AI/ML Applications in Astronomy & Astrophysics: IUCAA 2025

AI Aspirations:
Towards Better Sub-grid Models for Cosmic Structure Formation Simulations

Arif Babul
Distinguished Professor, Univ. of Victoria
Infosys Visiting Chair Prof., Indian Inst. of Science
Leverhulme Visiting Prof., Univ. of Edinburgh

1. My initiation into Scientific AI: Concussion research in Neuroscience.

Concussion Neuroscience: A micro-introduction

10^{11} neurons in the brain

Cosmic Structure
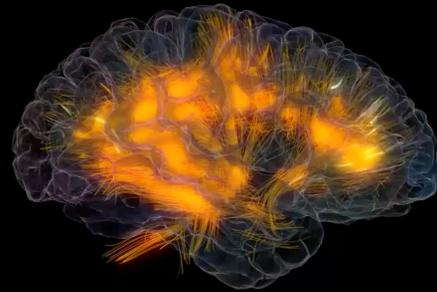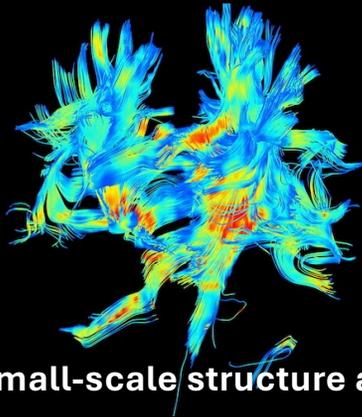
Neural Structure
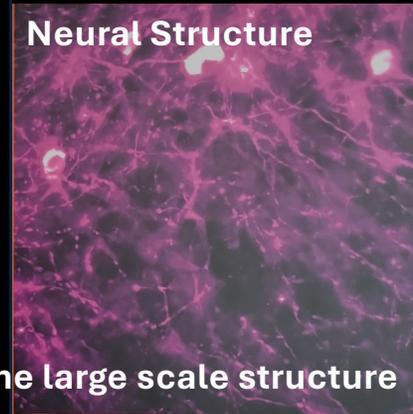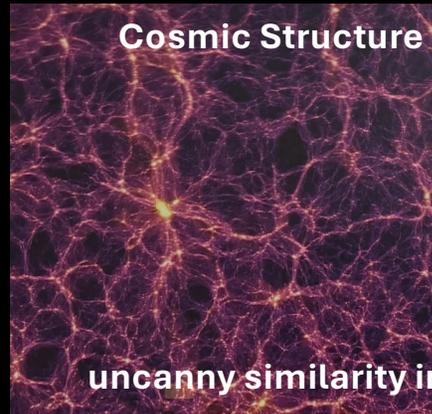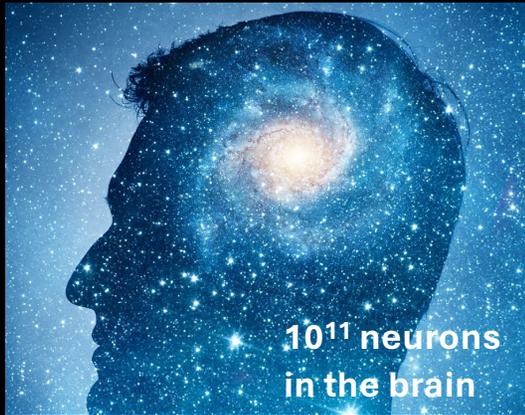
uncanny similarity in the large scale structure
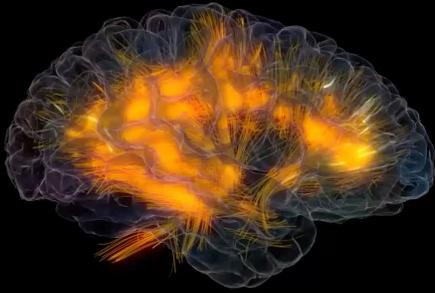
Main collaborators:

N. Virji-Babul
D. Hristopulos
K. Thanjuvar

Small-scale structure and dynamics is more complicated

EEG sensor net

Concussion Neuroscience: A micro-introduction

Power series: one channel
Healthy
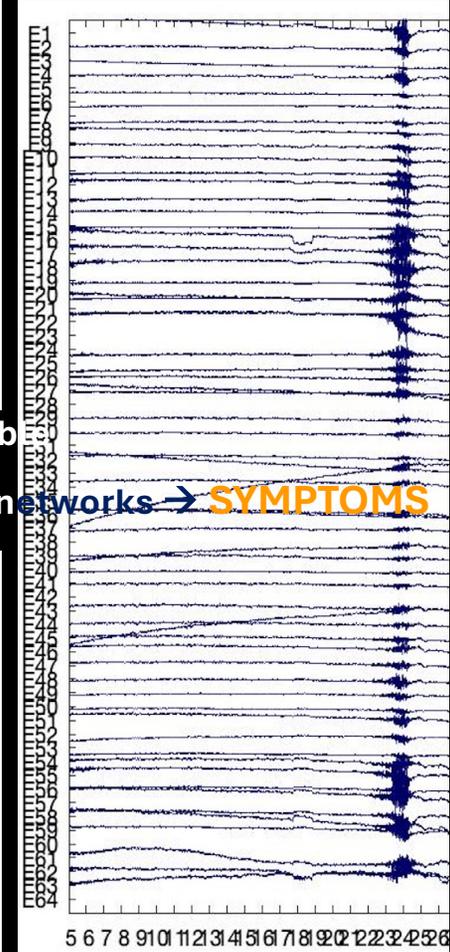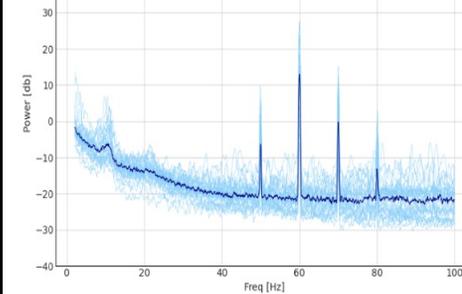
N time series data. Each time series is non-stationary, artifact contaminated, highly variable

