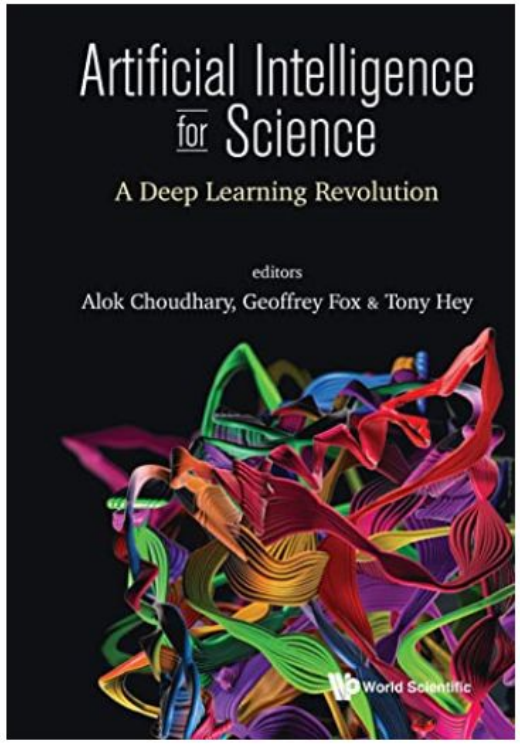
Generative AI for Digital Twins for Physics and Biology

Geoffrey Fox with Tien Comlekoglu, James Glazier,
and Javier Toledo

Kindle Store › Kindle eBooks › Computers & Technology



Artificial Intelligence for Science: A Deep Learning Revolution

Kindle Edition

by Alok Choudhary (Editor), Geoffrey Fox (Editor), Tony Hey (Editor) | Format: Kindle Edition

5.0 ★★★★★ 1 rating

See all formats and editions

Kindle
$37.49

Read with Our **Free App**

Hardcover
$148.00

2 Used from $158.12
8 New from $144.00

This unique collection introduces AI, Machine Learning (ML), and deep neural network technologies leading to scientific discovery from the datasets generated both by supercomputer simulation and by modern experimental facilities.

Huge quantities of experimental data come from many sources — telescopes, satellites, gene sequencers, accelerators, and electron microscopes, including international facilities such as the Large Hadron Collider (LHC) at CERN in Geneva and the ITER Tokamak in France. These sources generate many petabytes moving to exabytes of data per year. Extracting scientific insights from these data is

Read more

Read sample

Follow the Author

Alok N. Choudhary — Follow

ISBN-13	Sticky notes	Publisher	Publication date	Language	File size
978-9811265662	On Kindle Scribe	World Scientific Publishing Company	March 21, 2023	English	78220 KB

https://doi.org/10.1142/13123

40 Chapters across all areas in science and engineering dominated by deep learning

Veni, Vidi, Vici Regge Theory

Regge theory has, in the past few years, been compared with essentially all available data on two-body scattering at high energies. Indeed it is the only formalism (theory?) that is sufficiently flexible for such comprehensive quantitative tests to be feasible. Experimental data exist for so many reactions, and are now of sufficient precision, that the large number of parameters, inherent in Regge-pole theory, is no longer a hindrance in judging the scope of its validity. Let us define three possible models. Firstly we have the simple Regge-pole model in which no cuts are allowed. Secondly we consider the (as yet theoretically unjustified) absorption prescription for generating Regge-cut corrections to simple Regge poles. Finally we have the general framework of Regge poles (defined, say, by the particles created along the trajectory) plus arbitrary Regge cuts, restricted only by known general principles. We now state without proof three, almost everywhere valid, theorems.

(i) *The (cut) corrections to Regge-pole theory are at least as large as predicted by the absorption model.*

We point out that an absorptive-cut correction to, say an elastic amplitude, is some 20% of the pole at $t = 0$, while the cut becomes equal to the pole somewhere between $-t = 0.5$ and 1.0 (GeV/c)^2. If this is to be judged an important discrepancy, one can conclude that simple Regge-pole theory is insufficient.

(ii) *The predictions of the absorptive-cut model are generally incorrect.*

Although this model has had some interesting qualitative successes, most of these are shared by rather general models. In particular it is clear that Regge-pole predictions are generally *un*reliable for *low* (direct channel) partial waves. In fact, experimentally, these low partial waves are typically smaller than their Regge-pole values.

(iii) *The present fund of knowledge on the general properties of cuts is insufficient for meaningful phenomenology.*

Given the failure of the pole model and the inadequacy of the absorptive prescription, it is necessary to find a less specific framework with which to describe the increasingly accurate high-energy data. Such a formalism does not exist at present, as restricting a Regge-cut fit with

known principles, allows a ridiculous number of parameters. This is the essential difference between Regge cut and pole phenomenology. The latter had sufficiently few unknown parameters that fits to the data could determine them and hence properties of the poles without specific theoretical assumptions.

We will now discuss three examples that illustrate our three theorems.

(a) πN charge exchange

As van Hove has described,[1] the energy dependence, at fixed t, of the experimental $d\sigma/dt$ for πN charge exchange (CEX) is well described by

$$\frac{d\sigma}{dt} = A(t)\,(p_{\text{lab}})^{2\alpha_{\text{eff}}(t)-2}.$$

In fact, such a form gives an adequate fit to $d\sigma/dt$ for almost all reactions,‡ in the present high-energy range of $5 \leqslant p_{\text{lab}} \leqslant 30$ GeV/c. If there is but one pole exchanged, then of course Regge-pole theory would predict that $\alpha_{\text{eff}}(t) = \alpha(t)$, the trajectory function of the exchanged pole. However, for any data, $\alpha_{\text{eff}}(t)$ proves very useful for judging the relative contributions of cuts and/or different trajectories.

Figure 1 shows a plot of the experimental $\alpha_{\text{eff}}(t)$ versus t for πN CEX. As is well-known, this agrees remarkably well with a simple pole (the ρ-trajectory) which is roughly given by $\alpha_\rho(t) = 0.58 + t$. This was historically the first and, unfortunately, still essentially the only successful application of Regge-pole theory to fitting data over a wide range of t. The observation of nonzero polarization in this reaction led to a modification of this simple one-pole description.[2] In particular, various versions of the reggeized absorption model were advanced to successfully explain this anomaly. In Fig. 1, we have also shown the theoretical α_{eff} predicted by the two most popular of these calculations. The solid curve, marked E, uses exchange-degenerate (EXD) pole residues and so explains the dip in $d\sigma/dt$ for πN CEX at $t \sim -0.6$ (GeV/c)^2 by an intrinsic zero in the pole residue. The dashed curve, marked M, uses a model advocated by the Michigan group. Here the absorptive cut is much larger and so generates a greater deviation from a straight line in the theoretical α_{eff}. The dip in $d\sigma/dt$ is explained, quite differently, as the interference between the cut and a pole whose residue is nonvanishing at $t \sim -0.6$.

Figure 1 demonstrates that only the EXD version of the absorption model is consistent with πN CEX data. Even here, it is worth noting that this sophisticated cut model gives a fit to α_{eff} that is somewhat worse than in the original, pole only, model.

‡ I have collected empirical values of $\alpha_{\text{eff}}(t)$ for some 20 reactions in my Stony Brook talk.

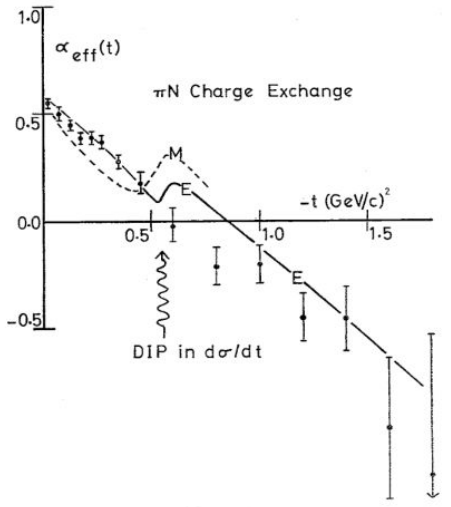


FIGURE 1

(b) Photoproduction

There is now an enormous amount of beautiful data on both forward[3] and backward photoproduction. Let us say, at once, that there is as yet no theoretical model which can explain anything but the very gross features of these data.

First of all we may dispose of Regge-pole theory. The π-exchange reactions, $\gamma p \to \pi^+ n$ and $\gamma p \to \pi^- \Delta^{++}$ are predicted to vanish at $t = 0$ if there are but well-behaved Regge poles exchanged. The size of $d\sigma/dt$ at $t = 0$ is thus a direct measure of the (cut) correction, and this is found to be from 1.5 to 3 times the simple absorption prediction. This allows us to confirm theorem (i) and further to rule out the EXD reggeized absorption model. Thus only the Michigan model which predicts such enhanced absorption to be universal remains to be considered.

However, let us complete the case for the prosecution by remarking that there is quite outstanding evidence that photoproduction is

dominated not by cuts at all, but rather by fixed poles in the j-plane. Drell[3] has described how fixed singularities, forbidden in strong interactions, are allowed in weak processes like photoproduction. In Figs. 2 and 3, we show the experimental α_{eff} for forward $\gamma p \to \pi^0 p$ and backward $\gamma p \to n\pi^+$, respectively. I believe that these are quite typical, and that all photoproduction data are as consistent with a fixed power in their energy dependence (i.e. a flat α_{eff}) as these two examples indicate. Notice the shift of the fixed power from $j = 0$ in forward processes (corresponding to the energy independence of $s^2 d\sigma/dt$) to $j = -1/2$ in the backward data (corresponding to the constancy of $s^3 d\sigma/dt$). This is an expected theoretical property of fixed poles in, respectively, meson and baryon exchange reactions. It is, of course, possible that these flat α_{eff} plots are not due to fixed poles but rather to some kinematic quirk of photoproduction. However, this seems unlikely, for, in $\gamma p \to \pi^0 p$, most theoreticians agree that one may expect the ω trajectory to dominate. Given this, it is easy to show that the kinematic structure of the unabsorbed ω in $\gamma p \to \pi^0 p$ is essentially identical to that of the embryo ρ in πN CEX. Moreover, the absorption is expected to be the same, and we at once predict roughly the same α_{eff} in both reactions. Of course, Figs. 1 and 2 are quite different.

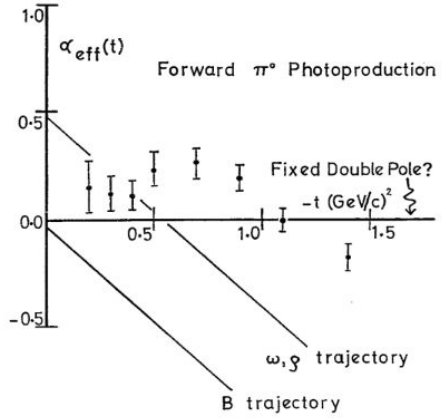


FIGURE 2

Fox, G.C., “Veni, Vidi, Vici Regge theory”,
Comments Nucl.Part.Phys. 3 (1969) 190-197.

