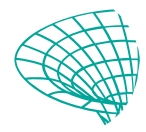


Centaur Science: Adventures in AI+Physics

Jesse Thaler



MASSACHUSETTS INSTITUTE OF TECHNOLOGY
MIT CENTER FOR THEORETICAL
PHYSICS - A LEINWEBER INSTITUTE



SIM NS
FOUNDATION



DatA Seminar, IJCLab, Université Paris-Saclay — May 4, 2026



The NSF Institute for Artificial Intelligence and Fundamental Interactions (IAIFI /aI-faI/ iaifi.org)



*Artificial intelligence
as a pathway to
scientific insight*

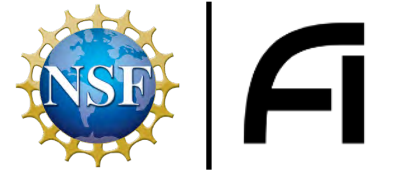
IAIFI

*Physics intelligence
as a pathway to
AI innovation*

Launched August 2020
















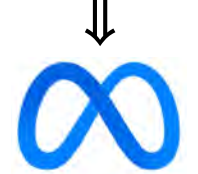

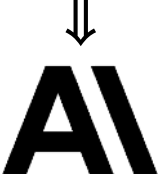


*Progress in AI+Physics driven by
early career talent with interdisciplinary training*

Empowering the Next Generation of AI + Physics Talent



IAIFI Postdoctoral Fellows

Plus 2026-29 fellows to be announced!

Albergo	Boyda	Bright-Thonney	Cuesta	Dogra	Feng	Gagliano	Gerdes	Golubeva	Grosso	Harvey	Luo	Micallef	Mishra-Sharma	Yang
														
AI and Statistical Physics		AI for Particle Physics		Mathematical Physics of AI	AI for Scientific Imaging	AI for Time-Domain Astronomy	AI for Theoretical Physics		AI for Collider Physics	AI for String Theory		AI for Neutrino Physics		AI Frontiers of Reinforcement Learning

IAIFI Summer School & Workshop

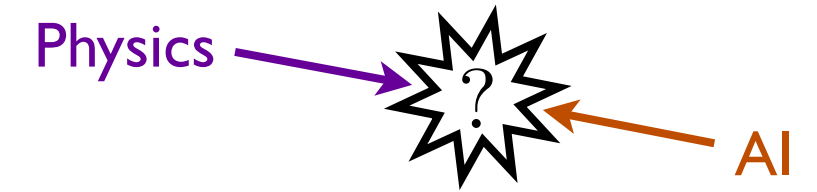
Photos from 2024 Edition @ MIT



AI IAIFI
Summer School
August 3–August 7
2026



Lessons from the Physics / AI Collision



To meet the standards of scientific rigor and performance,
we need to teach machines to “**Think like a Physicist**”

E.g.: symmetries, robustness to systematics, exactness guarantees, statistical inference, ...

But to fully capitalize on AI technologies,
we also need to teach physicists to “**Think like a Machine**”

E.g.: computational complexity, reframing via optimization/search, algorithmic reasoning, ...

The Power of “**Centaur Science**”

Progress in **computation and information theory** has long been intertwined with progress in the **physical sciences** (e.g. statistical mechanics, lattice gauge theory, quantum computers)

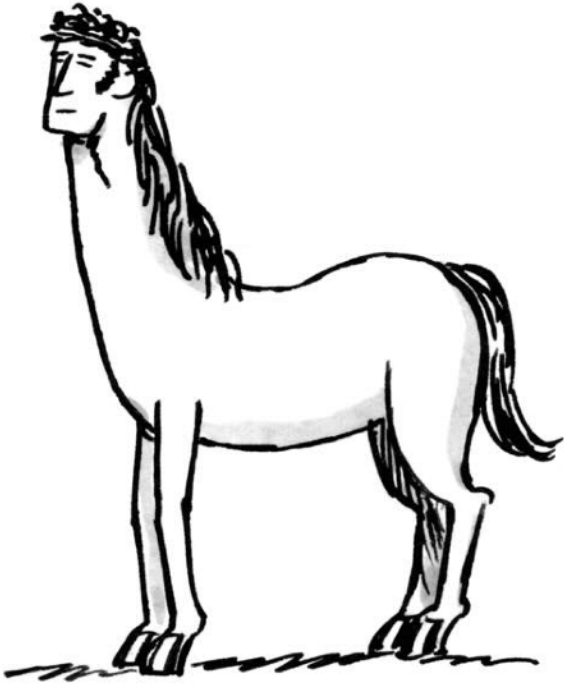
THE CENTAUR SCALE



NOT ENOUGH HORSE



THE RIGHT AMOUNT OF HORSE



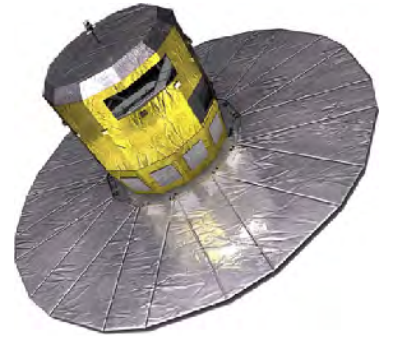
TOO MUCH HORSE

JOEDATOR

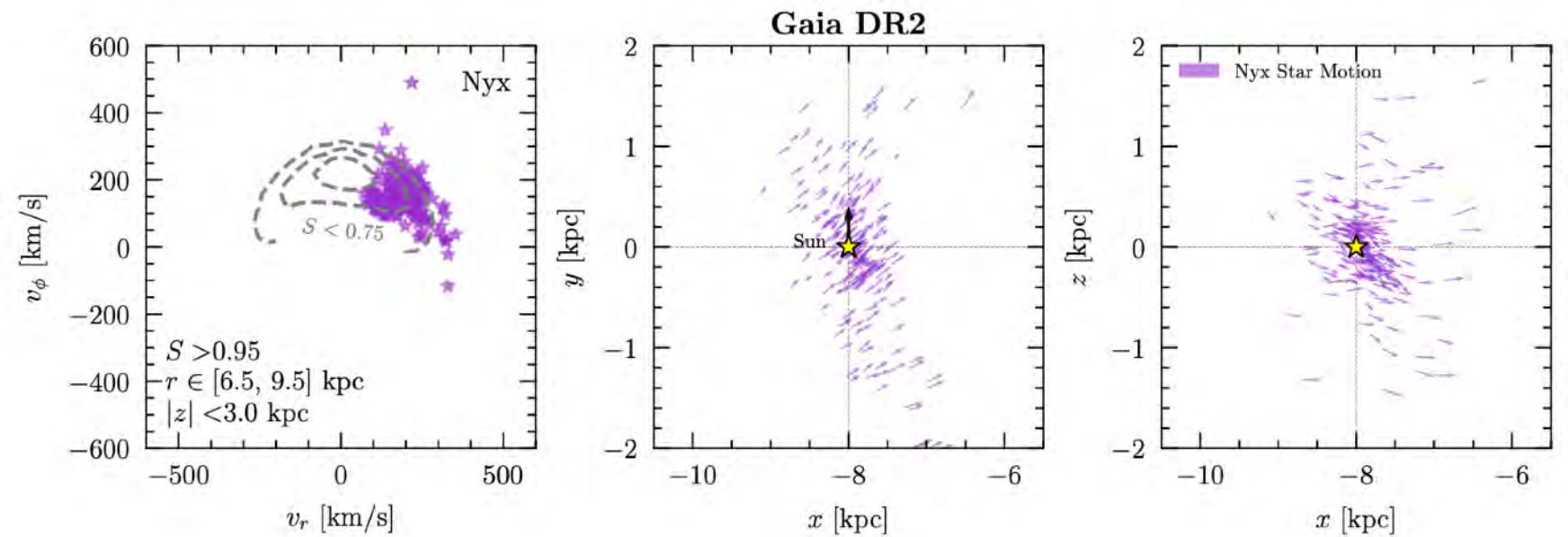
[Joe Dator, The New Yorker 2025; h/t Kyle Cranmer]

AI for Astrophysics: Galaxy Formation and Dark Matter

Scrutinizing stellar kinematics from Gaia to reconstruct history of the Milky Way



Discovery of Nyx!



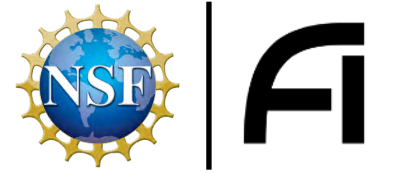
My former PhD student
(and current MIT colleague)



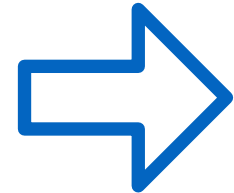
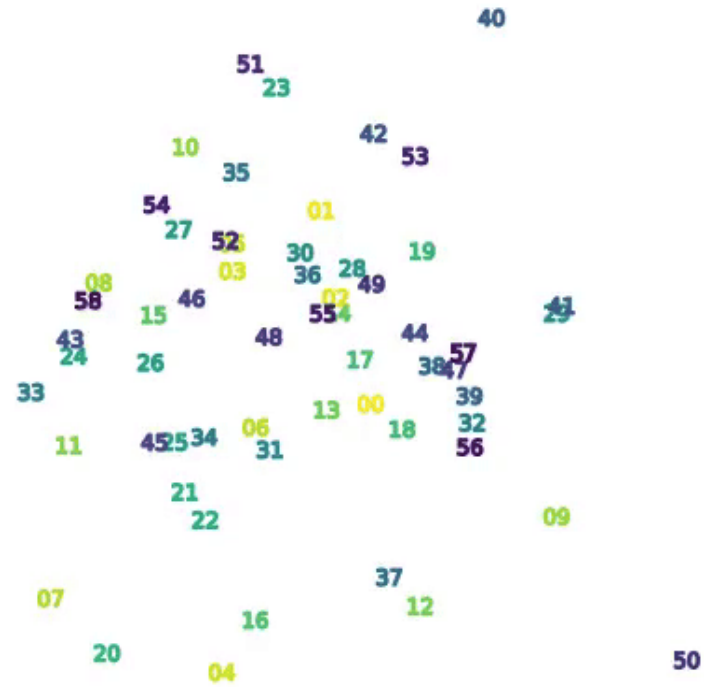
[Necib, et al., *Nature Astronomy* 2020; Ostdiek, et al., *A&A* 2020]

Physics for AI... and Back Again

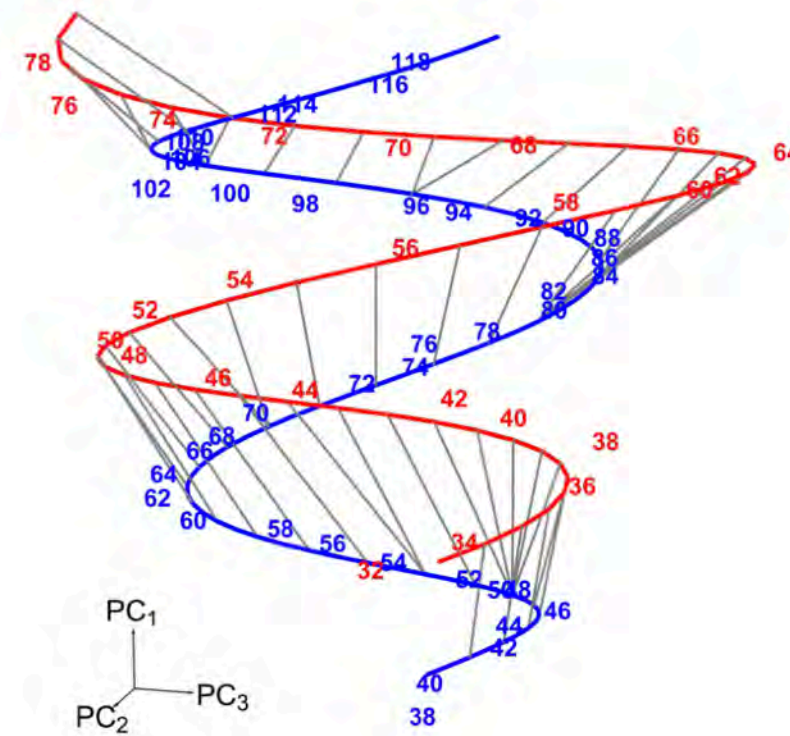
How do machines learn? What do they learn?