INJURY → microtears in white matter fibres (brain wiring); disruptions in communication networks → **SYMPTOMS**
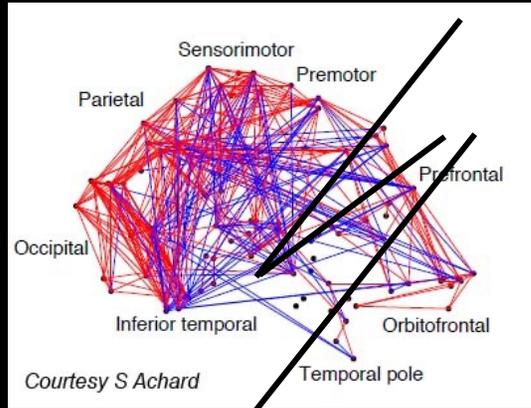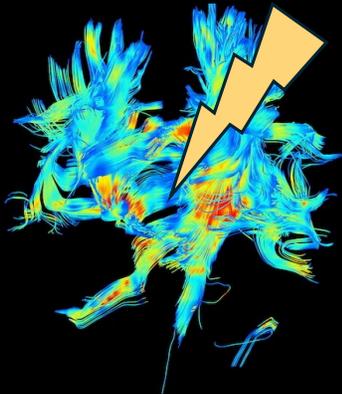
Power series: one channel
Concussed

Courtesy S Achard
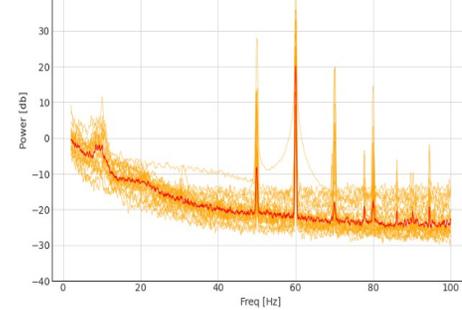
There are population differences but significant overlap. No way to identifiy a concussed person

Concussion Neuroscience: A micro-introduction

Current treatment is highly subjective and not consistent. This can be a serious problem leading to all sort of complications!

Started exploring AI/ML solution in 2014

SVM: 60% success. Discovered <u>FEATURE BIAS & FEATURE ENG.</u>
CNN: ADVANTAGE: POTENTIALLY KNOWABLE – marginal improvement (75%)
LSTM: 95% accuracy but irritatingly, a BLACK BOX

64 Channel EEG Sequential Input
Bidirectional LSTM - 1
Dropout-1
Bidirectional LSTM - 2
Dropout-2
FC-8 + RELU
FC-2
Output Layer
Concussed (1) - Control (0)
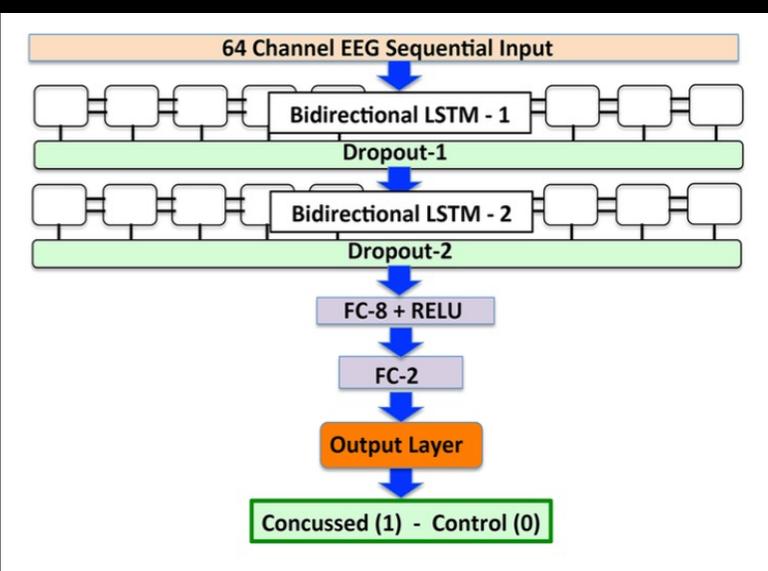
This is leading to several new directions of research:

- Interpretable?; extracting information leading to insight
- Combining datasets from different EEG devices
- Mapping networks for information flow in the brain
- Forecasting changes in the brain's structure and function in response to rehab treatments, disease, etc.

