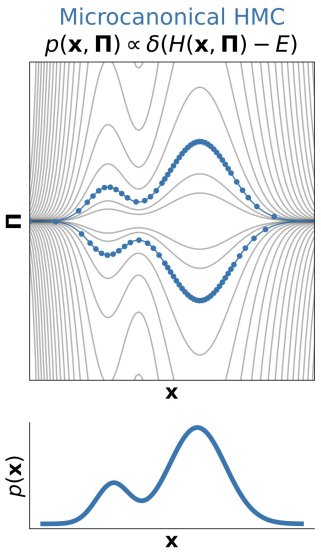


An invitation to dream big! What if AI doesn't just give an approximation to rigorous results but yields something even better?

Towards the "Physics of AI"

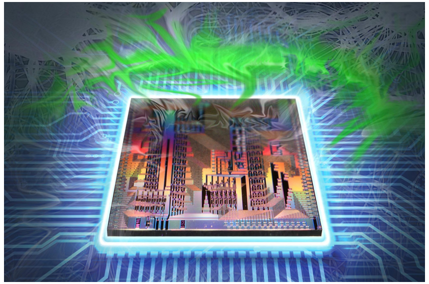
The Standard Story:

- Progress in computation and information theory has long been intertwined with progress in the mathematical and physical sciences
  - E.g. statistical mechanics, quantum computers
- **Physical systems are learning systems** (2024 Nobel Prize!)
  - E.g. diffusion and classical mechanics as generative models
- **Neural networks can be analyzed using physics tools**, since they share many similarities with quantum many-body systems and quantum field theories



Beyond the Standard Story:

- What if the next AI advance doesn't come from mimicry of a physical systems (or the brain) in silico, but from a new kind of **physical learning device**?
  - Cf. quantum mechanics → transistor (well before quantum computers!)



AI + PHY Priorities (preliminary)

- **Theoretical Frameworks for AI Use:** Establish theoretical frameworks to characterize how and when AI will accelerate discovery, including understanding the limits of AI and how its predictions can be trusted, especially using uncertainty quantification.
- **Experiment Development, Design, and Automation:** Integrate AI into experiment development and design, including developing reinforcement learning techniques for operations. Automation of experimental control could have a large impact in efficiency/uptime of operations, especially for advanced and next generation systems.
- **Interpretable AI:** Build models that are multi-use and interpretable, which will allow researchers to have an intuitive understanding of how an AI tool is manipulating data and give assurance of the mathematical rigor used by the tool.
- **Data Analysis:** Combine the ability of AI to process vast datasets with the human ability to scrutinize puzzling features in physical data.
- **Science of AI:** Develop a physics-based theory of complex AI systems and their emergent properties toward a multidisciplinary “Science of AI” effort.

AI + PHY Questions to Consider

- What are **flagship examples** of AI+PHY that will help demonstrate where the field is now?
- The **core research areas** in NSF PHY are as follows. Is this the right organization for the AI white paper? How to ensure coverage across all of these areas?
    - Atomic, Molecular and Optical Physics
    - Elementary Particle Physics
    - Gravitational Physics
    - Nuclear Physics
    - Particle Astrophysics and Cosmology
    - Physics of Living Systems
    - Plasma Physics
    - Quantum Information Science
- What are the most important **priorities and opportunities** in AI+PHY in the next 5 years?
- What are the **key challenges** to address in AI+PHY?
    - Do these answers differ between **large-scale experiments**, **small-scale experiments**, and **theory**?
- What does PHY have in common with the **other MPS disciplines** when it comes to AI?

Driving Question: How can the MPS domains best capitalize on, and contribute to, the future of AI?

Coming Up Next…

Today’s Schedule (Monday)

9:00–9:30 am: Welcome and Overview

9:30–10:30 am: Theme Overviews

- Interdisciplinary Research: Opportunities and Challenges: Lars Ruthotto
- Interdisciplinary Research: Resources Needed: Andrew Ferguson
- Education & Workforce Development: Yuan-Sen Ting
- Responsible AI: Pratyush Tiwary

10:30–11:00 am: Break

11:00 am–12:30 pm: Domain Overviews

- AST: Yuan-Sen Ting
- CHE: Pratyush Tiwary
- DMR: Andrew Ferguson
- DMS: Soledad Villar
- PHY: Jesse Thaler

12:30–2:00 pm: Lunch

2:00–5:30 pm: Theme Breakouts

- Interdisciplinary Research: Opportunities and Challenges: Room 801 North (here)
- Interdisciplinary Research: Resources Needed: Room 801 South
- Education & Workforce Development: Room 804
- Responsible AI: Room 812