Only notable as one of earliest phenomenology papers in area

Observables for the Analysis of Event Shapes in e^+e^- Annihilation and Other Processes

PDF

Share

Geoffrey C. Fox and Stephen Wolfram

44 Google Citations 2024

Show more

Phys. Rev. Lett. **41**, 1581 – **Published 4 December, 1978**

DOI: https://doi.org/10.1103/PhysRevLett.41.1581

Export Citation

Abstract

We present a set of rotationally invariant observables which characterizes the "shapes" of events, and is calculable in quantum-chromodynamics perturbation theory for final states consisting of quarks and gluons (G). We include the effects of fragmentation to hadrons in comparing the shapes of events from the processes $e^+e^- \to q\bar{q}$, $e^+e^- \to q\bar{q}G$, and $e^+e^- \to \text{heavy resonance} \to \text{GGG}$, and from heavy-quark and lepton production. We indicate how our analysis may be extended to deep-elastic lepton-hadron interactions and hadron-hadron collisions involving large transverse momenta.

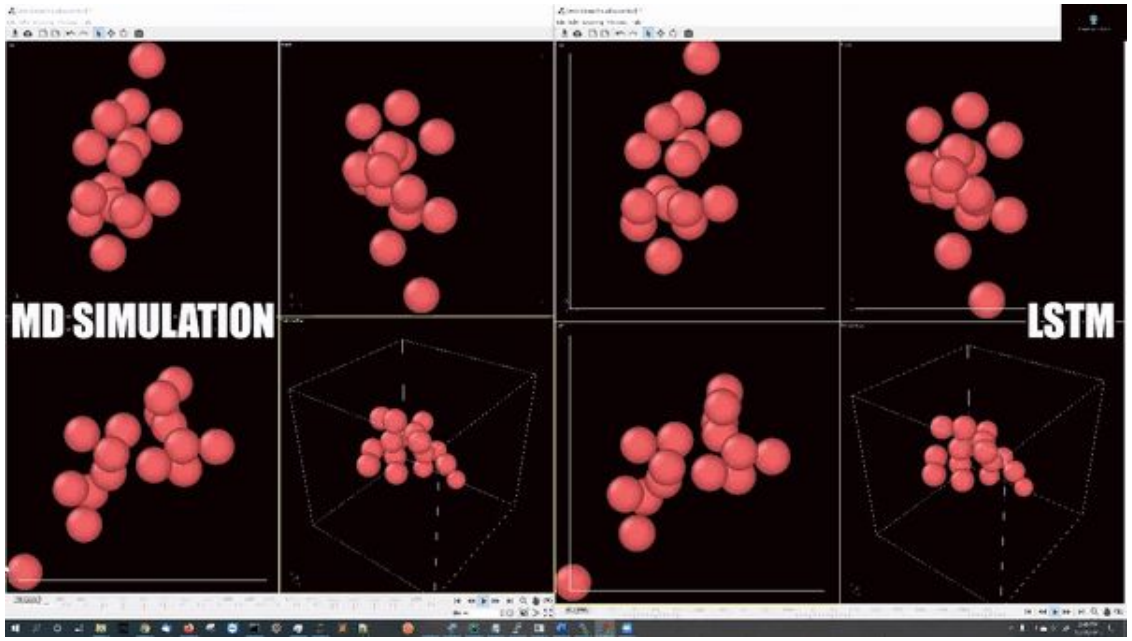
Operator Formulation of Deep Learning: Surrogates

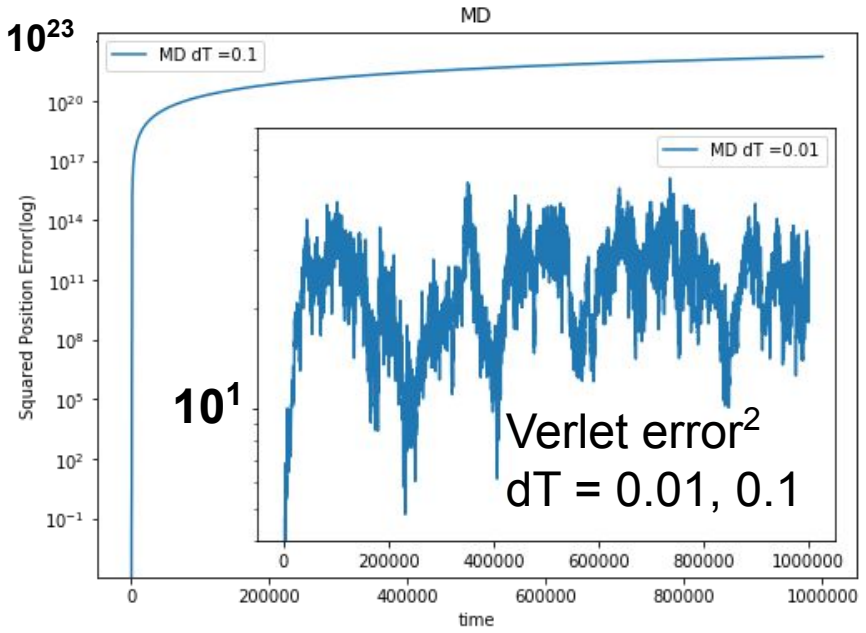
- Suppose we are solving PDE's or sets of coupled ODE's
- Typically we solve iteratively **New Values = (Differential Operator O) Previous Values**
- Classic applied math tells you nifty difference equations and spectral methods to represent Operator numerically
- Deep Learning surrogates learns the operator from classic numerics or observational data or their combination
- Inference is **New Values = (DL Operator O) Previous Values**
- This new nonlinear trained DL operator can allow much larger time steps, incorporate variations in parameters, learn potentials etc.
- **DL Operator O** is the new **theory** (Newton's laws) of science
- High order approximations are traditionally very sensitive to noise and one was taught to avoid but Deep NNs are the opposite – both verbose and robust
  - Note DL operator **O** with multiple LSTM layers has 100s-100,000 parameters
  - Newton's laws for this have 2-4 parameters

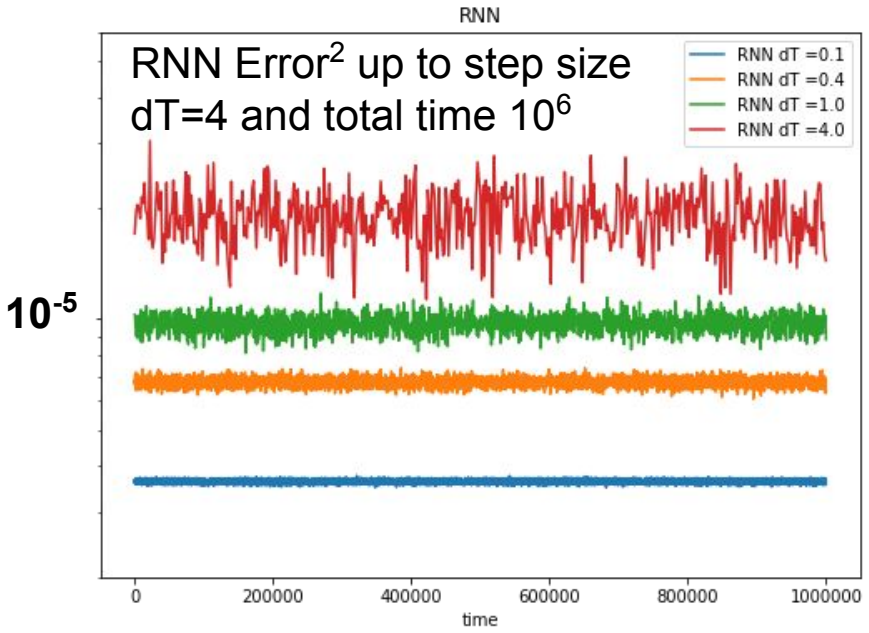
Learn Newton's laws with Recurrent Neural Networks

JCS Kadupitiya, Vikram Jadhao

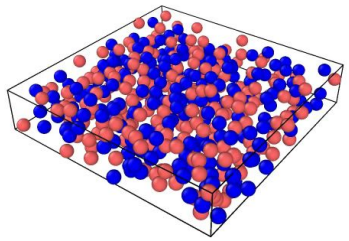
- **Deep Learning is revolutionizing (spatial) Time series Analysis**
- Good example is integrating sets of differential equations
- Train the network on traditional 5 time step series from (Verlet) difference equations
- Verlet needs time step .001 for reliable integration but
- Learnt LSTM network is reliable for time steps which are **4000 times longer** and also learn potential.
- **Speedup is 30000** on 16 particles interacting with Lennard-Jones potentials
- 2 layer-64 units per layer LSTM network: **65,072 trainable parameters**
- 5000 training simulations





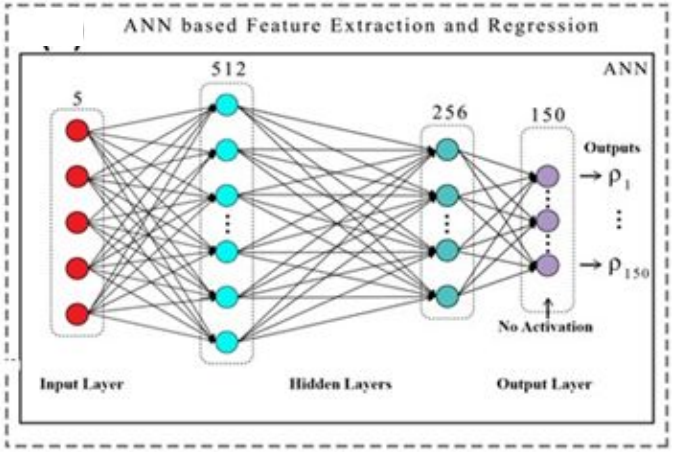


Examples of Simple Surrogates (work with JCS Kadupitiya, Vikram Jadhao)



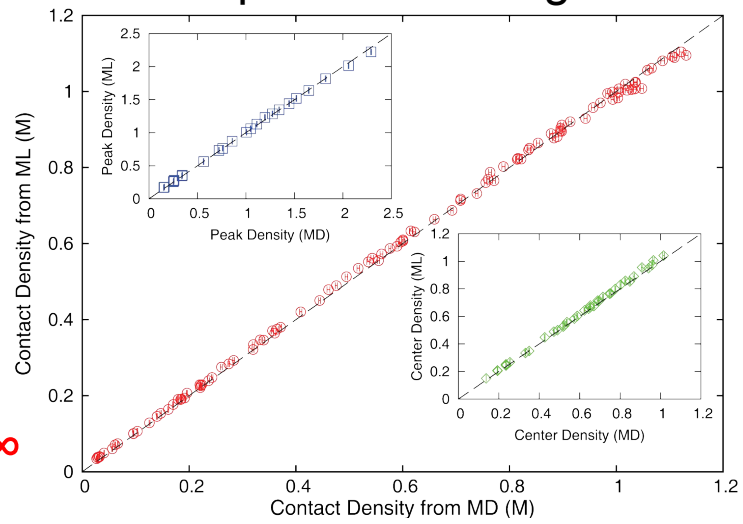
- Extraction of **ionic structure in electrolyte solutions confined by planar and spherical surfaces**.
- Classic HPC code written with C++ and accelerated with hybrid MPI-OpenMP.