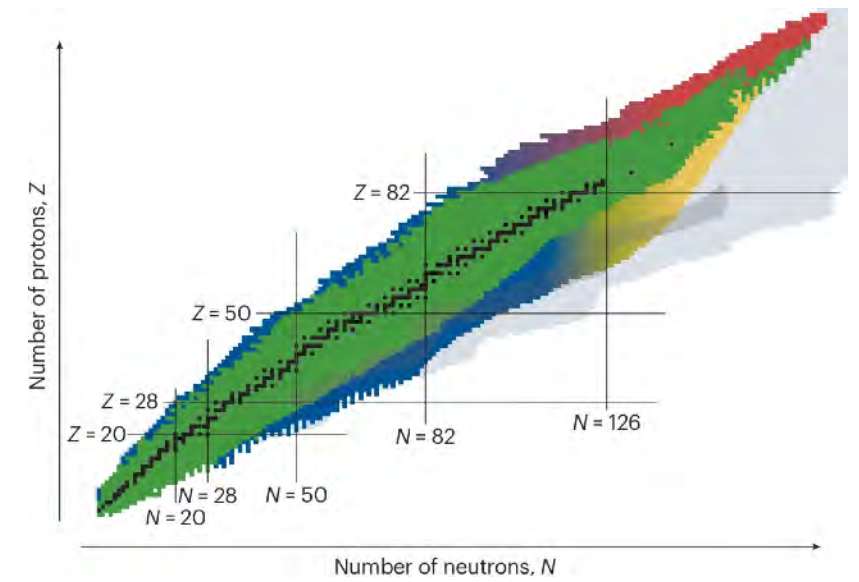
Loss: 4.29e+00|4.36e+00 Acc: 0.02|0.02



Machine-Learned Double Helix



cf. Nuclear Data Tables



Understanding “grokking”
(sudden learning) as phase transition

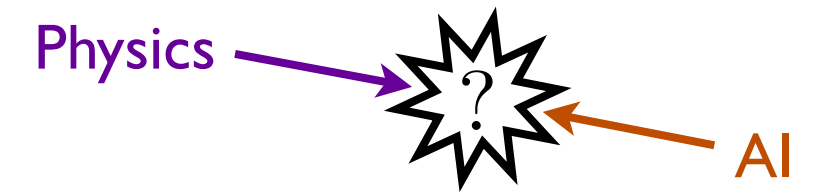
Predicting nuclear properties with high precision
(often better than specialized nuclear models)

IAIFI Interim Director

[Williams, Tegmark, Kitouni, Nolte, Michaud, Liu, Pérez-Díaz, Trifinopoulos, Kantamneni, Richardson, NeurIPS 2022 Spotlight Oral, ICML 2023, ICML 2024, arXiv 2025]

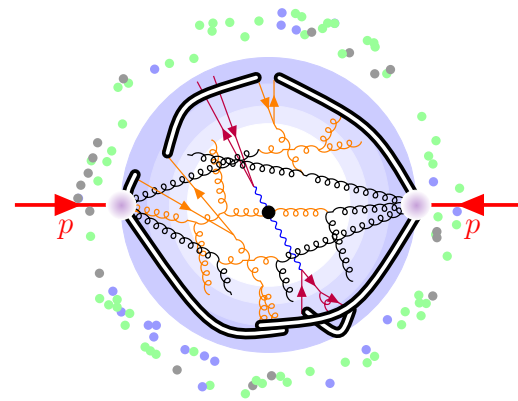


Centaur Science: Adventures in AI+Physics



Machine Learning through the Lens of Physics

*AI is transforming physics research, but many foundational aspects of AI can be translated into the **language of physics***



QCD Theory meets Information Theory

*To confront thorny challenges in theoretical physics, we can reframe them as well-defined **optimization problems***



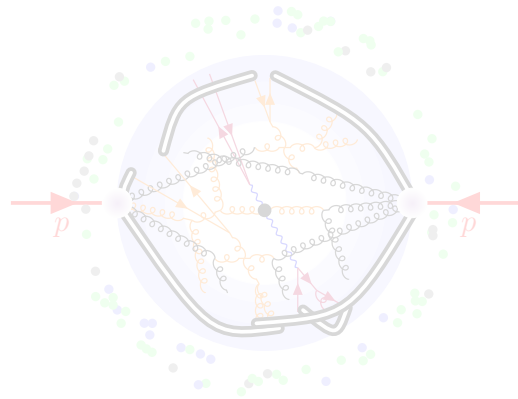
Scenes from my Sabbatical

*Being a “**centaur scientist**” means capitalizing on emerging AI technologies while insisting on scientific rigor and robustness*



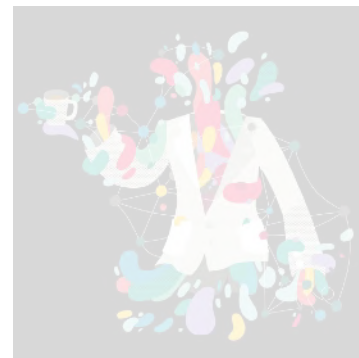
Machine Learning through the Lens of Physics

*AI is transforming physics research, but many foundational aspects of AI can be translated into the **language of physics***



QCD Theory meets Information Theory

*To confront thorny challenges in theoretical physics, we can reframe them as well-defined **optimization problems***



Scenes from my Sabbatical

*Being a “**centaur scientist**” means capitalizing on emerging AI technologies while insisting on scientific rigor and robustness*

Confronting AI/ML Hype vs. Reality

“AI is a **transformative technology**”

vs.

“ML is just **numerical optimization**”

*Both are true! And both are **changing the scientific enterprise***

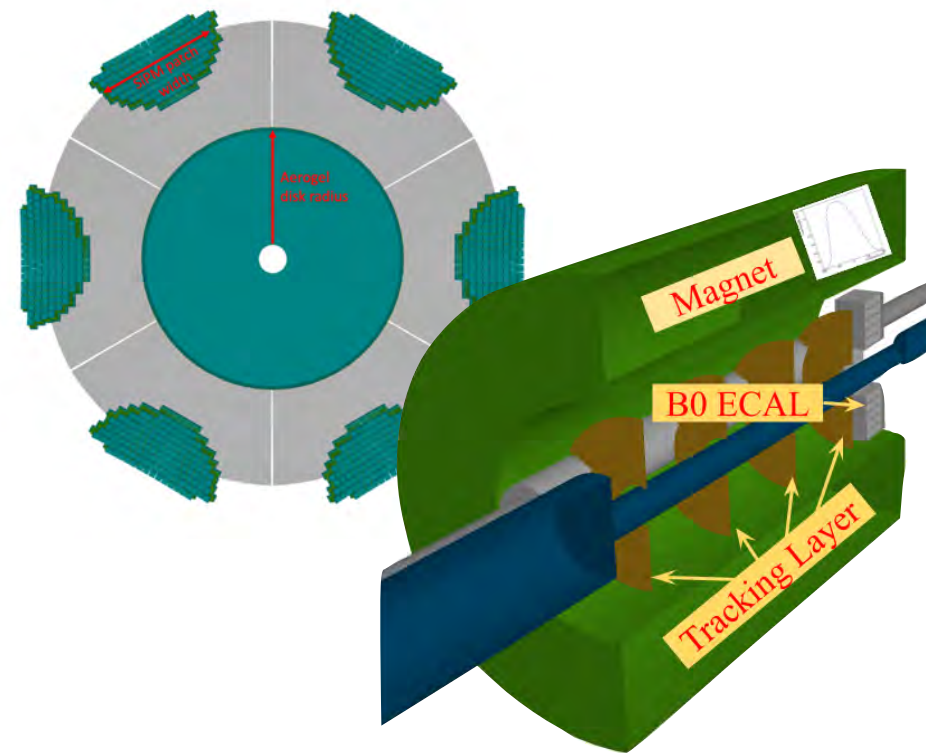
Direct AI/ML: Enables scientific investigations of **high-dimensional spaces**
e.g. simulation-based inference

Indirect AI/ML: Capitalizes on **emergent behaviors** in computational systems
e.g. foundation models, agentic AI

AI is Changing What it Means To...