2. My ^{main} research focus is: cosmic structure evolution (galaxies, groups, smbhs)

GALAXIES SIT AT THE CONVERGENCE OF COSMOLOGY AND ASTROPHYSICS

Stellar Dynamics

ISM and Star Formation

Stellar Feedback

Stellar Evolution

Black Holes

Chemical Enrichment

Gas/hydro Dynamics
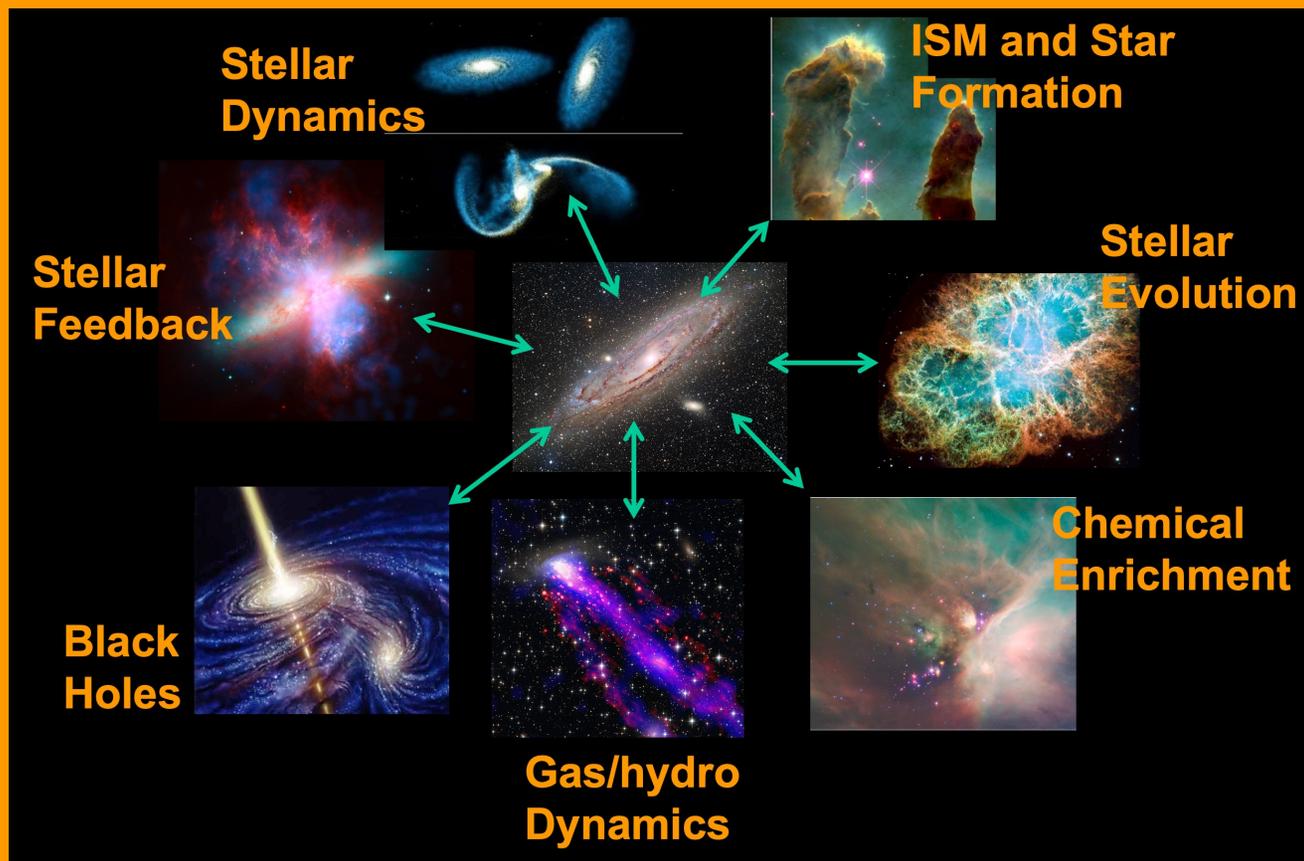
- Galaxies are essentially the building blocks.

- They are shaped by internal conditions and external ones.

- And they also re-engineer their surroundings.

SUB-GRID MODELS

Romulus Simulation
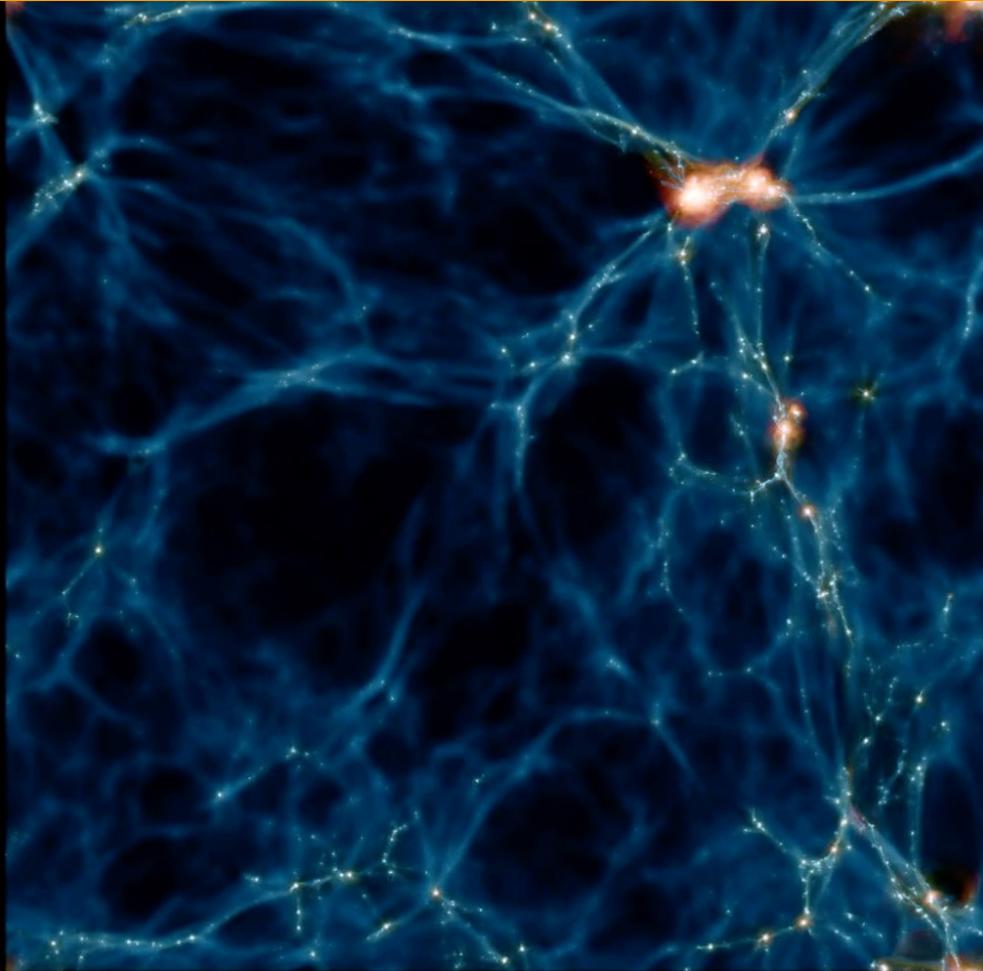
- Complex networks of physical processes, with non-linear interactions, cutting across nearly all of physics → numerical cosmological hydro sims