The Learning Net



- Uses quite small Multi-layer perceptron DL Network to predict 150 observables from 5 input parameters (~5000 in training set)
- DL outperforms other AI choices
- Deployed on nanoHUB for **education** (an attractive use of surrogates so students get answers fast)
- General Electric uses similar approach to give interactive Engine design options (200 in training set)

Direct simulation compared to Surrogates



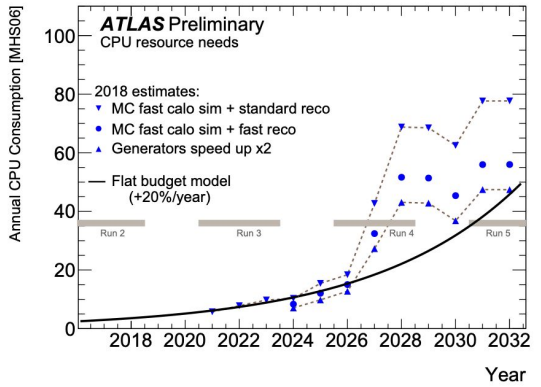
$$\text{Effective Speedup } S = \frac{T_{seq}(N_{lookup} + N_{train})}{T_{lookup}N_{lookup} + (T_{train} + T_{learn})N_{train}} \rightarrow 10^6 \text{ as } N_{lookup} \rightarrow \infty$$

Summary of Initial Surrogate Research

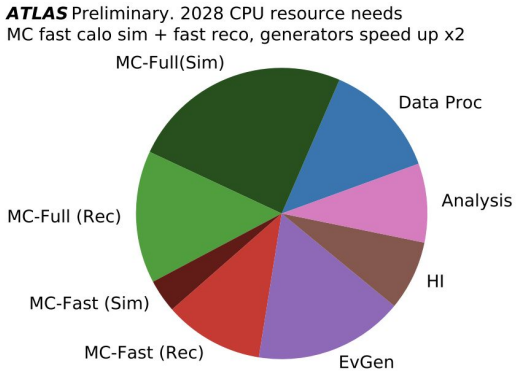
- Useful if need to run a particular job fast (real time warnings) or lots of times (education or parameter searches)
- These are **classical non-generative Deep Learning-based AI** and sometimes just generalize classical differential equation solvers
- If evolution statistical either in dynamics or in uncertainties in data or equations of motion, then **generative AI** needed

Generative AI for Kaggle Calorimeter Surrogates

- Generation of simulated events is a significant computing load at LHC
- These are typically generated by GEANT4 from known physics of particle material interactions
  - AI surrogates must be generative to mimic a Monte Carlo
  - **Errors** are largely proportional to √**Energy** and there are **significant correlations**; often ignored

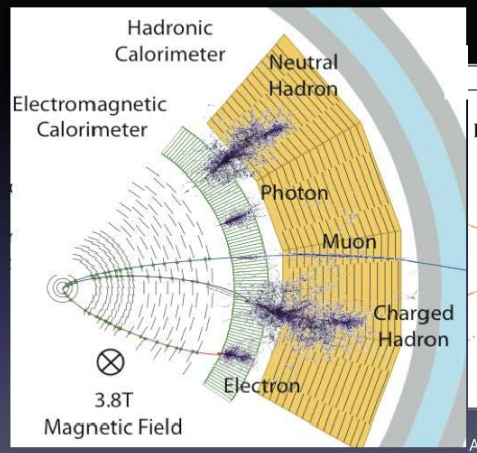


(a)

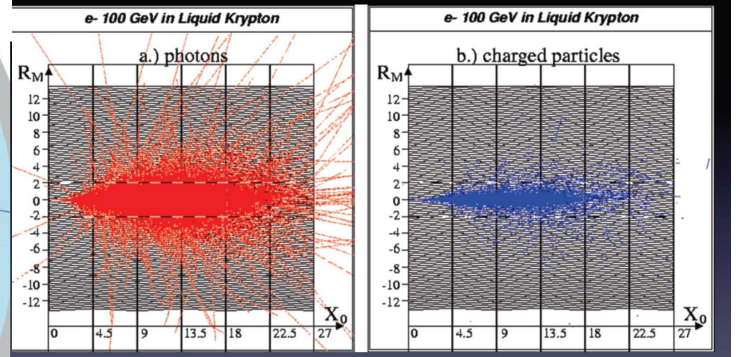


(b)

- In nuclear and particle physics calorimetry refers to the detection of particles through total absorption in a block of matter
  - The measurement process is destructive for almost all particle
  - The exception are muons (and neutrinos) → identify muons easily since they penetrate a substantial amount of matter
- In the absorption, almost all particle's energy is eventually converted to heat → calorimeter
- Calorimeters are essential to measure neutral particles



EM shower development in liquid krypton (Z=36, A=84)



ANT simulation of a 100 GeV electron shower in the NA48 liquid Krypton calorimeter (D.Schinzel)

Generative AI Surrogate Methods for Kaggle Calorimeter

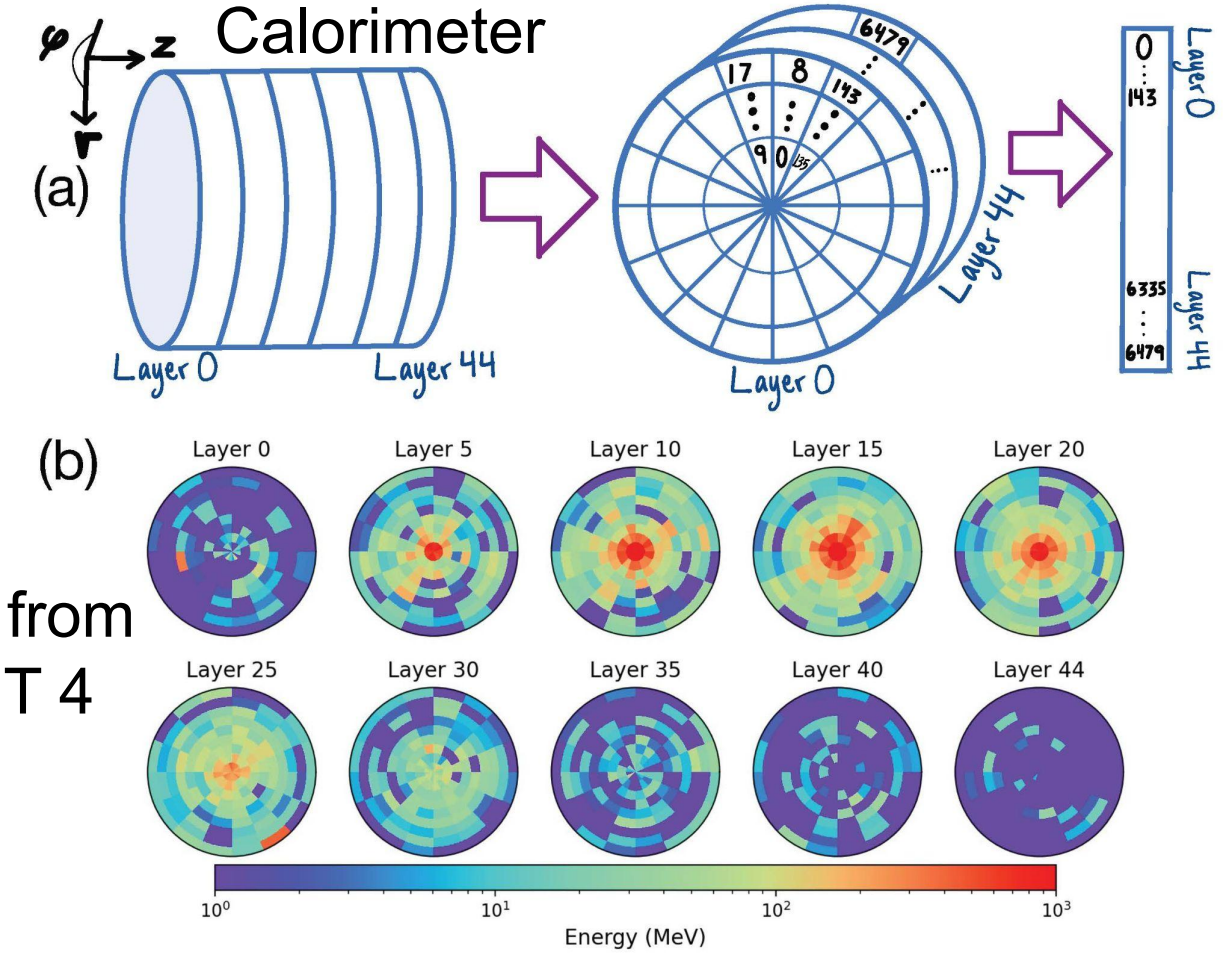
Extensive review 2410.21611v1.pdf

Timing of Methods
Geant4 ~ 1s
GPU (A100) ~ 2ms
Calo4pQVAE 0.2ms

- GAN
- VAE Variational Autoencoder
- Normalizing Flow
- Diffusion models (Currently best accuracy but slow)
- **Calo4pQVAE Quantum Spin Generator from Dwave + VAE**

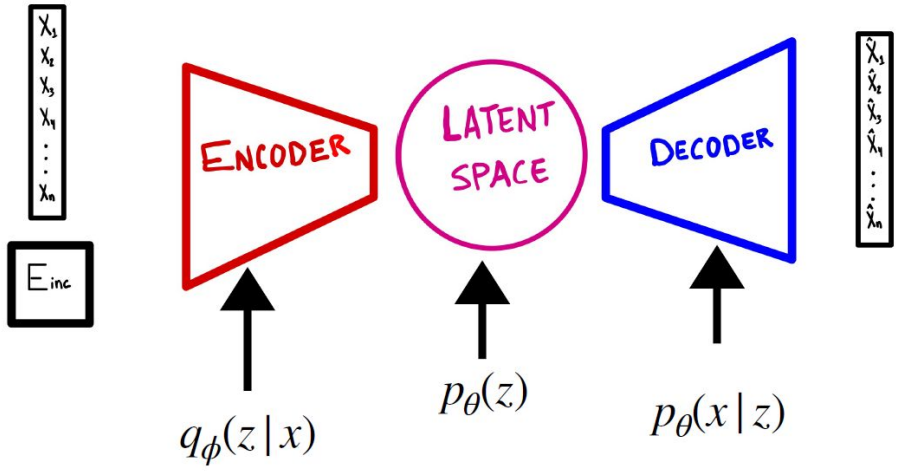
- Kaggle Dataset 2
- Each event 6480 voxel energies
- Incident electrons normal incidence to calorimeter; 1 Gev to 1 TeV energy
- Can extend usefully by varying incident particle, angle of incidence, position of incidence, calorimeter design

typical signal from GEANT 4



Calo4pQVAE: A Quantum-assisted 4 Partite VAE Surrogate for Particle-Calorimeter Interactions