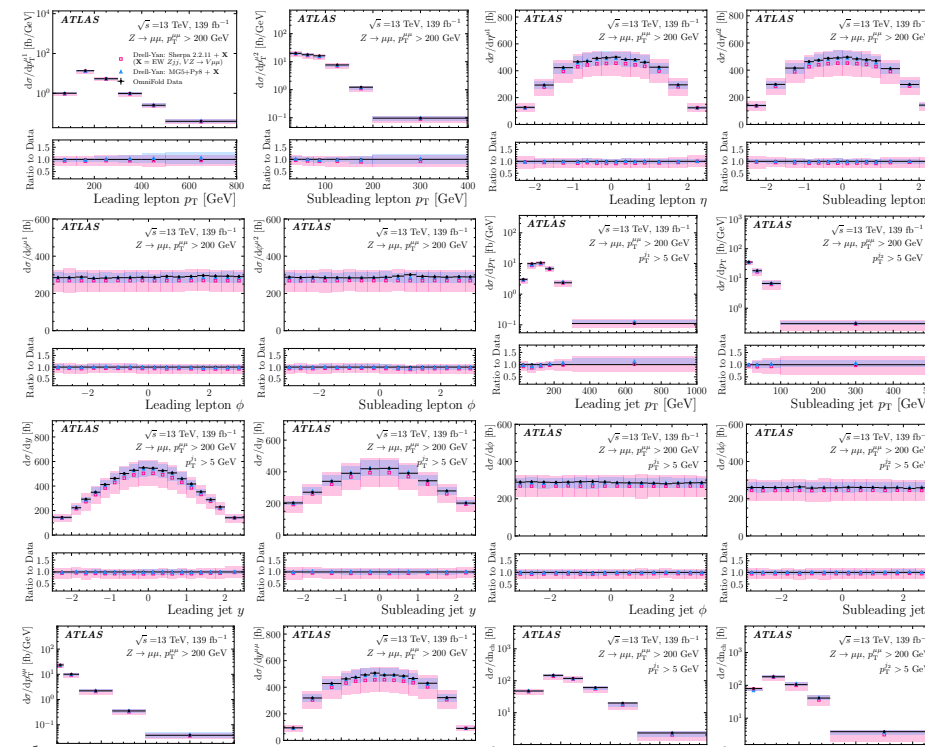
Representative examples,
far from exhaustive!

Design/Build/Operate Experiments



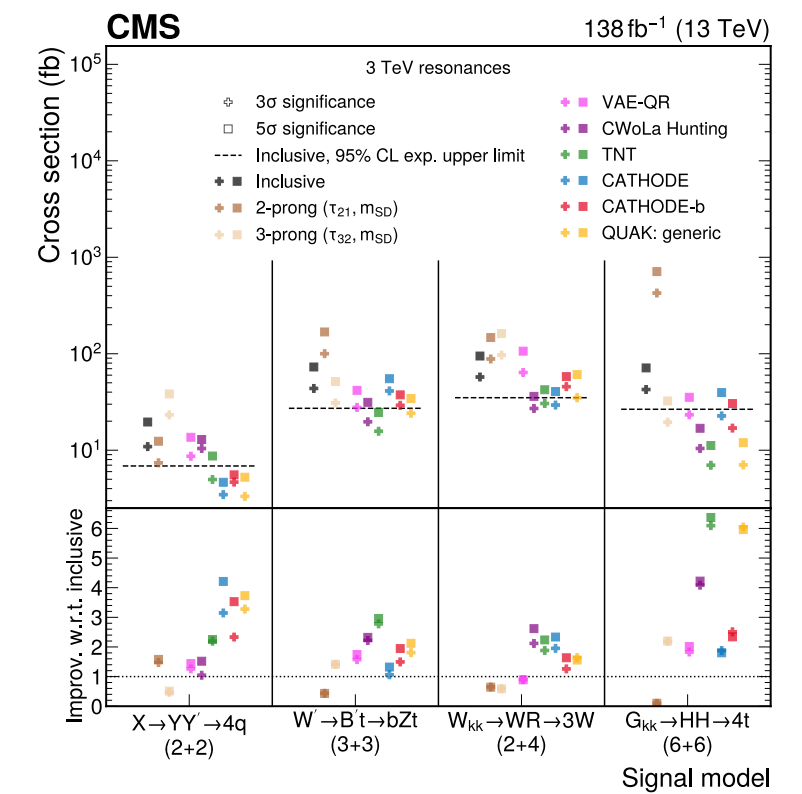
e.g. ePIC detector
optimization for EIC
[AID(2)E Collaboration, JINST 2024]

Perform/Report Measurements



e.g. ATLAS 24-dimensional
unboxed unfolding
[ATLAS, PRL 2024]

Search for New Physics



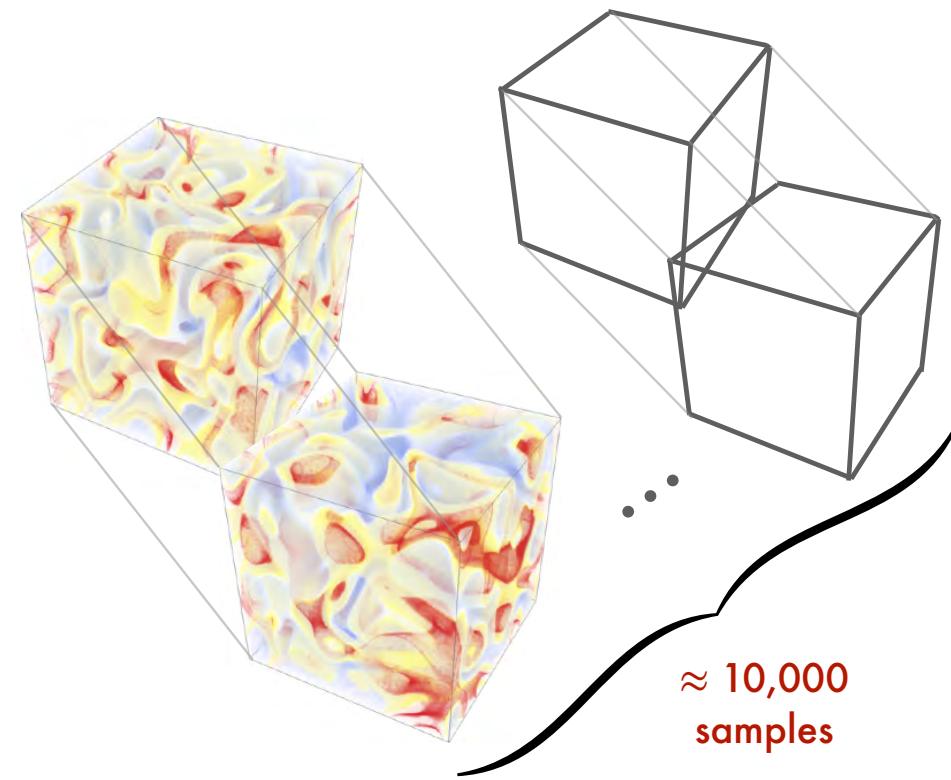
e.g. CMS anomalous dijet
resonance search
[CMS, RPP 2025]

[from my GGI lectures, January 2026; more examples at [HEP ML Living Review](#) and Ferguson, LaFleur, Ruthotto, JDT, Ting, Tiwary, Villar, et al., [MLST 2026](#)]

AI is Changing What it Means To...

Representative examples,
far from exhaustive!

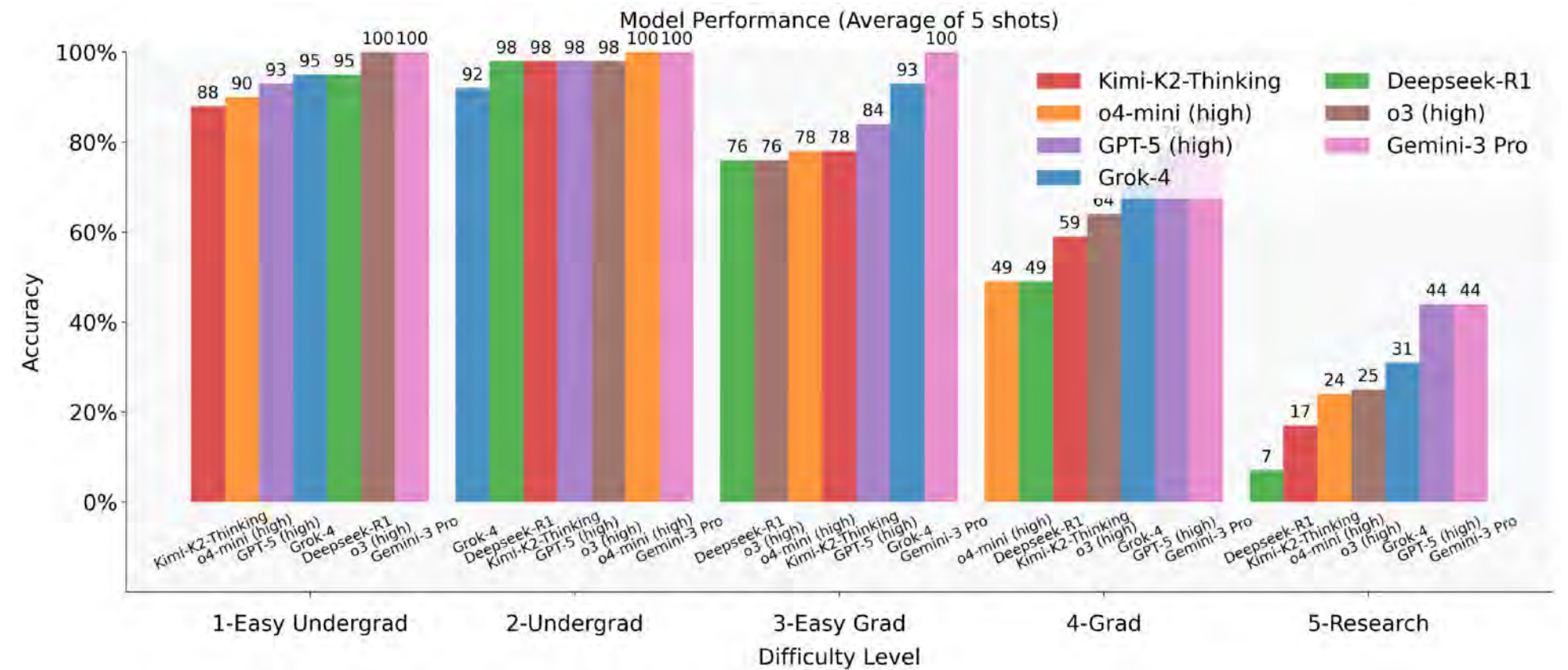
Make Theory Predictions



e.g. generative modeling
for lattice field theory

[Cranmer, Kanwar, Racanière, Rezende, Shanahan,
[Nature Reviews 2023](#)]