- Multi-component, multi-scale problem with physics spanning >15 order of magnitudes in spatial scales. The best we can do is ~10^7 and with limitations, ~10^9.

- Need sub-grid models to treat/capture processes that cannot be directly modelled.

WHY IS THIS HARD?

- **Most sub-grid processes are themselves complex (e.g. turbulence) and some we just plain don't understand well (jet/wind launch and initial evolution).**

- **Presently, these processes are modelled using simple "spherical cow" schemes and *ad hoc* models.**

- **In detail, the results can vary dramatically**

SIMBA

$z = 9.94$

Astrid

Magneticum
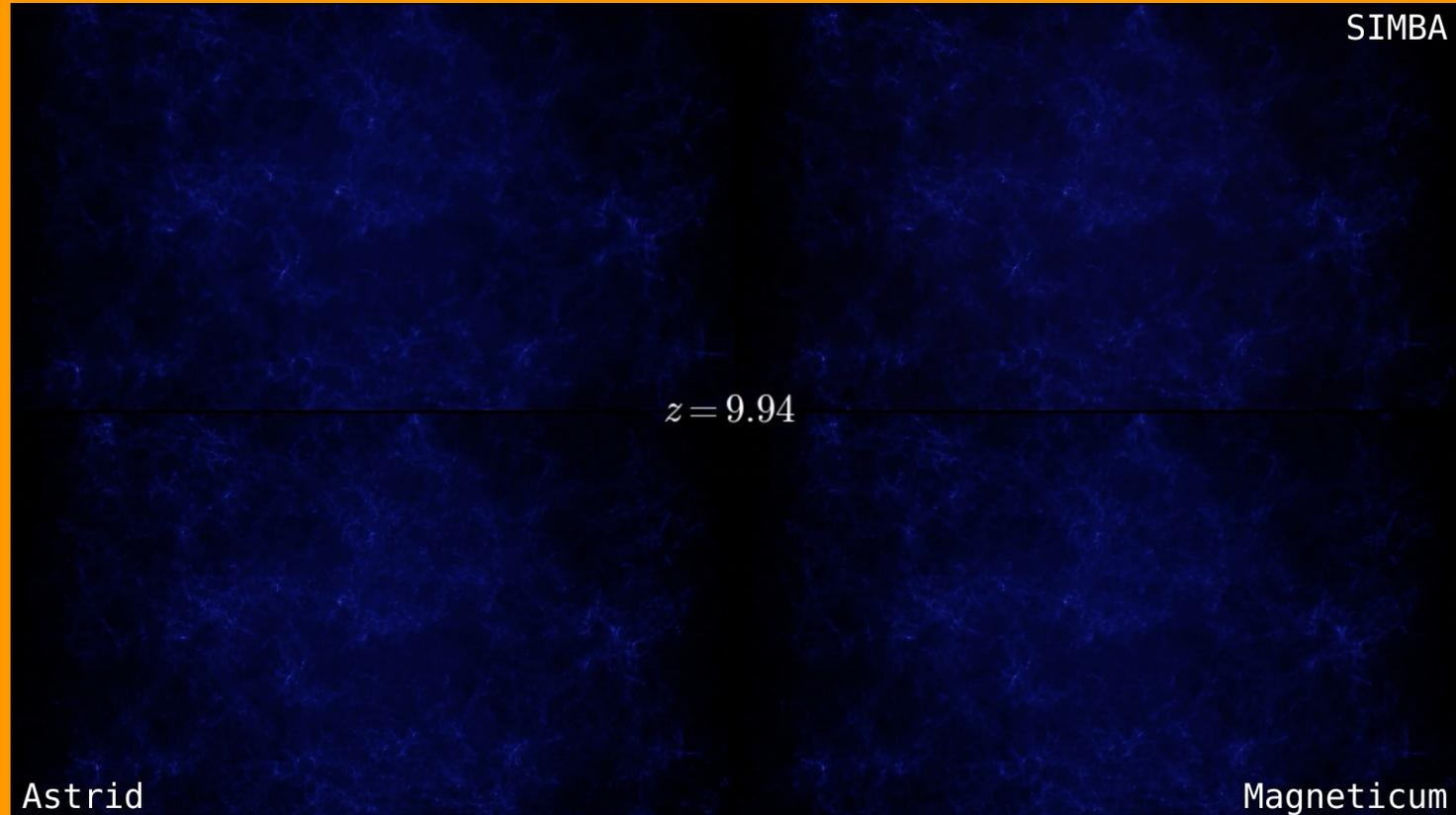
Temperature slices (CAMELS)
Movie by F. Villaescusa-Navarro

AI OFFERS A POTENTIAL WAY TO IMPROVE THE SITUATION

Two of the problems that I am especially interested in:

- **SMBH merging, feedback, …**
  [Suvodip Mukherjee, Lucio Mayer, Doug Rennehan, Romeel Dave…]

- **CGM / IGrM / ICM Turbulence (mixing, multiphase structure…)**
  [Prateek Sharma, Lucio Mayer, Vikram Khaire, Doug Rennehan…]