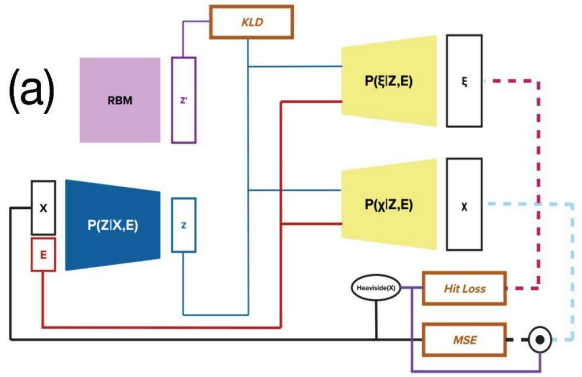
Variational Autoencoders

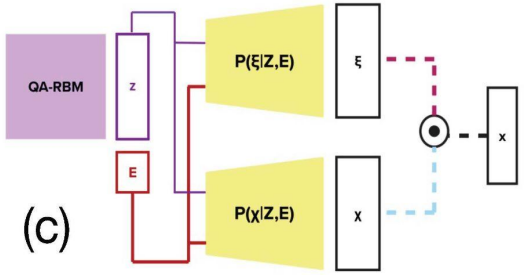


$$\mathcal{L}_{\phi,\theta}(x) = \langle \ln p_\theta(x\,|\,z) \rangle_{q_\phi(z|x)} - \langle \ln \frac{q_\phi(z\,|\,x)}{p_\theta(z)} \rangle_{q_\phi(z|x)}$$

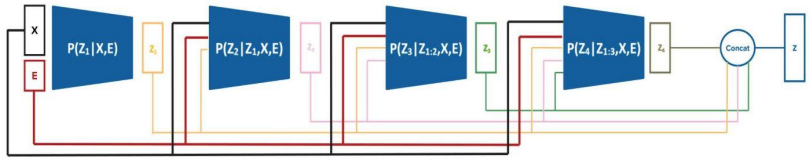
- There are two terms in QVAE loss
- Classic MSE loss and a KL Divergence that is forcing the distribution to be correct

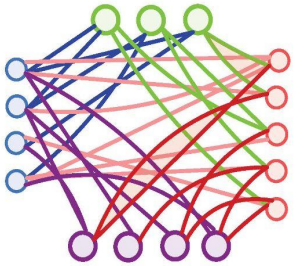
a) Data flows from left to right: it is first encoded into the latent space, then decoded to reconstruct the voxel energies, with **incidence energy conditioning** both the encoder and decoder.





b) The encoder is hierarchical with 4 pipelined sub-encoders, each integrating previously generated data and the original input. The encoded data is utilized to train the RBM. (Restricted Boltzmann Machine)

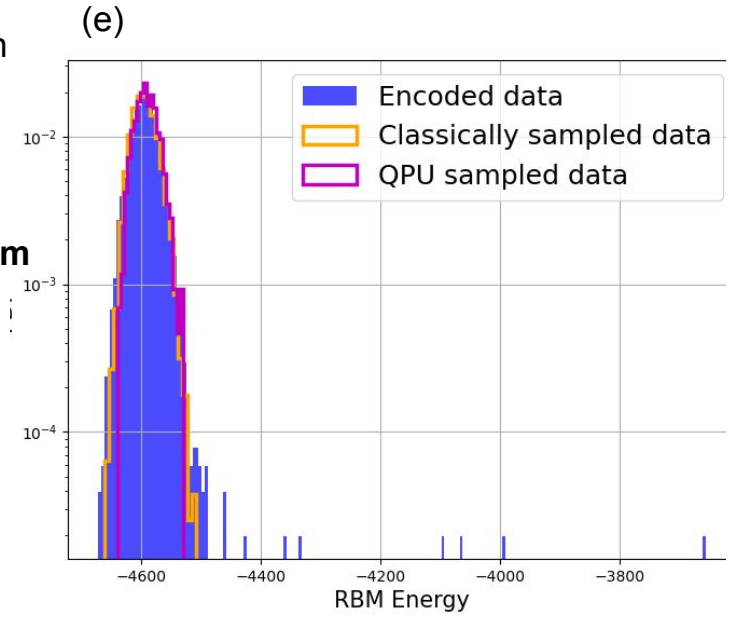




c) Inference

d) Calo4pQVAE uses a discrete binary latent space and a Boltzmann distribution prior, where the energy function correspond to a sparse 4-partite graph, mimicking the **DWAVE Pegasus topology Quantum Annealer**

e) RBM energy histogram in the case of the encoded validation data set after training in blue. The yellow and purple histogram shows the RBM energy histogram for classically sampled data via block Gibbs sampling and via the QA



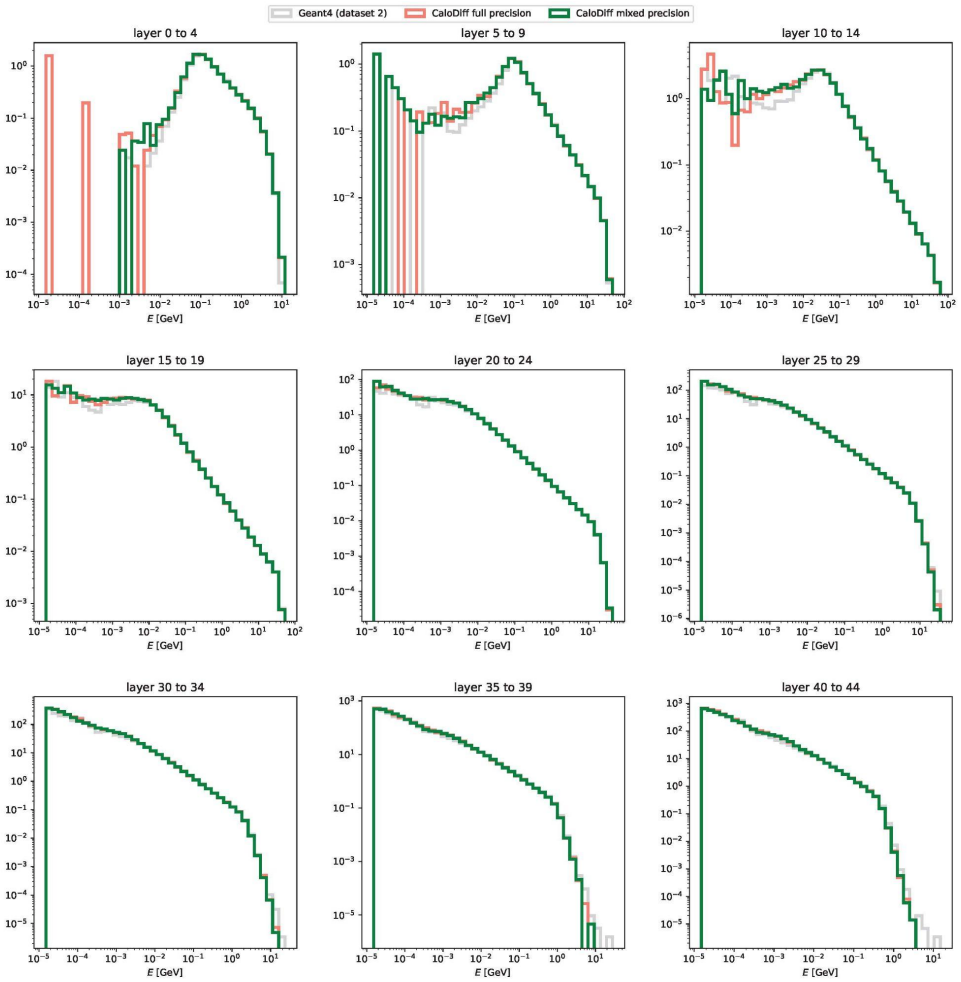
Why use Diffusion model

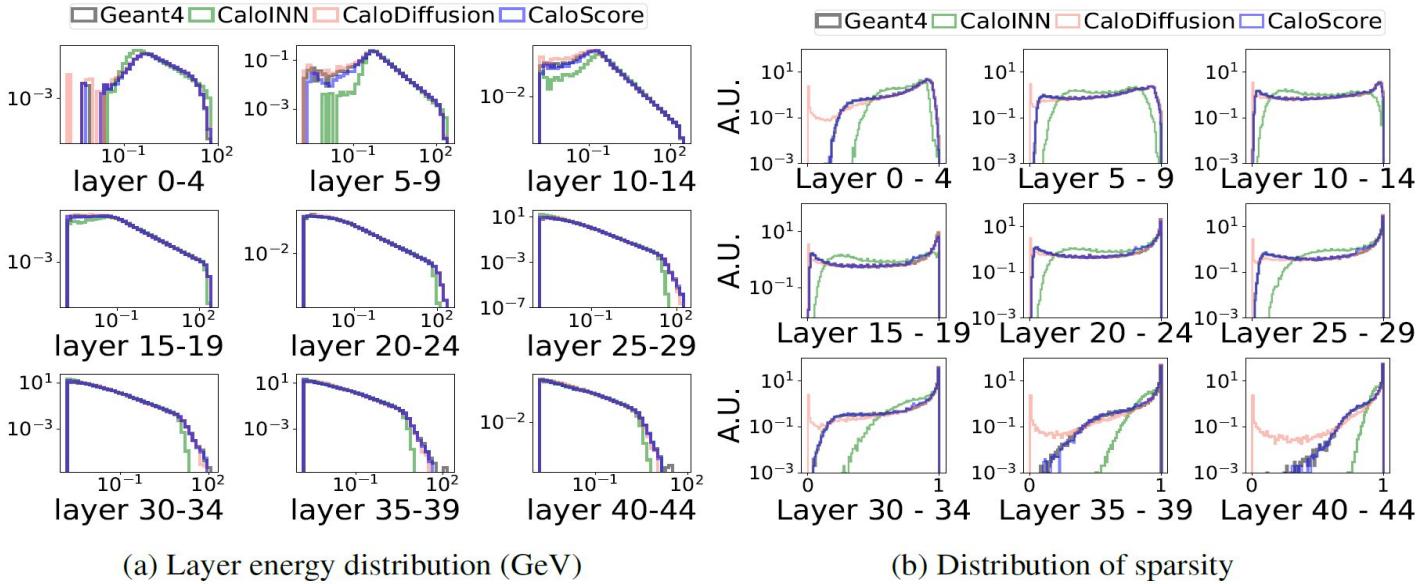
- We using a method that **starts with events of interest, lands in a latent space, and then returns to same event of interest**
  - The last step is decoder used to generate new events
- Surrogates aim to find a latent space so that generating points with a uniform Gaussian distribution in latent space, leads to events with correct distribution as output of neural network
- Diffusion equation generates points in such a latent space automatically
- **So diffusion model only needs to minimize MSE loss for neural network to generate correct final state**
- All other methods need the **composite loss** which ensures that both the latent space is a Gaussian and that the final states are correct
- In calorimeter and some other cases, diffusion model has best quality results with **fewer neural network parameters** but longer execution time