Be a Physicist



e.g. evaluating AI reasoning
abilities on TPBench

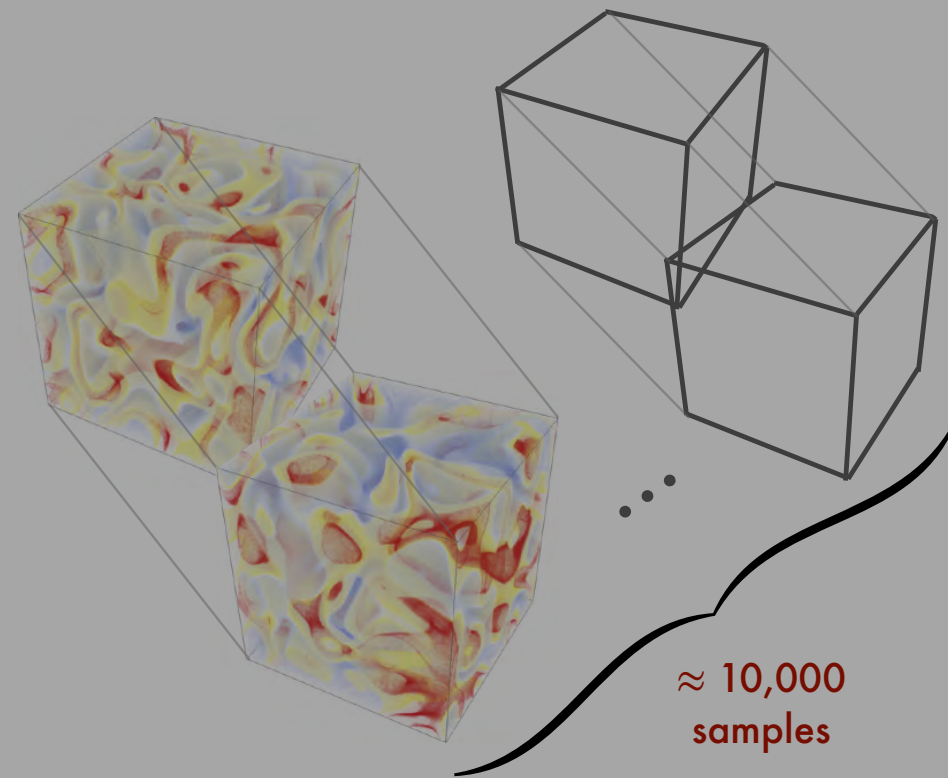
[Chung, Gao, Kvasiuk, Li, Münchmeyer, Rudolph, Sala, Tadepalli, [MLST 2025](#)]

[from my GGI lectures, [January 2026](#); more examples at [HEP ML Living Review](#) and Ferguson, LaFleur, Ruthotto, JDT, Ting, Tiwary, Villar, et al., [MLST 2026](#)]

AI is Changing What it Means To...

Representative examples,
far from exhaustive!

Make Theory Predictions



e.g. generative modeling
for lattice field theory

[Cranmer, Kanwar, Racanière, Rezende, Shanahan,
[Nature Reviews 2023](#)]

Be a Physicist



Resummation of the C-Parameter Sudakov Shoulder
Using Effective Field Theory

Matthew D. Schwartz^{1,2}

¹*Department of Physics, Harvard University, Cambridge, MA 02138, USA*

²*Institute for Artificial Intelligence and Fundamental Interactions (IAIFI)*

schwartz@g.harvard.edu

AI RESEARCH ASSISTANT: Claude Opus 4.5 (Anthropic)

January 7, 2026

[Schwartz, [arXiv 2026](#); see also Shih, [arXiv 2026a](#), [arXiv 2026b](#)]

[from my GGI lectures, [January 2026](#); more examples at [HEP ML Living Review](#) and Ferguson, LaFleur, Ruthotto, JDT, Ting, Tiwary, Villar, et al., [MLST 2026](#)]

ML is “Just” Optimization in Response to Data

*Surprisingly powerful,
both directly and indirectly!*

ML Framing

Physics Translation

Training Data

$$\frac{1}{N} \sum_{i=1}^N (\dots) = \int d\Phi (\dots) + \mathcal{O}\left(\frac{1}{\sqrt{N}}\right)$$

Monte Carlo
Integration

Loss Function(al)

$$\mathcal{L}_i[f(\Phi)] \Leftrightarrow \frac{\delta \mathcal{L}}{\delta f} = 0$$

Lagrangian Mechanics
e.g. Euler-Lagrange equation

Learnable Function
e.g. neural networks

$$f_{\text{NN}}(\Phi; b) \Rightarrow f_{\text{physics}}(\Phi; b)$$

Physics Knowledge
e.g. symmetries

Optimizer
e.g. gradient descent

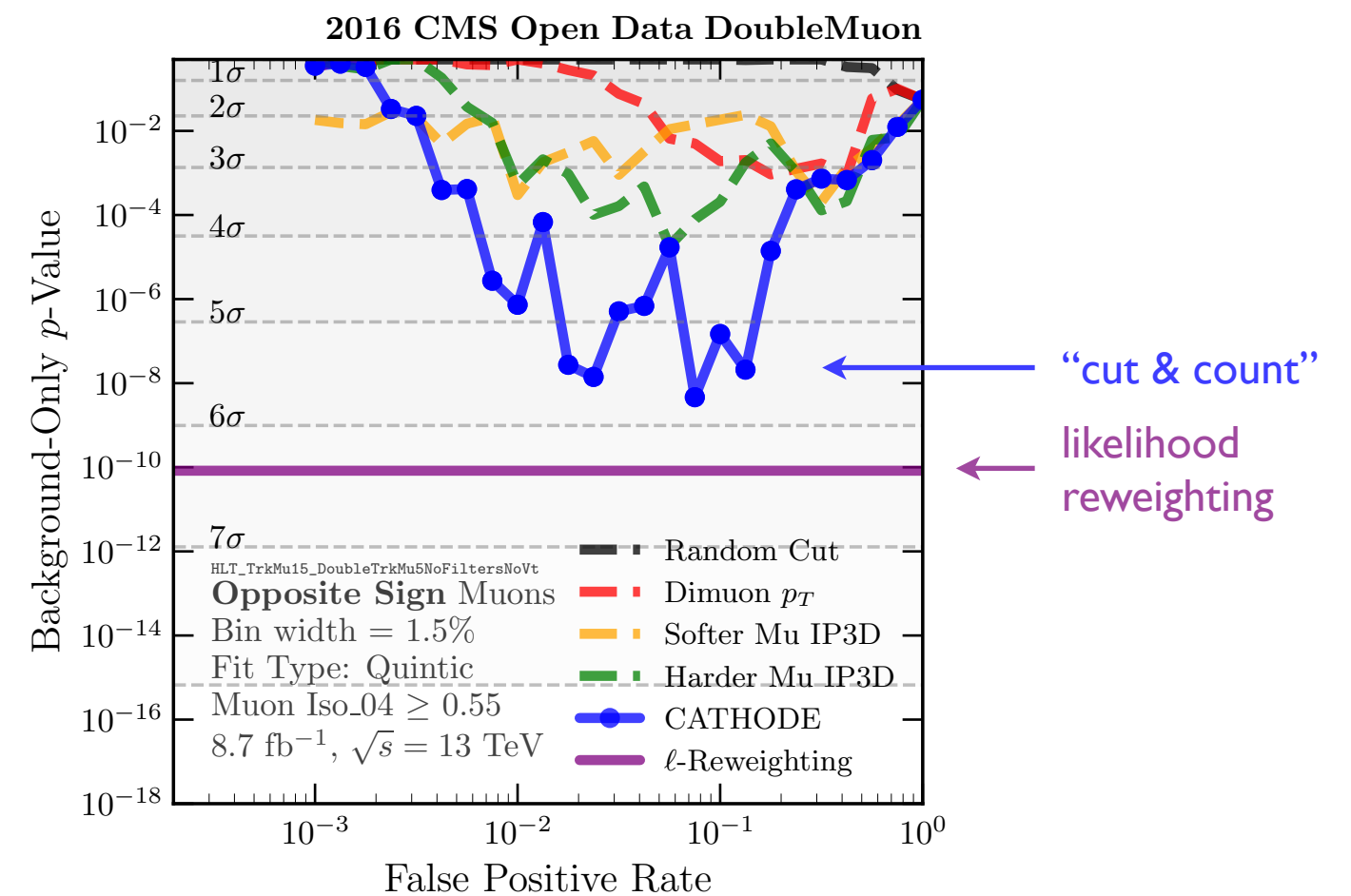
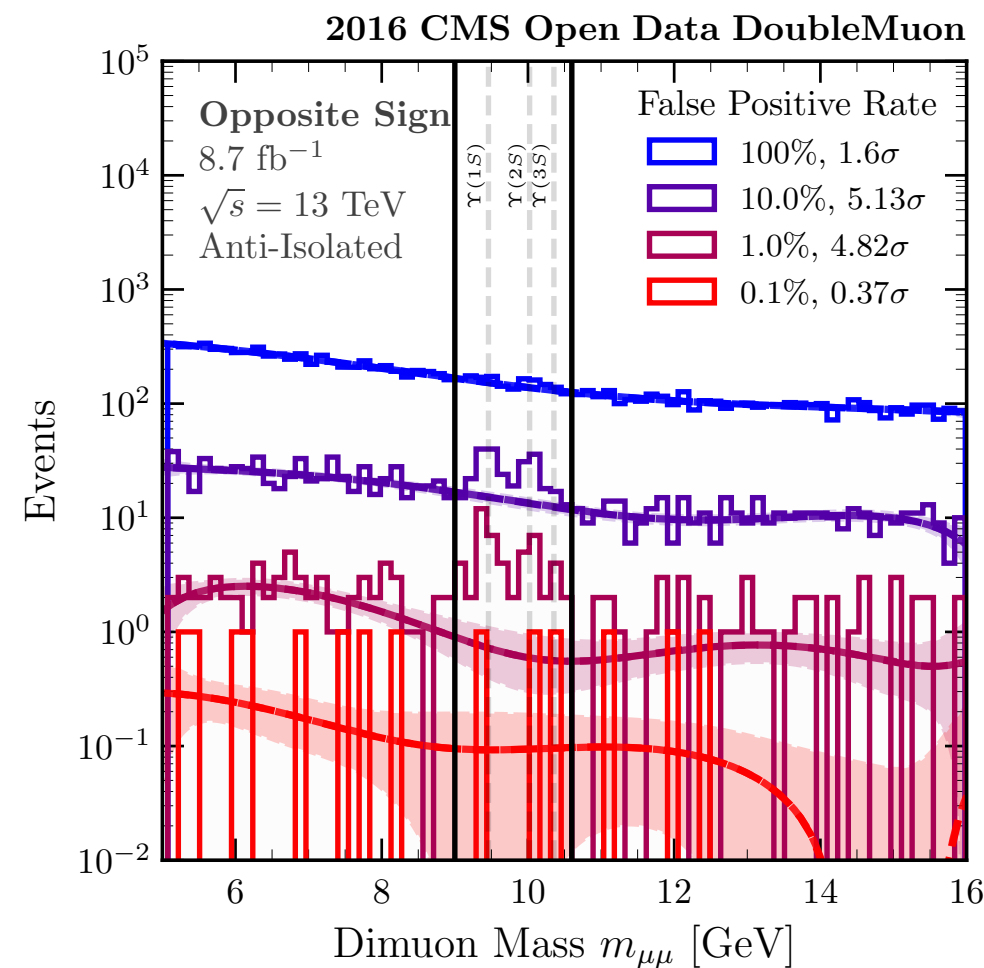
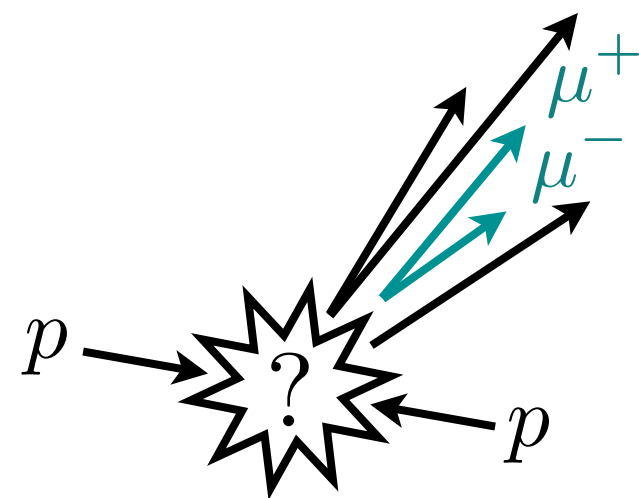
$$\Delta b = -\eta \nabla_b \mathcal{L} \Leftrightarrow m \ddot{b} + \gamma \dot{b} + \frac{dV}{db} = 0$$

Newtonian Dynamics
e.g. overdamped oscillator

[from my GGI lectures, [January 2026](#)]

Direct AI/ML: Rediscovering Upsilon with Anomaly Detection

Background predictions from **normalizing flows**, signal classification through **boosted decision trees**...