SMBH MERGING → GRAVITATIONAL WAVES + EM SIGNATURES

Shining light on the hosts of the nano-Hertz gravitational wave sources: a theoretical perspective

Vida Saeedzadeh [1], Suvodip Mukherjee [2]★ Arif Babul [1,3] Michael Tremmel [4]
and Thomas R. Quinn[5]

[1]*Department of Physics and Astronomy, University of Victoria, 3800 Finnerty Road, Victoria, BC, V8P 1A1, Canada*
[2]*Department of Astronomy and Astrophysics, Tata Institute of Fundamental Research, 1, Homi Bhabha Road, Colaba, Mumbai 400005, India*
[3]*Infosys Visiting Chair Professor, Department of Physics, Indian Institute of Science, Bangalore 560012, India*
[4]*School of Physics, University College Cork, College Road, Cork T12 K8AF, Ireland*
[5]*Astronomy Department, University of Washington, Box 351580, Seattle, WA 98195-1580, USA*

Imprints of supermassive black hole evolution on the spectral and spatial anisotropy of nano-hertz stochastic gravitational-wave background

Mohit Raj Sah [1]★ Suvodip Mukherjee [1]★ Vida Saeedzadeh [2] Arif Babul,[2,3] Michael Tremmel [4]
and Thomas R. Quinn[5]

[1]*Department of Astronomy and Astrophysics, Tata Institute of Fundamental Research, Mumbai 400005, India*
[2]*Department of Physics and Astronomy, University of Victoria, 3800 Finnerty Road, Victoria, BC, V8P 1A1, Canada*
[3]*Infosys Visiting Chair Professor, Indian Institute of Science, Bangalore 560012, India*
[4]*School of Physics, University College Cork, College Road, Cork T12 K8AF, Ireland*
[5]*Astronomy Department, University of Washington, Box 351580, Seattle, WA 98195-1580, USA*

SMBH MERGING → GRAVITATIONAL WAVES + EM SIGNATURES

MICECAT Simulation (Gpc scale)
$H_0 \quad \Omega_m \quad \Omega_\Lambda \quad \sigma_8$

ROMULUS Simulation (Mpc scale)
$\dot{M}_* \quad \epsilon_{SN} \quad \epsilon_r \quad \epsilon_f$

BH binary modeling in gas dynamic (< pc scale)
$\beta \quad \kappa \quad \gamma$
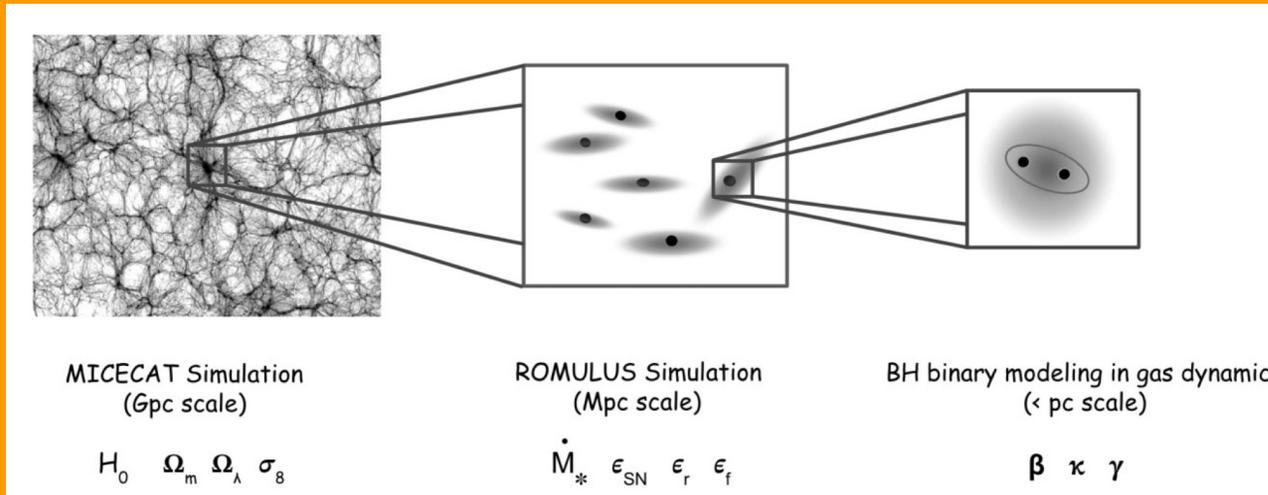
Multi-scale problem:

Running a single simulation that covers the full dynamic range is impossible.

SMBH MERGING → GRAVITATIONAL WAVES + EM SIGNATURES

Multi-scale problem

Our base hydro sim is Romulus → has "free-range" SMBHs"

Can track BH dynamics, including BBHs, to ~300 pc

Then, post-process via analytic sub-grid model to get nHz signals

The ROMULUS Simulations

CHANGA

Certified organic, free-range, locally grown supermassive black holes

✓ Early Seeding in low mass halos
✓ Self-consistent and physically motivated dynamics, growth, and feedback
✓ Shown to make realistic M < 1.E+12 Msun galaxies in detailed sense (zoom and cosmological boxes)
✓ Organic M > 1.E+12 Msun systems
✓ No hormones or additives
✓ Naturally produces large-scale outflows

ROMULUSC
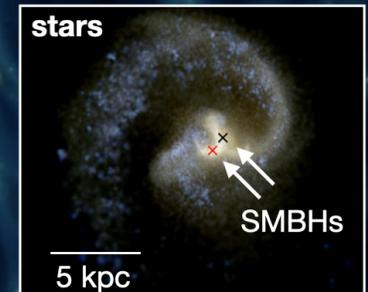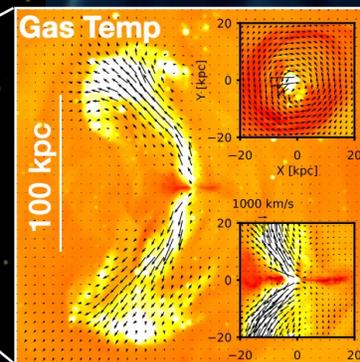10^14 Msun Galaxy Cluster
Tremmel+2019
Chandayammuri+2020