Comparison of Geant and Generative Calorimeter Simulators

Physics GenAI and Precision

- FP16 (.12 sec)versus FP32 (.22 secs) speeds up as in LLM

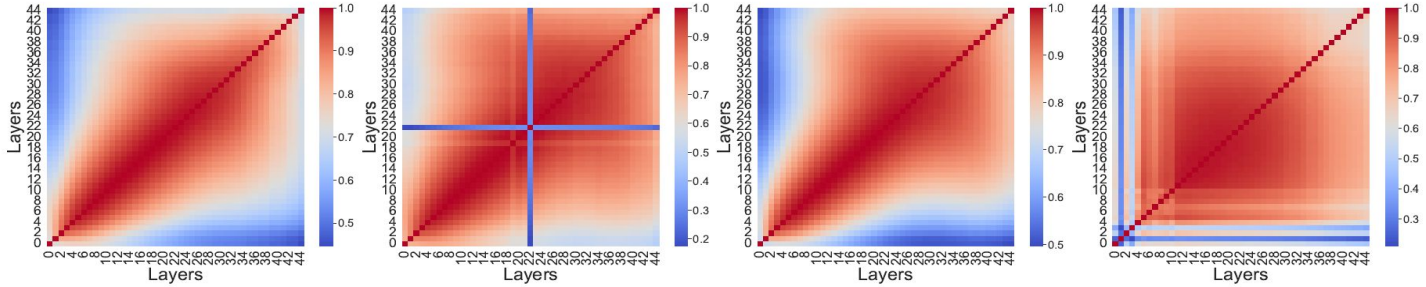




(a) Layer energy distribution (GeV) (b) Distribution of sparsity

Correlations

Sparsity = Ratio of non zero voxels to total number of voxels per layer



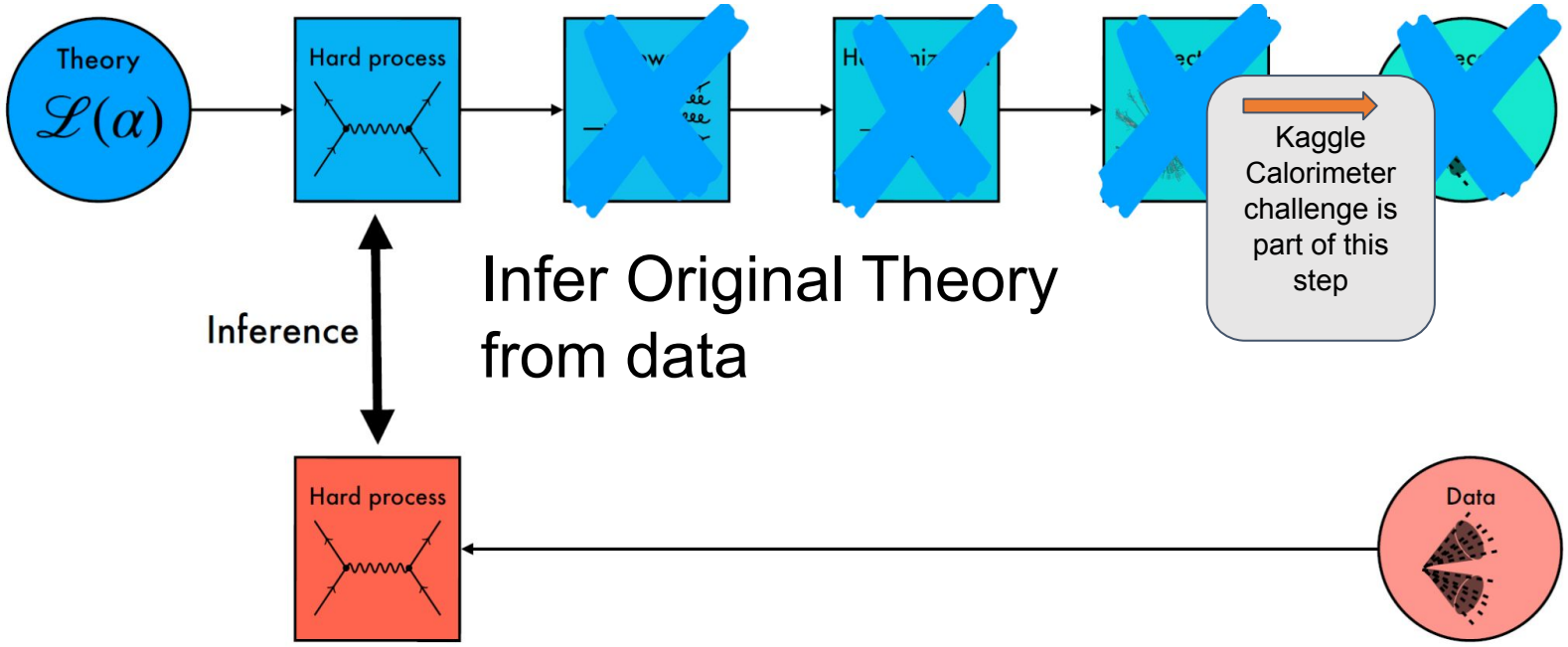
(a) Geant4 (b) CaloScore (c) CaloDiffusion (d) CaloINN

- **Chisq χ^2 = Σ (Observed - Predicted)^2 / Error^2** is classic MSE
- **MSE Unfair to small Energies**; 1 GeV /1 TeV weights off by factor 1000
- Observables should not have diagonal MSE but **Correlated**

$\Sigma_{i,j}$ **(Observed(*i*) - Predicted(*i*)) (Observed(*j*) - Predicted(*j*)) $C^{-1}(i,j)$**

Generative AI in Science I

- Kaggle Calorimeter illustrates **generative AI as a surrogate for statistical simulations**
- But **generative AI** heart of deep learning for **data assimilation** which is probably a larger field
- Google WeatherBench 2 Weather Benchmark
- ACM Gordon Bell Prize for Climate Modelling 2023-2032
- AI is revolutionizing CERN LHC Data analysis. Quantum Physics is a statistical hierarchy



Top blue boxes are typical data simulation pipeline Bottom red boxes are **data assimilation process** to infer each step from data using simulation-based inference

Generalized Unfolding

- **ML4Jets workshop** is excellent starting point and we replace **Partons** by **Regge poles and Resonances**
- When we have a hierarchical set of transformations, one can in principle **unfold to any step of process**
- **OMNIFOLDHI** unfolds detector and background process
- Classically (me 1967) one constructs a model and **fits the model to the data** using a few parameters and never steepest descent
- With AI, one builds model, generates events according to model and **unfolds to any step of model** with steepest descent and lots of parameters
- **LHC:** Model is QCD plus parton hadronization
- **Jefferson Lab:** Model is Regge plus Resonance model
- Note both are phenomenological and from 1970's
- One can add backgrounds like **final state interactions** and unfold these away if in **probability not amplitude**
- Jefferson Laboratory needs **amplitude based unfolding** not just probabilities

Quantum-chromodynamic approach for the large-transverse-momentum production of particles and jets

PDF

Share

R. P. Feynman, R. D. Field, and G. C. Fox

Show more

Phys. Rev. D **18**, 3320 – **Published 1 November, 1978**

DOI: https://doi.org/10.1103/PhysRevD.18.3320

Export Citation

Abstract

It is shown that if, in a calculation of high-transverse-momentum ($p_\perp$) meson production in hadron-hadron collisions, one includes not only the scale-breaking effects that might be expected from asymptotically free theories but also the effects due to the transverse momentum of quarks in hadrons *and* further adds contributions from quark-gluon and gluon-gluon scattering to those of quark-quark scattering *then* the results are not inconsistent with the data. The approach yields the correct magnitude and an apparent approximate $\frac{1}{p_\perp^8}$ behavior in accord with single-particle data for the energy range currently observed. Two-particle correlations are examined. Because of scale-breaking effects and the presence of gluons, the theory does not have the problem of predicting too many away-side hadrons at large $p_\perp$ as did an earlier quark-quark scattering "black-box" approach. We conclude that the quantum-chromodynamics approach is in reasonable accord with the data although theoretical uncertainties (especially at low $p_\perp$) make incontrovertible conclusions impossible at present. Crucial tests of the theory require higher $p_\perp$ than are now available; estimates for this region are made.

Jefferson Laboratory Scientific Assistant

- Start with **AuroraGPT** for general science (**Trillion Parameter Consortium**)
- Collect **Jefferson Lab Physics papers and workshops**
  - Have **new workshops** to capture material as much material not on web
  - e.g. Many of my papers and all my physics presentations lost and S-Matrix theory stopped in 1980 before the web
  - Collect Jefferson Laboratory Questions and Answers from places like this workshop
- Repeat previous related work using AuroraGPT to get **AuroraGPT-Physics**
  - **PhysBert** from LBL
  - https://huggingface.co/thellert/physbert_cased
  - https://arxiv.org/abs/2408.09574
  - This is referenced by https://arxiv.org/abs/2411.14877 **Astro-HEP-BERT**
- Obtain **HadronGPT** by fine tuning **AuroraGPT-Physics** with Jefferson Laboratory knowledge
  - Teach HadronGPT **names of particles**
- Teach **HadronGPT** how to do unfolding for
  - **Reaction Dynamics**
  - **Reaction Dynamics plus Spectroscopy**
  - **Inclusive reactions**
- Fine tune on **Questions and Answers**

Goals?

From Franck Cappello

Aurora is:
166 Racks
10,624 Nodes
21,248 CPUs
63,744 GPUs
84,992 NICs
8 PB HBM
10 PB DDR5c