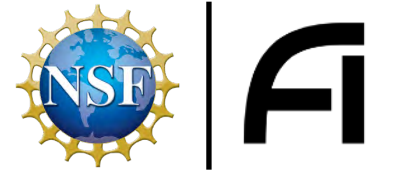


...with statistical significance estimated using **physics domain knowledge** about localized resonances

[Gambhir, Mastandrea, Nachman, JDT, [PRL 2025](#);
 using CATHODE in Hallin, Isaacson, Kasieczka, Krause, Nachman, Quadfasel, Schlaffer, Shih, Sommerhalder, [PRD 2022](#)]



Indirect AI/ML: Contrastive Learning for Hubble Data



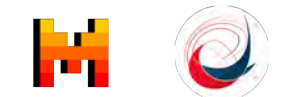
CLIP fine-tuning

With Hubble *observation-proposal abstract pairs*

Hubble proposal abstracts

Category: Galaxies. We propose WFC3/UVIS F336W, F438W, and F814W observations for 8 Luminous Infrared Galaxies (LIRGs) in the Great Observatories All-Sky LIRG Survey (GOALS) scheduled for JWST Cycle 1 (G01) observations. With a proprietary period of 0 days for 50% of the G01 LIRGs, observations taken now will provide the concurrent WFC3/UVIS imaging necessary to reliably age-date the star ...

Mixtral + Outlines



Optional summarization with constrained LLM generation

Luminous Infrared Galaxies, star clusters, nuclear regions, extranuclear regions, hydrogen recombination lines; measure fraction of star formation in clusters, determine nuclear and extranuclear cluster destruction rates, ...

Text encoder

Hubble observations

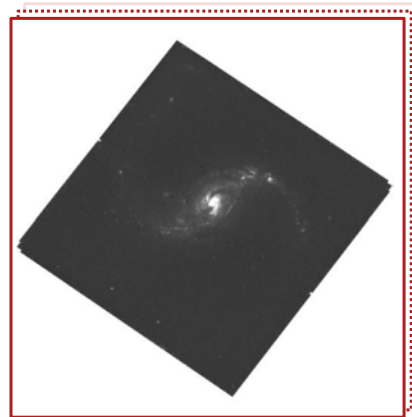
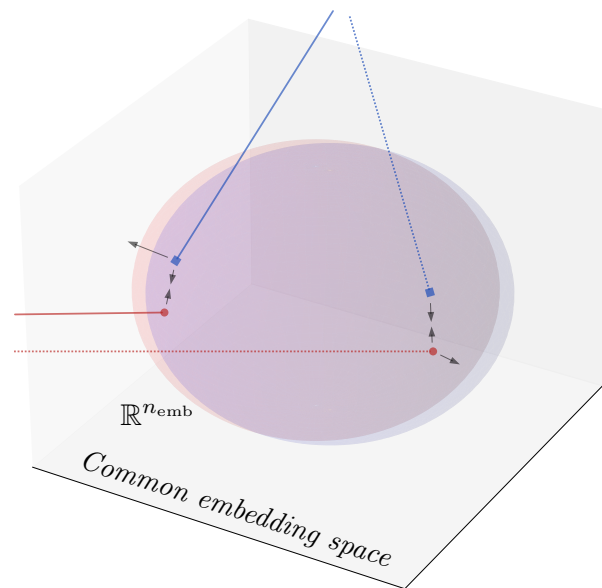


Image encoder

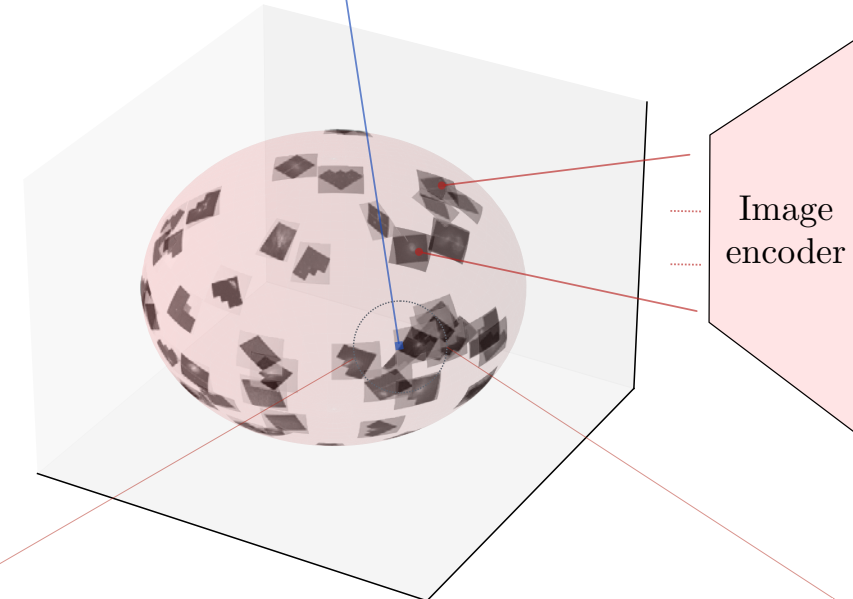


Downstream task: observation retrieval

Given natural language text query

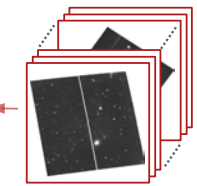
Query: "barred spiral galaxy"

Text encoder

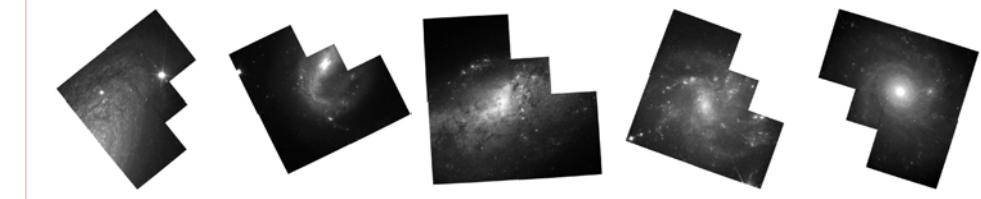


Candidate observations

Image encoder



Closest observations



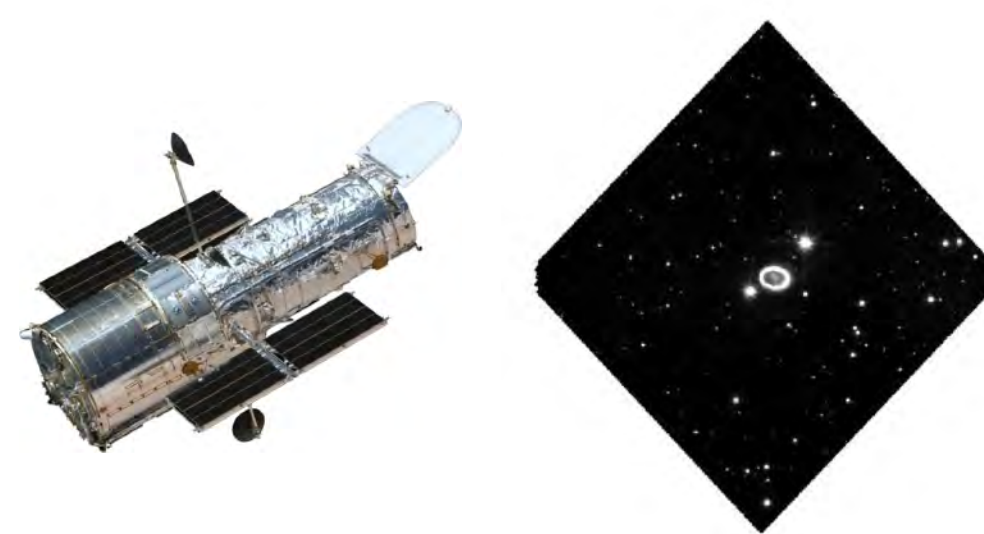
[PAPERCLIP in Mishra-Sharma, Song, JDT, COLM 2024; based on CLIP in Radford, Kim, Hallacy, Ramesh, Goh, Agarwal, Sastry, Askell, Mishkin, Clark, Krueger, Sutskever, PMLR 2021; see also AstroCLIP in Parker, Lanusse, Golkar, Sarra, Cranmer, Bietti, Eickenberg, Krawezik, McCabe, Ohana, Pettee, Regalado-Saint Blancard, Tesileanu, Cho, Ho, MNRAS 2024]