ROMULUSG1/G2
10^13.5-15 Msun
Jung+2022

ROMULUS25
25 Mpc Volume
Tremmel+ 2017
Butsky+2019; Jung+2022

Resolution:
250 pc (grav)
50 pc (hydro)
~1e5 Msun

stars

SMBHs

5 kpc

Gas Temp

100 kpc

Y [kpc]

X [kpc]

1000 km/s

SMBH MERGING → GRAVITATIONAL WAVES + EM SIGNATURES

- Merger dynamics is complicated: M_1 M_2 M_* dM_*/dt M_{gas} $L_{orbital}$ J_1 J_2
- Train an AI model to predict pre (BBH) and post BH dynamics and GW - and same model can even be trained to give EM → would require many many small-scale simulations [plus waveforms]
- But it is an [easier] problem – small scale is decoupled from large-scale → post-processing via AI model

Caveat: (1) probabilistic – diversity of differential distribution for given global parameter [eg. ρ_{gas} (r) for given M_{gas}]

(2) Are 1000s of sims to cover parameter space even possible, and if not, how can we overcome this obstacle?   AI to rescue?

TURBULENCE IN CGM / IGrM / ICM

In galaxies, groups and clusters, turbulence extends from "effectively incompressible subsonic" to "compressible supersonic"

8 Mpc box   Cluster gas density   4.02 Gyr

Poole et al 2006-2008
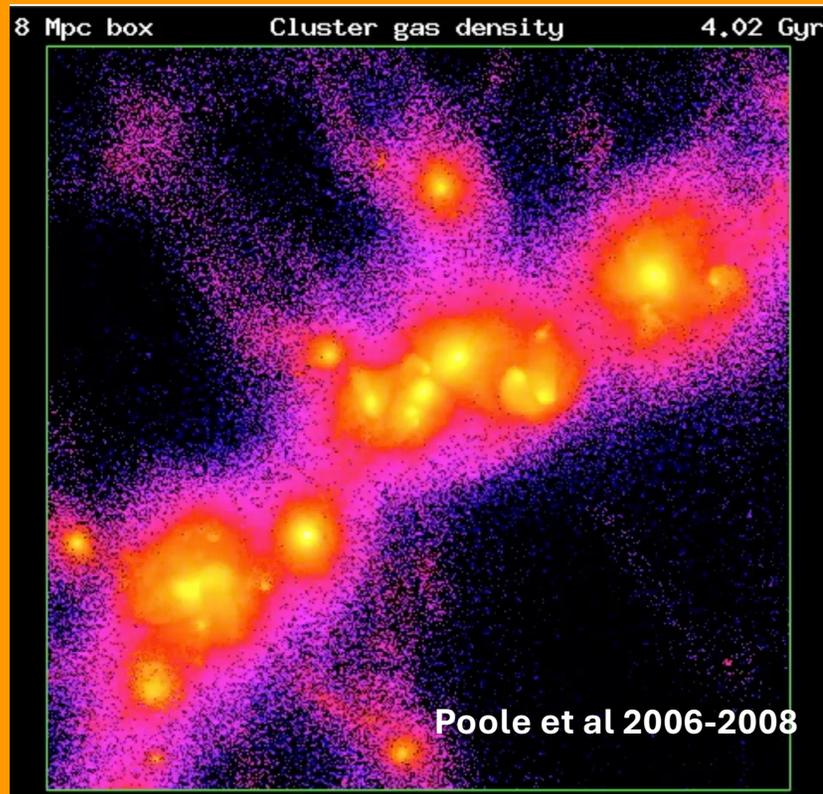
Prasad+2015, "Cool Core Cycles: Cold Gas and AGN Jet Feedback in Cluster Cores"
Prasad+2017, "AGN jets driven stochastic cold accretion in cluster cores"
Prasad+2018, "Cool-Core Clusters: Role of BCG, ... & AGN-driven turbulence"
Prasad+2020, "Cool-Core Cycles and Phoenix", MNRAS, 495, 594