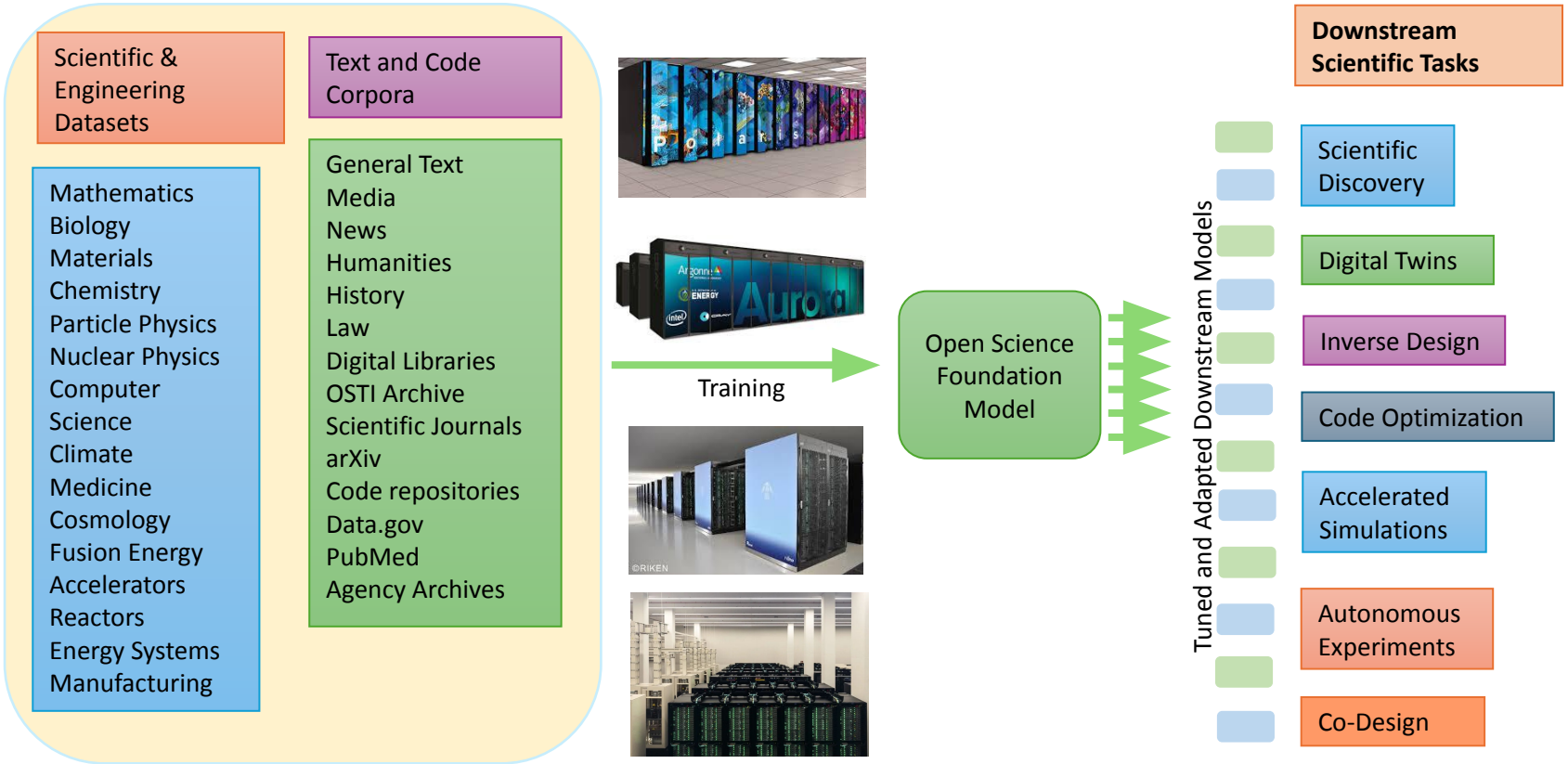
AuroraGPT*:

- **General purpose scientific LLM** – broadly trained – general corpora plus scientific papers and texts and structure science data
- **SAFE:** Trustworthiness, Safety, Security, Robustness, Privacy, Machine Ethics
- **Explore pathways** towards a "Scientific Assistant" model *(could be LLM + Monte Carlo Tree Search & Reinforcement Learning, Symbolic reasoning)*
- **Build with international partners** (RIKEN, BSC, others)
- **Multilingual** – English, 日本語, Français, Deutsche, Español, Italiana
- **Multimodal** – images, tables, equations, proofs, time-series, graphs, fields, sequences, etc.

*named after the Leadership Class Supercomputer at Argonne that will be used for much of the training

What?

- **Open Science Foundation Models**
- **Initially with textual interface** (scientific prompts), later multi-modal
- **Single model or Mixture/Combination of Experts**
  (scale progressively: 7B, 70B, … ~1 T parameter)

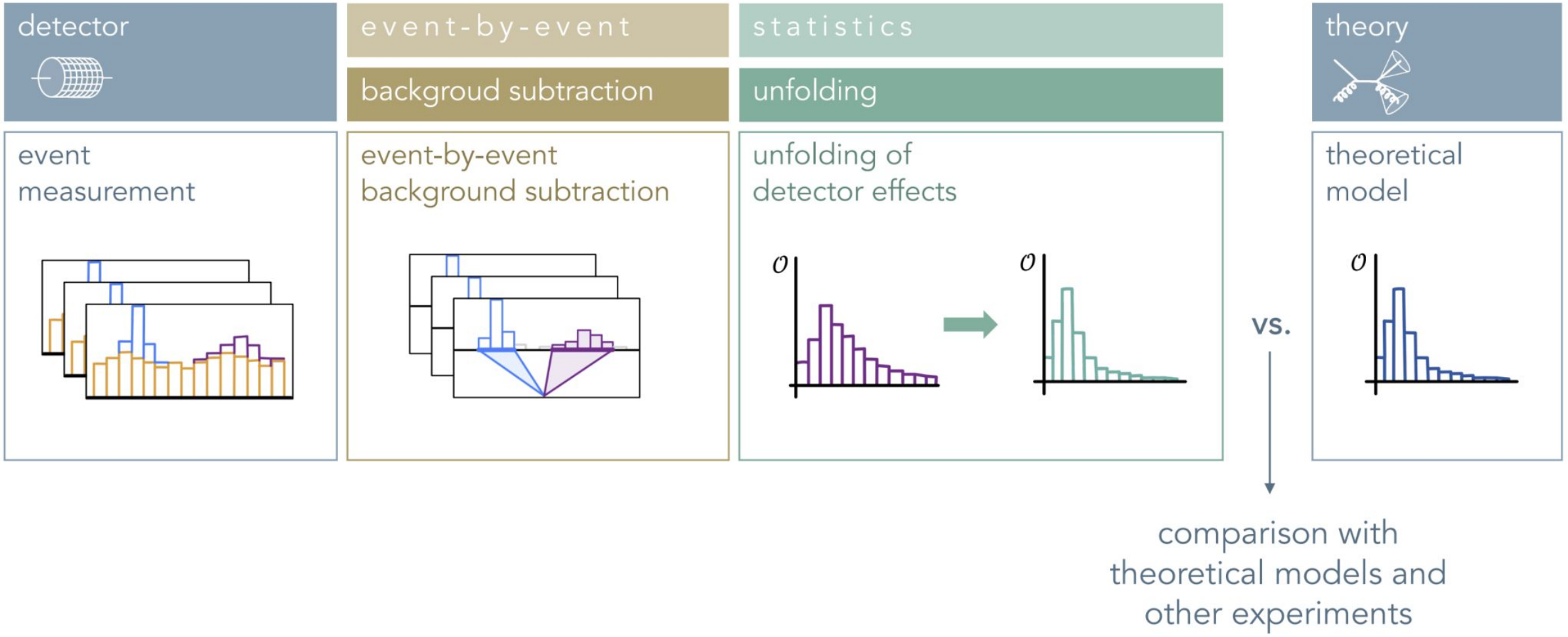


For what outcomes?

- **Datasets and data pipelines** for preparing Science training data
- **Software Infrastructure and workflows** to train, evaluate, and deploy LLMs at scale for scientific research purposes
- **Evaluation of state-of-the-art LLM models** to determine where they fall short in deep scientific tasks and where deep data may have an impact
- **Assessment of the approach** of augmenting web training data with two forms of data specific to science
  - Full-text scientific papers
  - Structured scientific datasets (suitably mapped to narrative form)
- **Research grade artifacts (models)** for scientific community and adaptation for downstream uses
- **Promotion of responsible AI** best practices where we can figure them out
- **International collaborations** around the long-term goal of AGI for science

Unfolding at LHC Example

This heavy ion experiment did not unfold to the theory. Rather it unfolded away both the background and the detector effects i.e. halfway down the chain



Generative AI in Science II

- Unfolding of theory from observed data is generalize**d data assimilation** and further current **best results come from diffusion models**
- **Errors** are "**revolutionized**" by these ideas: **store data as an ensemble** of possible configurations and not unrealistic values plus errors or better but even less realistic, values plus correlation matrix
  - Size of correlation matrix is square of data size
- **Generative AI and Diffusion models are not very well understood**
- Mathematics clear but hyperparameters etc. very unclear
  - We don't see a discussion as to **why diffusion model best for Calorimeter surrogates** (and other cases) where it has **fewer neural network parameters** but more random numbers so runs slower
  - **Need more experimentation and more educational examples e.g. well documented benchmarks**
  - For those exploring other applications diffusion equations don't easily work well

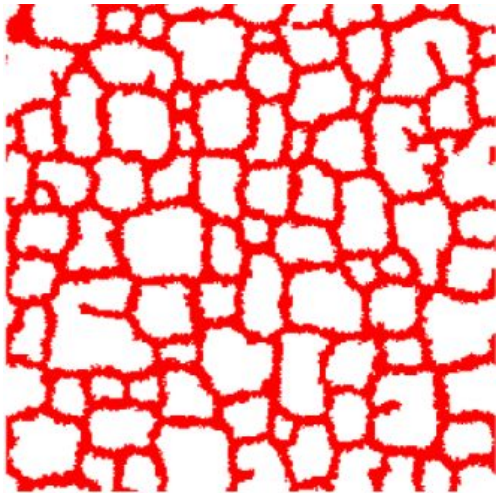
Diffusion models as a surrogate model for Biology

Forward problem
(model surrogates)

Biological system

Parameters

(x,y)



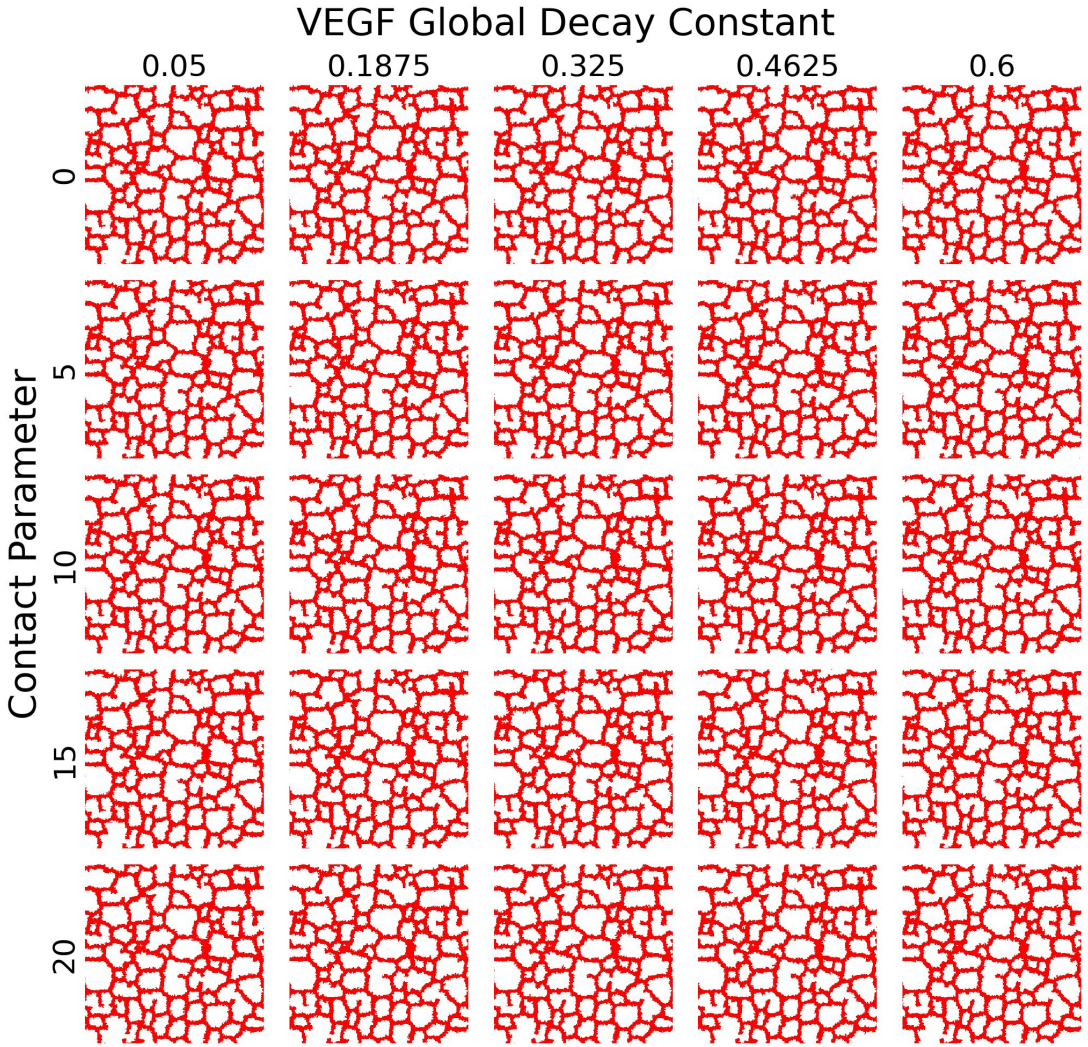
Inverse problem

Why use generative AI for Biology?