Indirect AI/ML: Contrastive Learning for Hubble Data



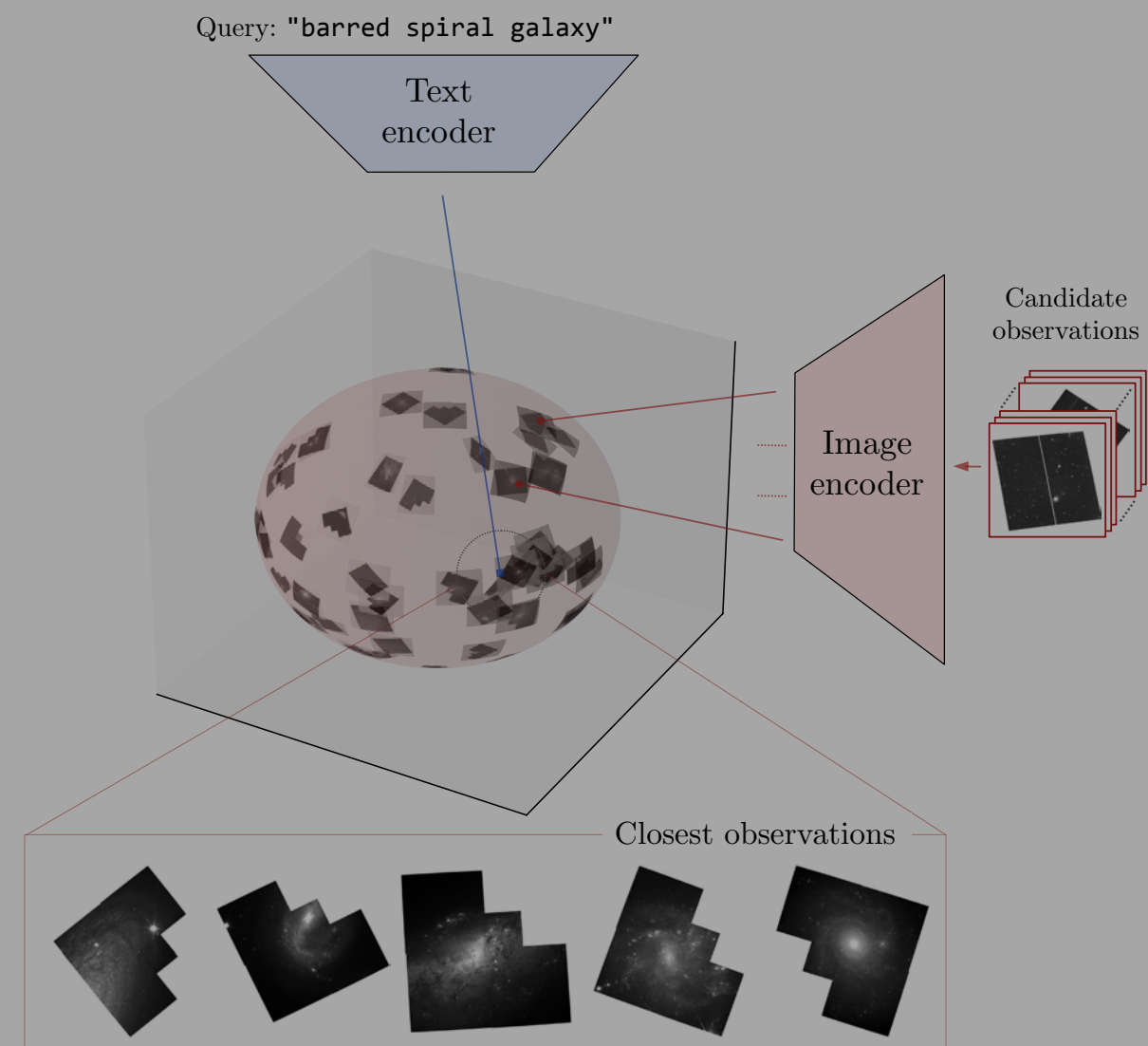
“What is this a picture of?”



CLIP-ViT-B/16: “gravitational lens”
Fine-Tuned: “supernova remnant”

[image from Supernova 1987A INTensive Survey, [HST Proposal 11653](#)]

Downstream task: observation retrieval *Given natural language text query*



[PAPERCLIP in Mishra-Sharma, Song, JDT, [COLM 2024](#);
based on CLIP in Radford, Kim, Hallacy, Ramesh, Goh, Agarwal, Sastry, Askell, Mishkin, Clark, Krueger, Sutskever, PMLR 2021;
see also AstroCLIP in Parker, Lanusse, Golkar, Sarra, Cranmer, Bietti, Eickenberg, Krawezik, McCabe, Ohana, Pettee, Regalado-Saint Blancard, Tesileanu, Cho, Ho, [MNRAS 2024](#)]



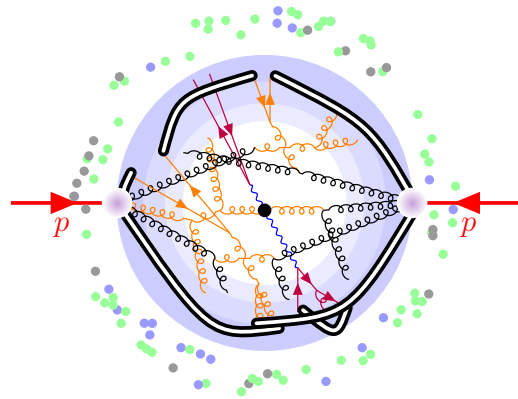


*The rise of **centaur scientists** who synthesize
physics intelligence and **artificial intelligence***



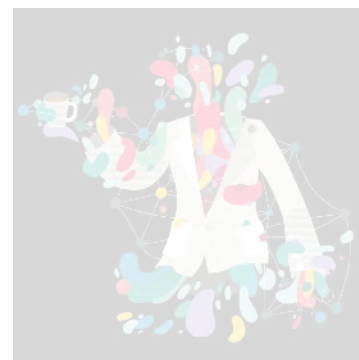
Machine Learning through the Lens of Physics

*AI is transforming physics research, but many foundational aspects of AI can be translated into the **language of physics***



QCD Theory meets Information Theory

*To confront thorny challenges in theoretical physics, we can reframe them as well-defined **optimization problems***



Scenes from my Sabbatical

*Being a “**centaur scientist**” means capitalizing on emerging AI technologies while insisting on scientific rigor and robustness*

Danger of “Thinking Like a Machine”?

An inspiring/infuriating paper!

***Position:* Is machine learning good or bad for the natural sciences?**

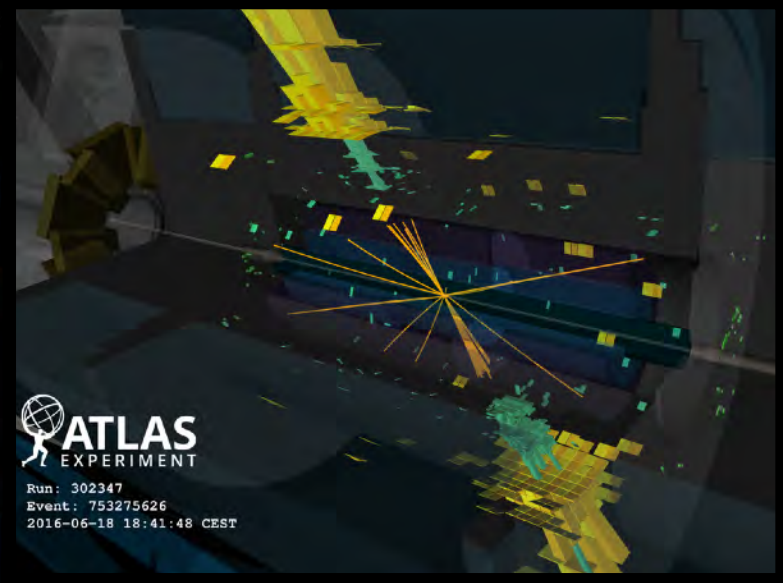
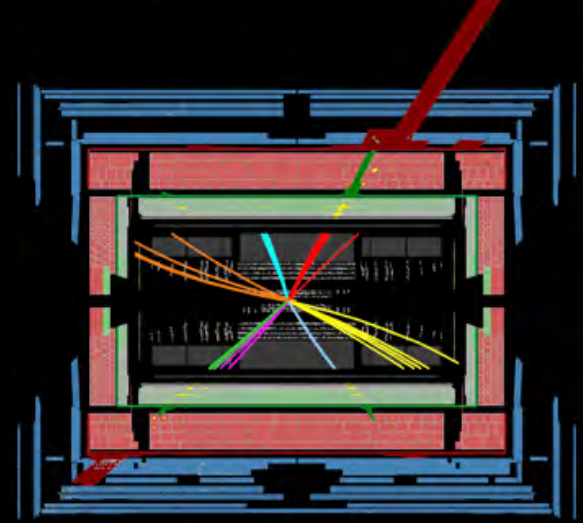
Abstract

Machine learning (ML) methods are having a huge impact across all of the sciences. However, ML has a strong ontology—in which only the data exist—and a strong epistemology—in which a model is considered good if it performs well on held-out training data. These philosophies are in strong conflict with both standard practices and key philosophies in the natural sciences. . . .

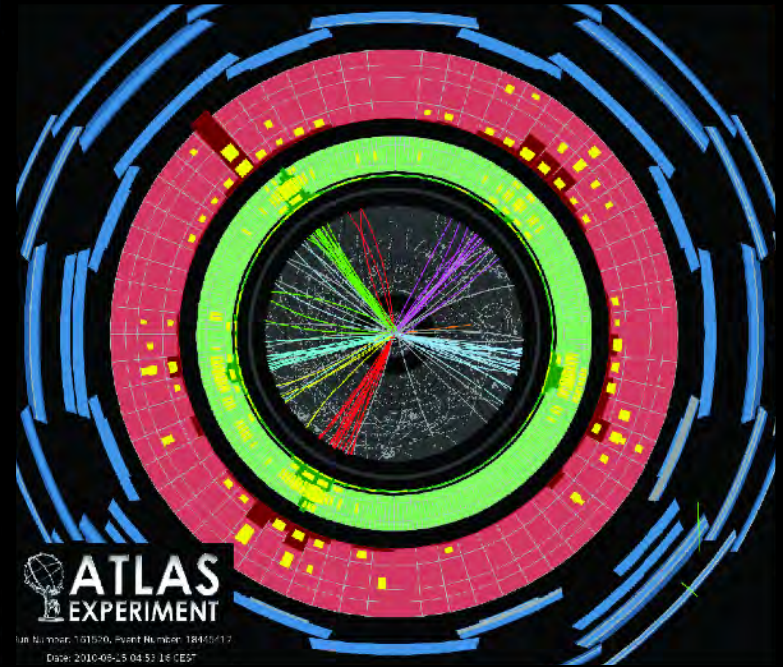
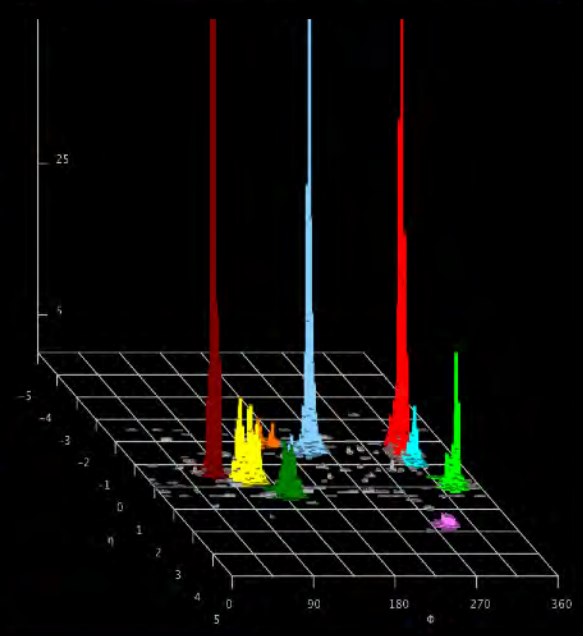
*My attempt at a
cartoon summary:*