Origin of multiphase CGM in idealised simulations:
AGN Jets-driven perturbations

Sharma et al. 2012;
McCourt et al. 2012

if $t_{cool}/t_{ff} > 10$, only hot phase;
if $t_{cool}/t_{ff} < 10$, multiphase

TURBULENCE IN CGM / IGrM / ICM

In galaxies, groups and clusters, turbulence extends from "effectively incompressible subsonic" to "compressible supersonic"
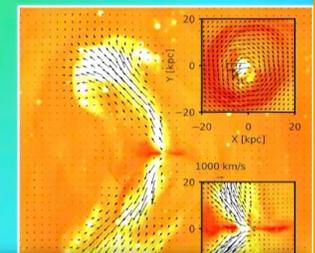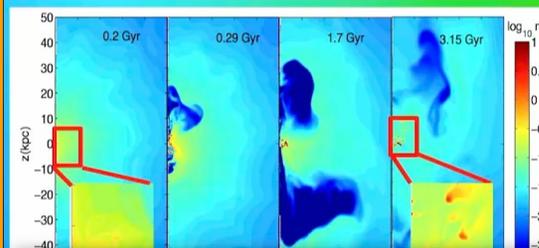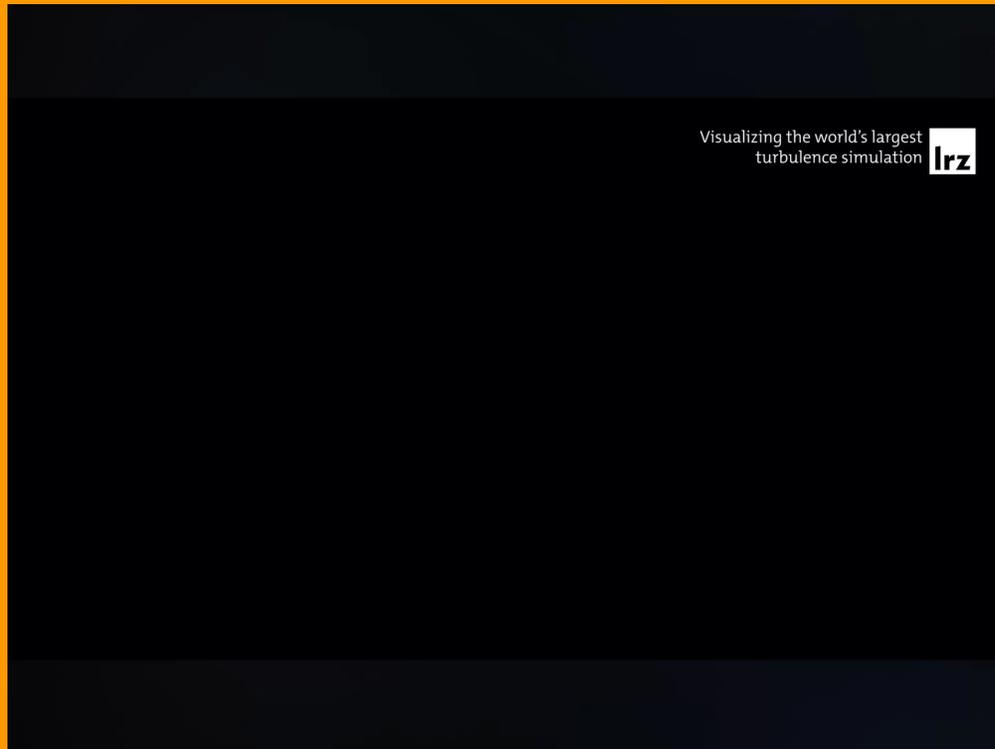
Subsonic and Supersonic
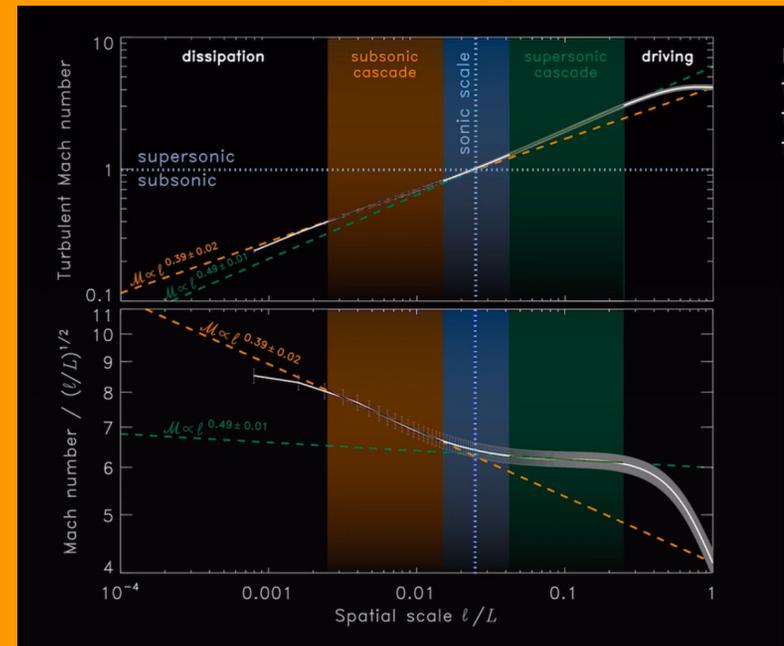
Salvatore Cielo, Luigi Iapichino, Johannes Günther, Christoph Federrath, Elisabeth Mayer, Markus Wiedemann (arXiv:1910.07850)

TURBULENCE IN CGM / IGrM / ICM

IMPORTANCE: mixing of metals,
                         energy and momentum re-distribution
                         density sub-structure → multiphase structure

In cosmological simulations, we have limited dynamic range

➢ **This breaks the physical coupling between scales that are resolved (> h) and unresolved (< h)**

➢ **They require sub-grid models of (i) the kinetic energy flux from the resolved to the unresolved scales and (ii) the effect of unresolved eddies on the resolved scales**

 **ii makes this a hard problem: needs small scale to be evolving consistently over time <u>and</u> at each t, small scale acts on large scale.

TURBULENCE IN CGM / IGrM / ICM

SPH has no inherent diffusion and therefore, by construction, explicitly requires additional transport terms

In moving-meshes (MM) and higher-order mesh-free methods (MF), numerical diffusion arises as a by-product of the numerical scheme used to solve the Riemann problem between fluid elements but this inherent numerical diffusion does not, in general, reproduce the correct turbulent properties of a simulated flow.

In most cosmological simulations – if this is addressed, it is done at all, via the eddy-viscosity hypothesis: The action of the unresolved eddies can then be modelled by a viscous term in the fluid equations of motion which diffuses momentum and dissipates kinetic energy.

KEY ASSUMPTION: SMALL SCALES DISSIPATE INSTANTANEOUSLY ALL THE ENERGY THEY RECEIVE FROM LARGE SCALES

TURBULENCE IN CGM / IGrM / ICM

Most common of this class of models: Smagorinsky Model where

$$\nu = D = (C_s\, h)^2\, |\text{S*}|$$

where |S*| is the trace-free shear tensor and C_s is the Smagorinsky constant.