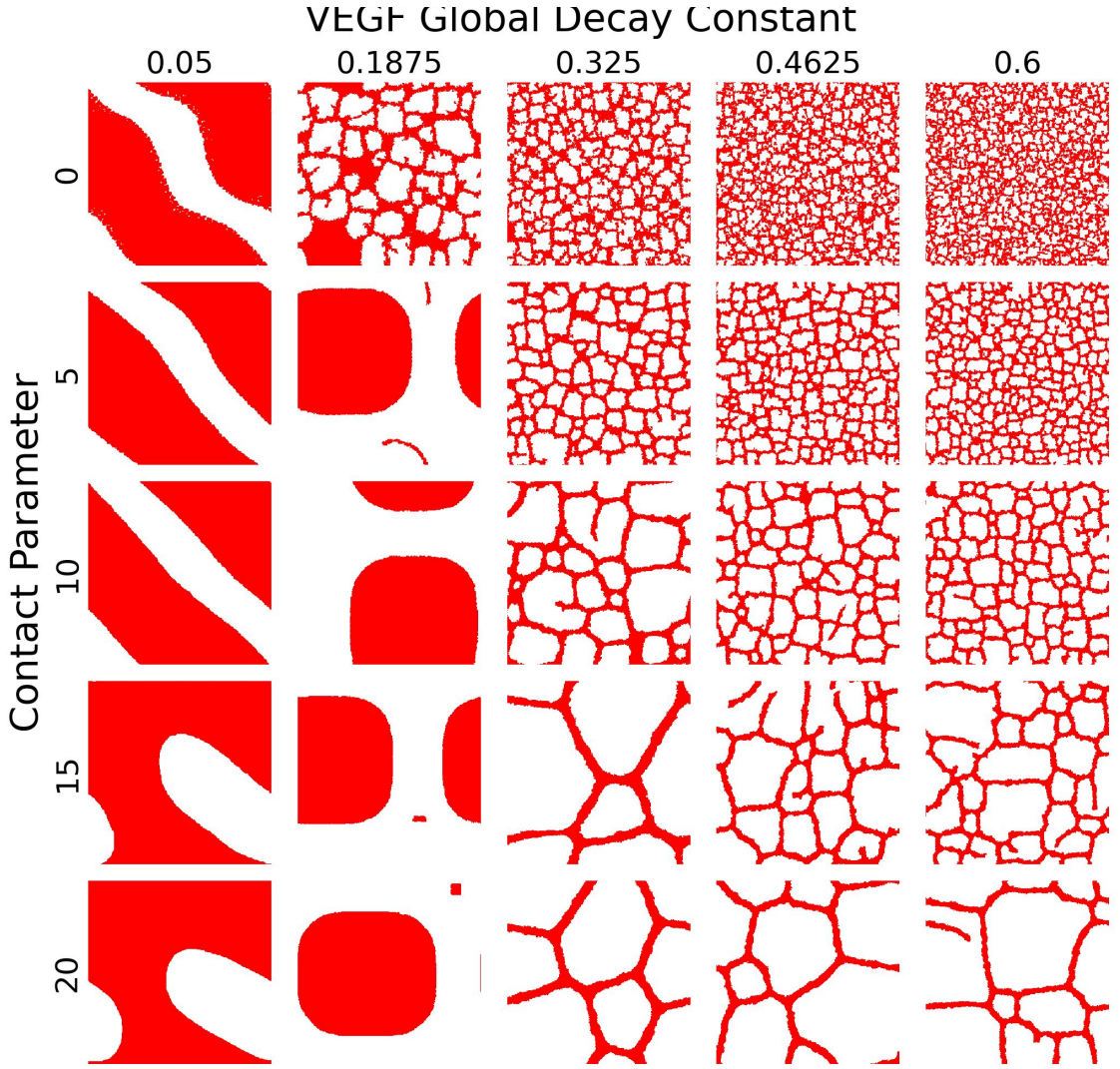
- *in silico* models of biological systems are often computationally expensive
- AI based model surrogates can significantly reduce the computational cost of evaluating a model
- *in silico* models of biological systems are often stochastic. Deterministic neural networks may be unable to adequately reproduce stochastic biological systems
- Generative AI techniques learn distributions of data and may be able to be effective candidates for surrogate models of biological systems
- Cellular-Potts agent-based model of *in vitro* vasculogenesis. **Aim at personalized medicine based on template models with patient specific customizations**