	Machine Learning	Natural Sciences
<i>What “exists”?</i>	Data	Latent Structures
<i>What is “success”?</i>	Accurate Modeling	Understanding

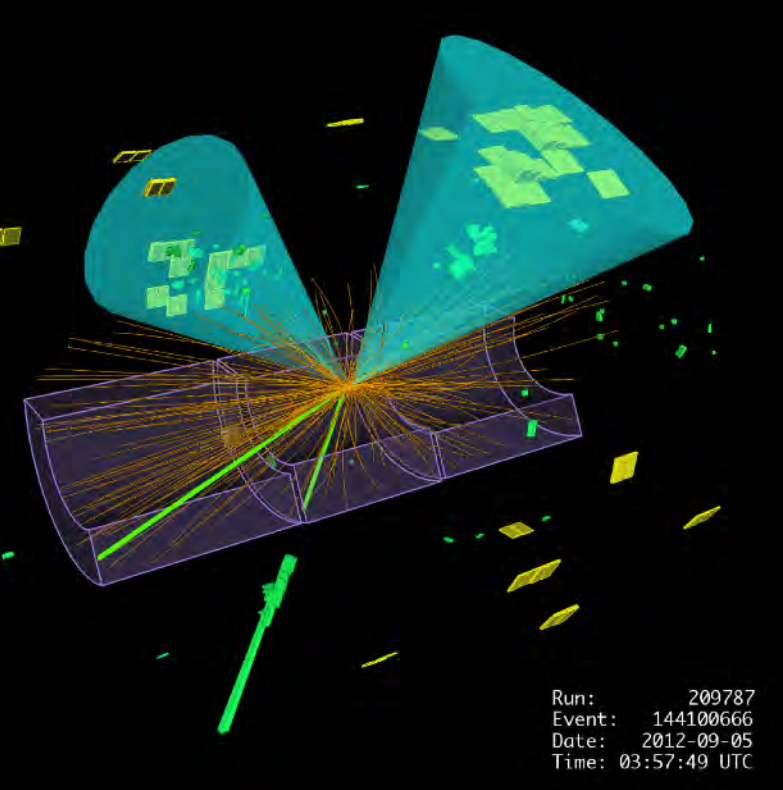
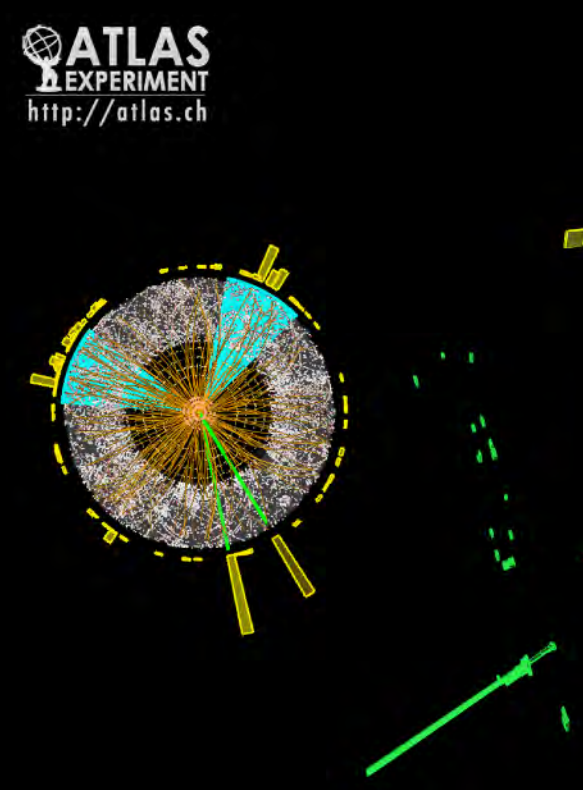
[Hogg, Villar, [PMLR 2024](#)]



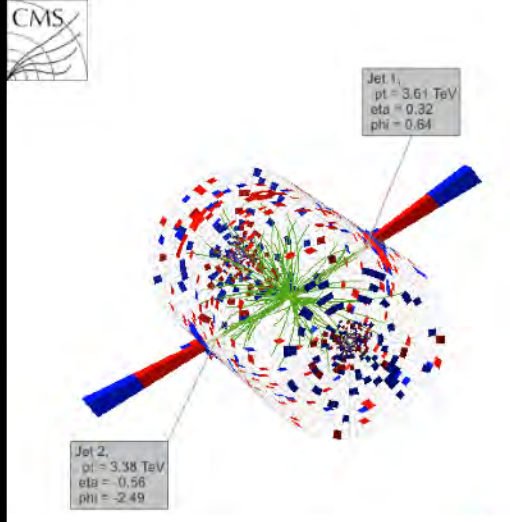
ATLAS EXPERIMENT
Run: 302347
Event: 753275626
2016-06-18 18:41:48 CEST



ATLAS EXPERIMENT
Run Number: 103100, Event Number: 18415117
Date: 2010-05-15 04:53:14 CEST



Run: 209787
Event: 144100666
Date: 2012-09-05
Time: 03:57:49 UTC

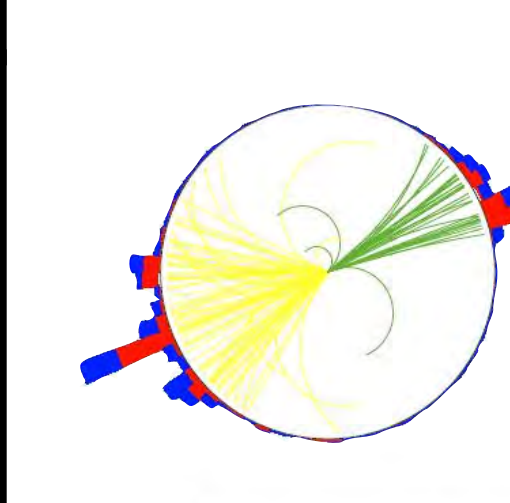


Jet 1,
pt = 3.61 TeV
eta = 0.32
phi = 0.64

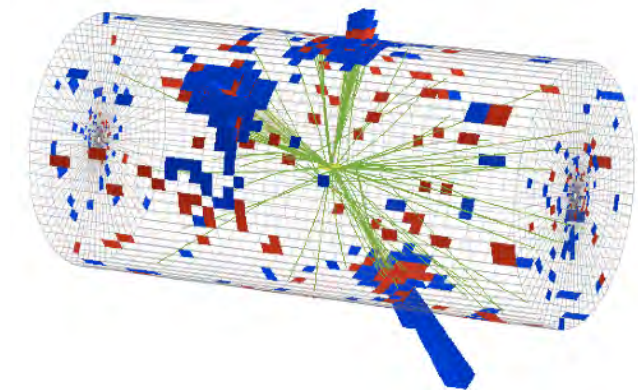
Jet 2,
pt = 3.38 TeV
eta = -0.56
phi = -2.49

CMS Experiment at LHC, CERN
Data recorded: Sat Aug 28 23:03:34 2010 EDT
Run/Event: 144899 / 67178071
Lumi section: 574