DRAWBACKS: Too dissipative!

(i) A single valued Cs cannot correctly describe different types of turbulent flows. Studies show that the Smagorinsky model introduces too much diffusion into the flow in almost all cases except for homogeneous, isotropic turbulence.

(ii) the sub-grid eddy viscosity does not vanish for laminar shear flows

(iii) Treats metal and energy, not momentum.

TURBULENCE IN CGM / IGrM / ICM

FIX → DYNAMICAL SMAGORINSKY MODEL   (Rennehan et al. 2019)

Dynamic Localised Turbulent Diffusion and its Impact on the Galactic Ecosystem

Douglas Rennehan,[1][*] Arif Babul,[1] Philip F. Hopkins,[2] Romeel Davé[3,4,5] and Belaid Moa[6]

$C_s \rightarrow C_s(x,t)$

DRAWBACKS

Diffusive process assumed isotropic and scale-invariant but in reality, neither is true.

Anisotropic diffusion is possible to compute (see Rennehan 2021) but it is costly and is still limited by simplifying assumptions.

The fluid mechanics, ocean physics and climate science communities are moving ahead rapidly:

DNN emulator that can emulate sub-grid viscosity and diffusivity.

Deep learning emulation of subgrid scale processes in turbulent shear flows.

Anikesh Pal[1,2]

[1] National Center for Computational Sciences, Oak Ridge National Laboratory, TN USA
[2] Department of Mechanical Engineering, Indian Institute of Technology Kanpur, U.P. INDIA

Key Points:

- DNNs are built to emulate the SGS eddy viscosity and diffusivity for turbulent stratified shear flows.
- These DNNs compute the SGS eddy viscosity and diffusivity $2-4$ times faster than the Dynamic Smagorinsky Model.
- These DNNs emulate the SGS processes accurately such that the energy and variance budgets match with the Dynamic Smagorinsky Model.

Artificial neural-network-based subgrid-scale model for large-eddy simulation of isotropic turbulence

Lei Yang (杨磊),[1] Dong Li (李栋),[1,2,a] Kai Zhang (张凯),[1,2] Kun Luo (罗坤),[3] and Jianren Fan (樊建人)[3,a]

AFFILIATIONS

[1] Department of Mechanics, School of Aerospace Engineering, Beijing Institute of Technology, Beijing 100081, China
[2] Tangshan Research Institute, Beijing Institute of Technology, Tangshan 063099, China
[3] State Key Laboratory of Clean Energy Utilization, Zhejiang University, Hangzhou 310027, China

[a] Authors to whom correspondence should be addressed: bitld@bit.edu.cn and fanjr@zju.edu.cn

ABSTRACT

This study is concerned with accurately predicting the subgrid-scale (SGS) stress using an artificial neural network (ANN) with a linear eddy-viscosity term and a nonlinear term as the input variables. A priori and a posteriori tests are conducted to examine the prediction performance of the ANN-based SGS stress model in decaying homogeneous isotropic turbulence. In a priori test, the present ANN-based SGS model shows high correlation coefficients between the true and predicted SGS stresses, and excellent predictions of the SGS stress and dissipation. In a posteriori test, it is found that the ANN-based SGS model can predict the turbulence statistics more accurately than the traditional dynamic SGS models. The generalization capabilities of the model to untrained flow conditions and unstrained types of turbulent flow have been evaluated. It is found that the proposed ANN-based model can provide an accurate prediction of the SGS stress under different Reynolds numbers and flow types. A comparison among several existing ANN-based models with different input variables is presented, demonstrating a significant advantage of the present model.

Published under an exclusive license by AIP Publishing. https://doi.org/10.1063/5.0212096

Finds that DNN is able to extrapolate correctly to 10x higher Re than training set, is more accurate than traditional Dynamical Smag Model, and it is typically 2.5-3x faster than computing the Dynamical Smag.

TRAINING?
Direct Numerical Simulations, AI-evolved Turbulence? Quantum Computing ?

No doubt hybrid ML/AI + traditional simulations will become the norm

Embedding machine-learnt sub-grid variability improves climate model precipitation patterns
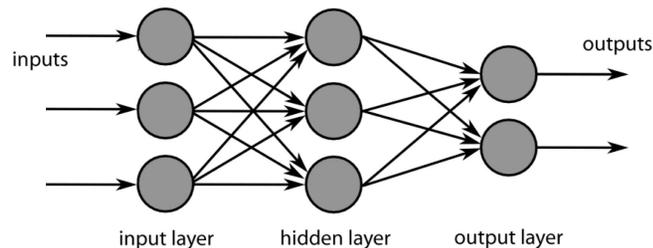
Check for updates

Daniel Giles [1] ✉, James Briant [2], Cyril J. Morcrette [3,4] & Serge Guillas[2]

OUTSOURCING SUB-GRID CLOUD PHYSICS TO NEURAL NETWORKS

August 18, 2020
Feature Story (https://e3sm.org/category/feature/)

inputs → input layer ... hidden layer ... output layer → outputs

(https://e3sm.org/wp-
content/uploads/2020/08/MultiLayerNeuralNetworkBigger_english_904x356.png)
Diagram of a two-layer feedforward neural network.

Deep Neural Networks: Powerful machine learning emulators of high-dimensional nonlinear functions disrupting industry and climate modeling

---

- **But sub-grid model forms are themselves simplifications.**
- **Still trying to understand whether DSM can give us density fluctuations forming naturally in simulations but …**

FINAL THOUGHTS:

…but even if it does, the spectrum of fluctuations and the nature of multi-phase structure depends on resolution of the simulation.

HOW TO OVERCOME?