Direct Simulation

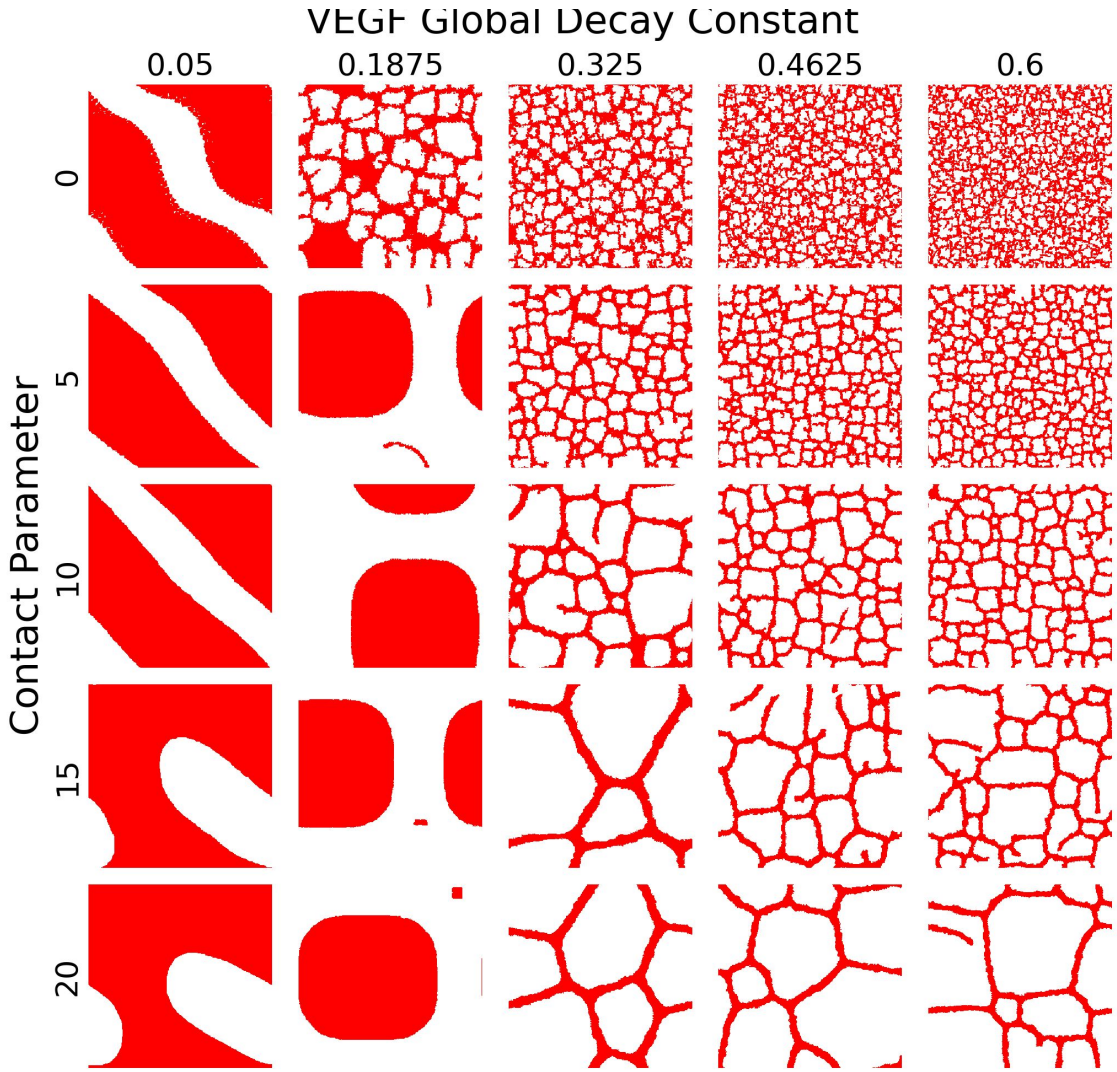
Initial Condition



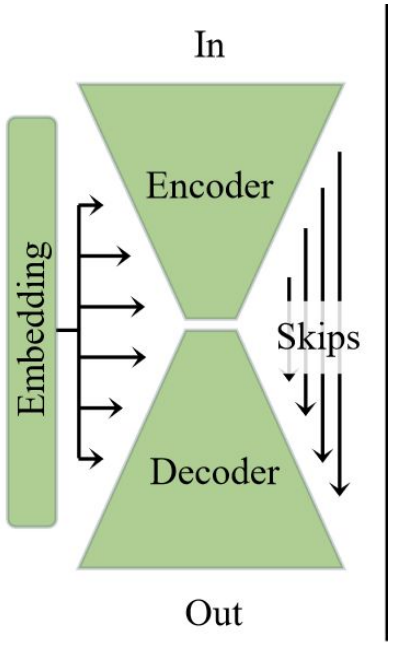
Steady State Configuration



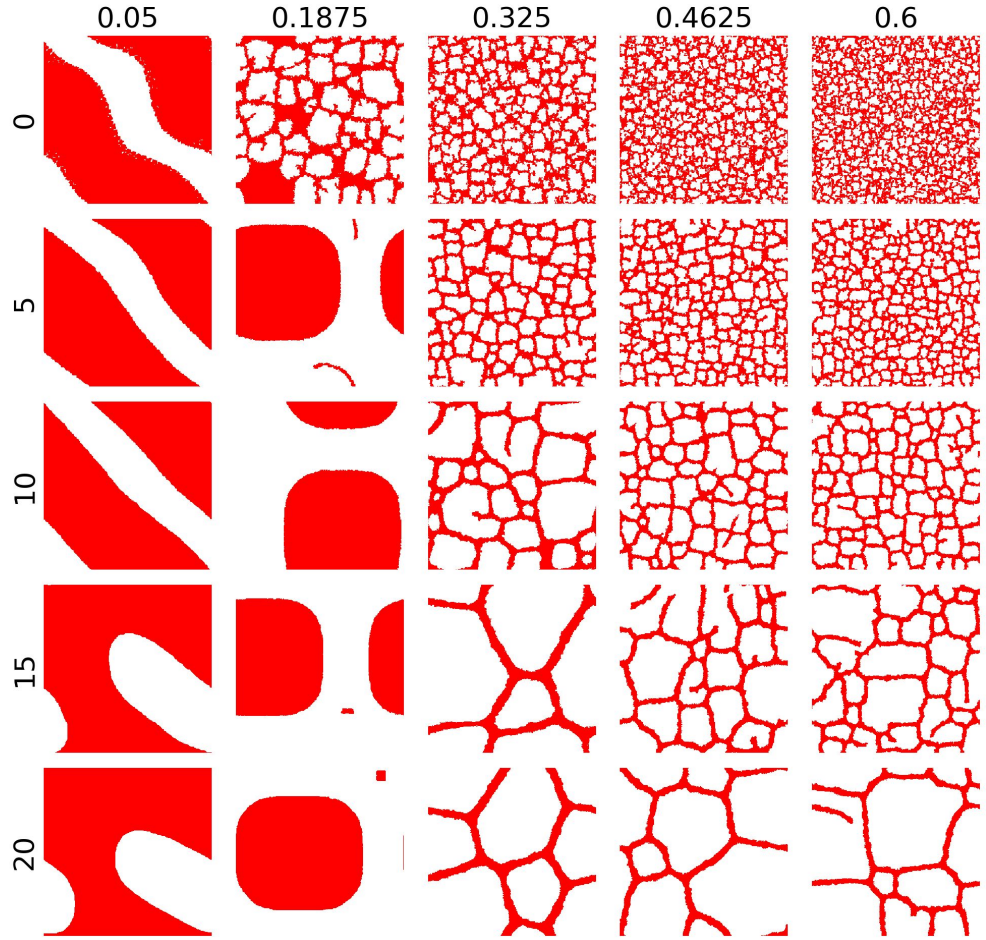
Forward problem - arriving at the steady state condition



Model training

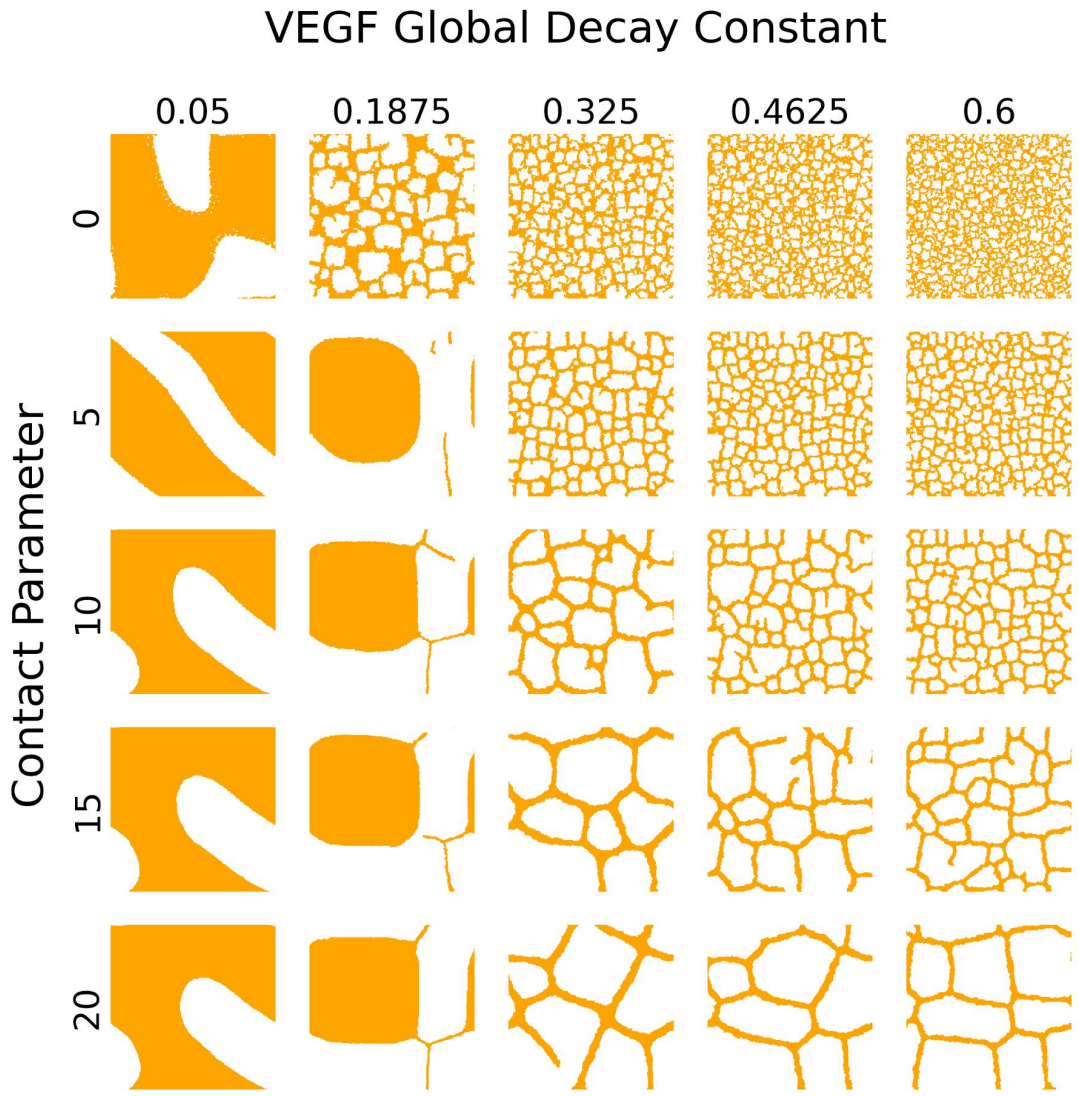


- Training data consisted of 5,100 images per class
- We trained a class-conditional diffusion model using the modified ADM network architecture
- Training followed the NVIDIA EDM training framework[1]

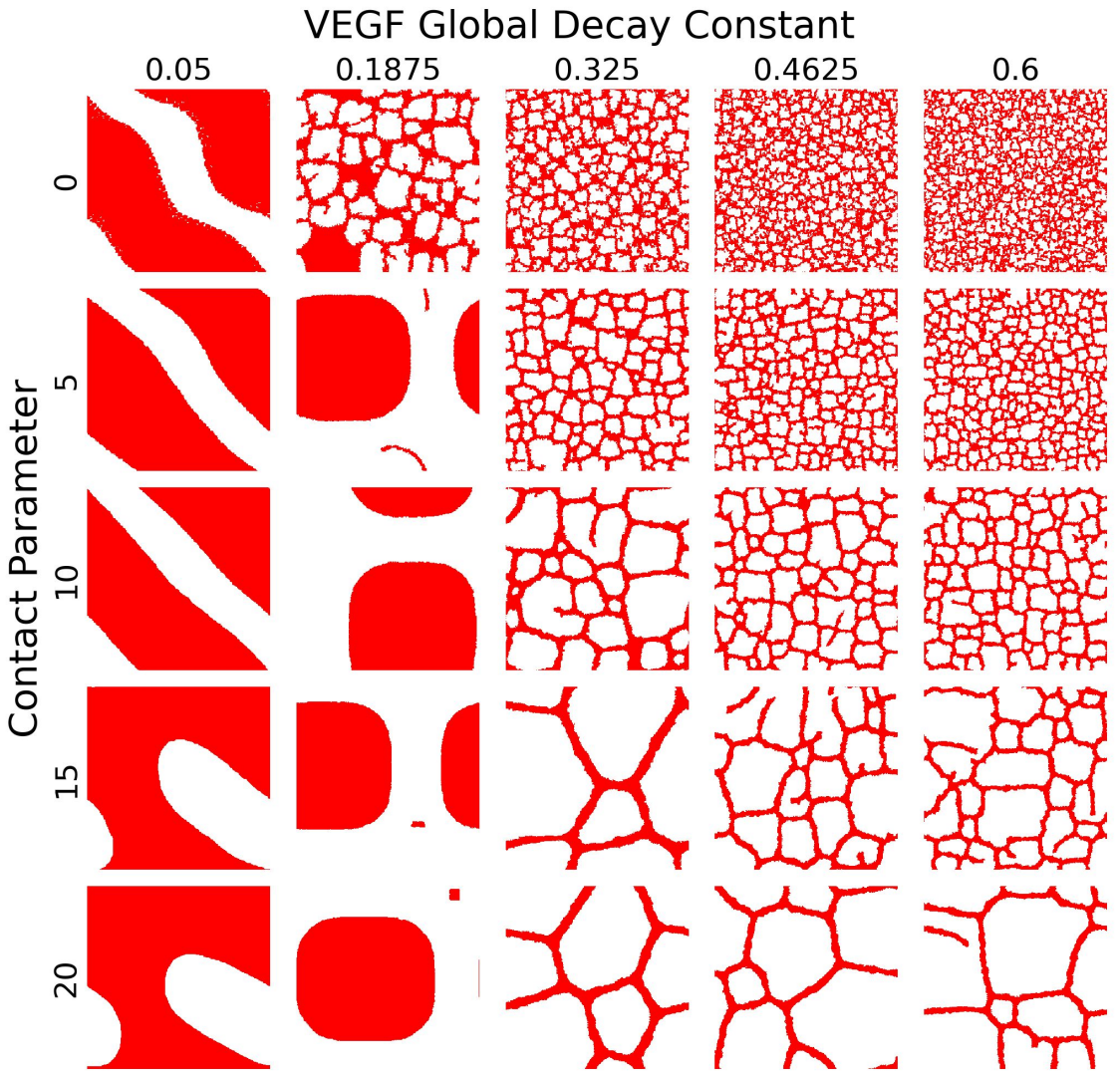


[1]Karras et. al, CVPR (2024), arXiv:2312.02696

Diffusion Model



Ground Truth



Bottom left hard

Conclusions from biology work-in-progress

- We are able represent most of the parameter space from a relatively small training dataset
- Sample generations took 20 seconds compared to 8 minutes per simulation (fast as in Julia)
- Denoising diffusion models may have potential as surrogate models for stochastic computational models of biological virtual tissue systems