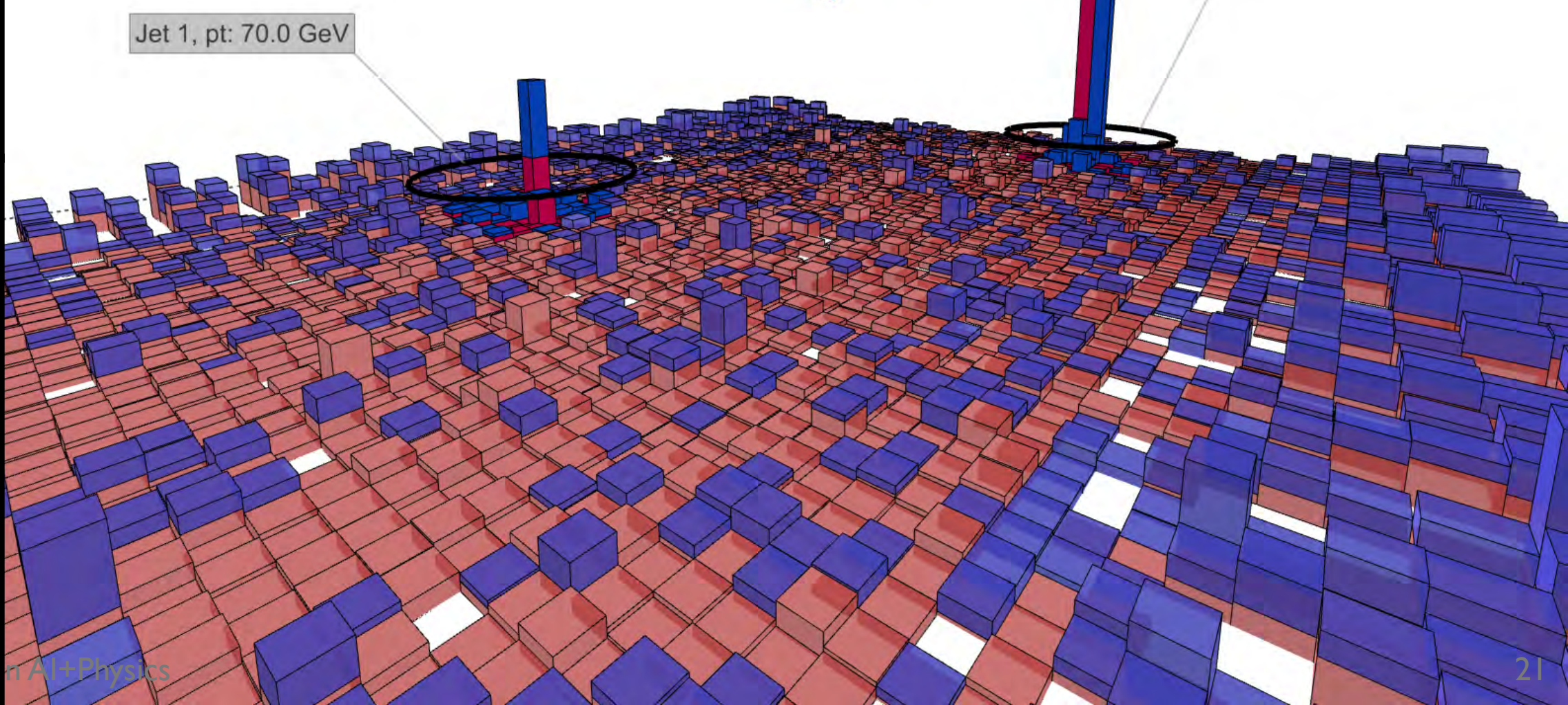
CMS



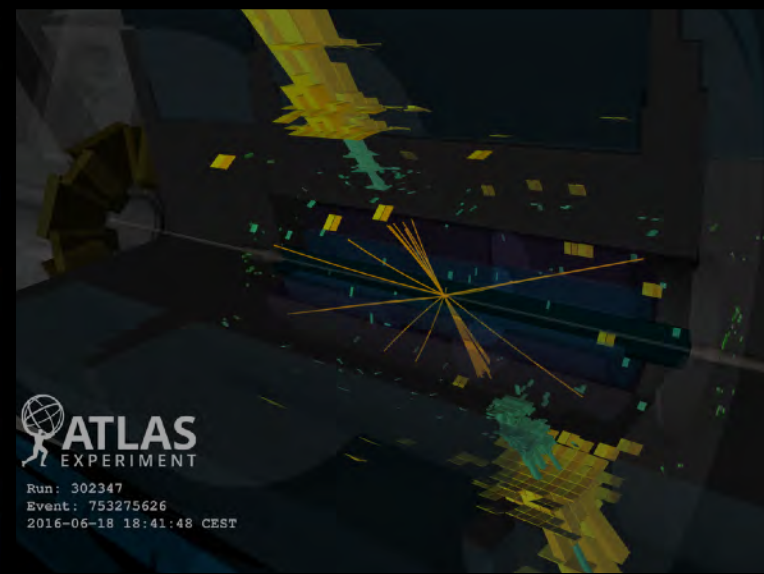
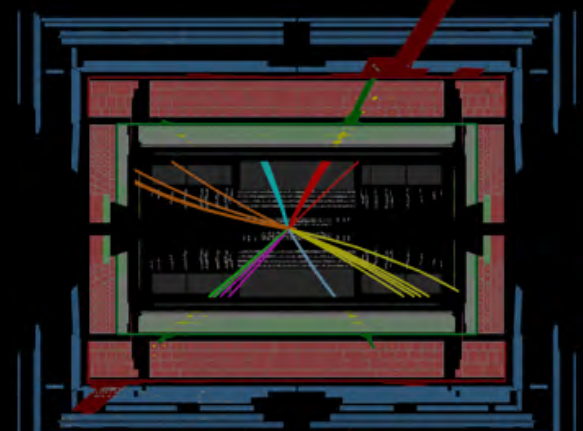
Jet 1, pt: 70.0 GeV



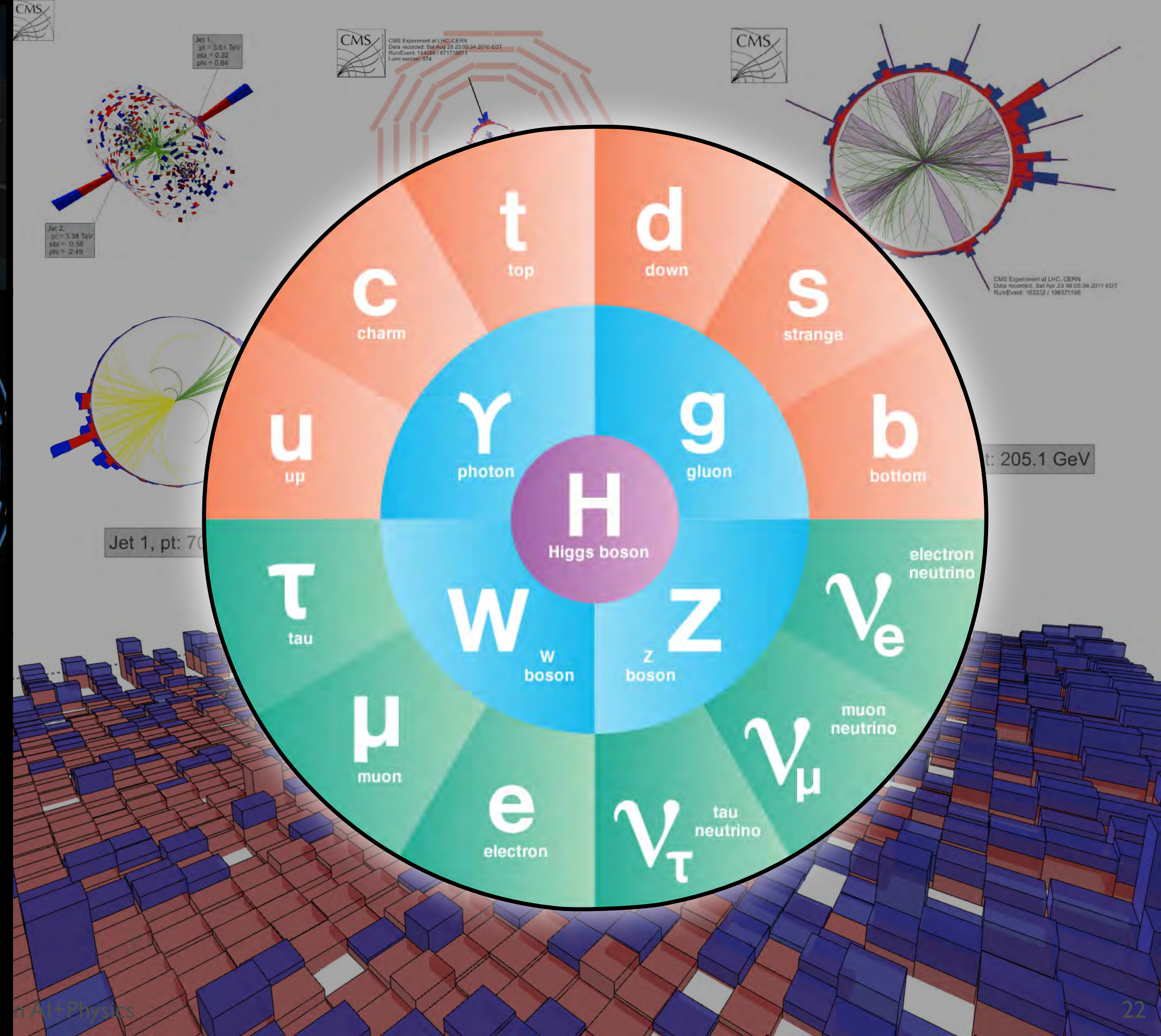
Jet 0, pt: 205.1 GeV



CMS Experiment at LHC, CERN
Data recorded: Sat Apr 23 08:05:38 2011 EDT
Run/Event: 163332 / 19637106



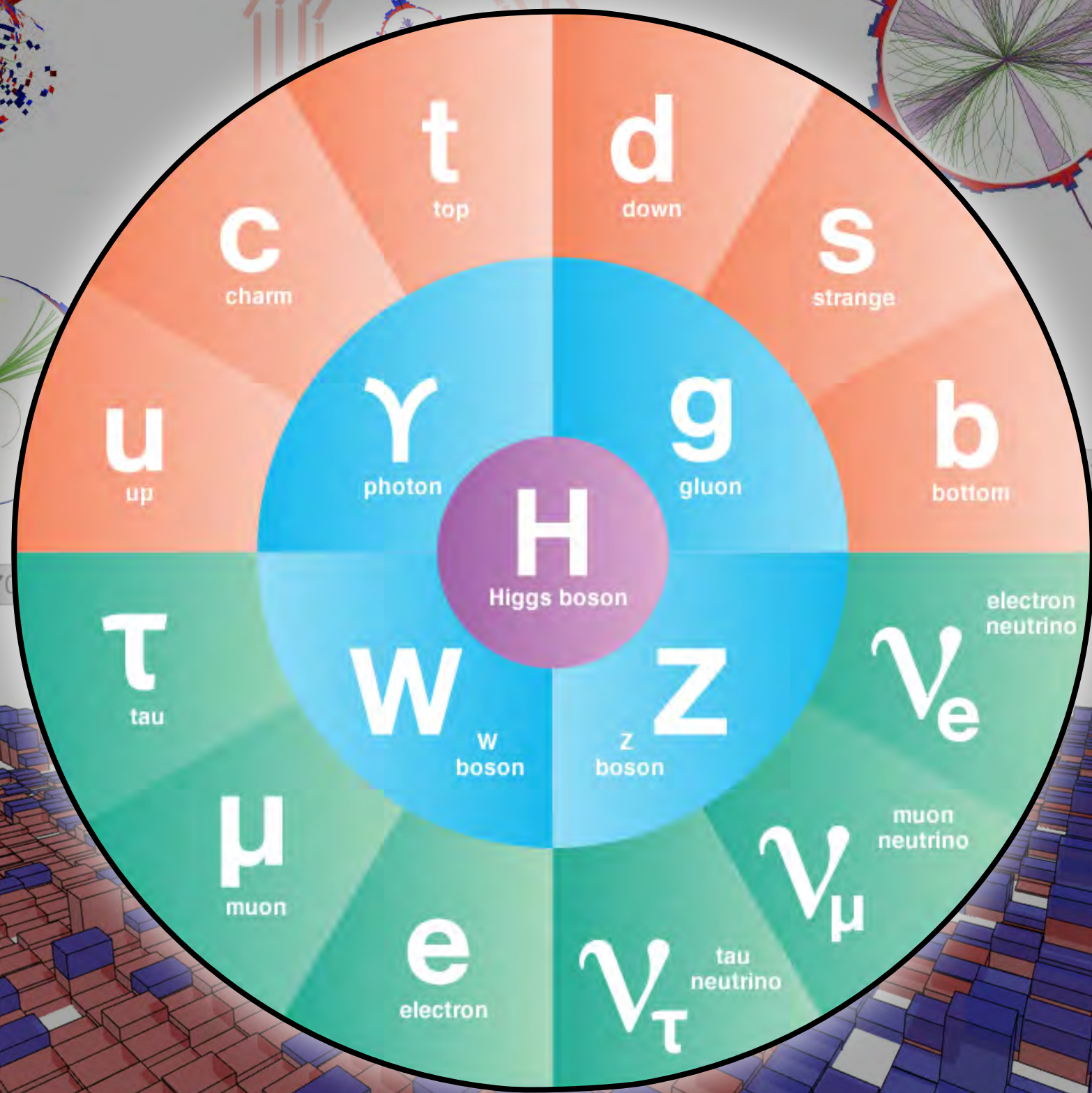
ATLAS EXPERIMENT
Run: 302347
Event: 753275626
2016-06-18 18:41:48 CEST


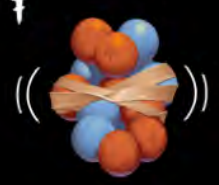


CMS Experiment at LHC, CERN
Data recorded: Sat Aug 28 23:03:34 2010 EDT
Run/Event: 144895 / 67178071
Lumi sector: 574

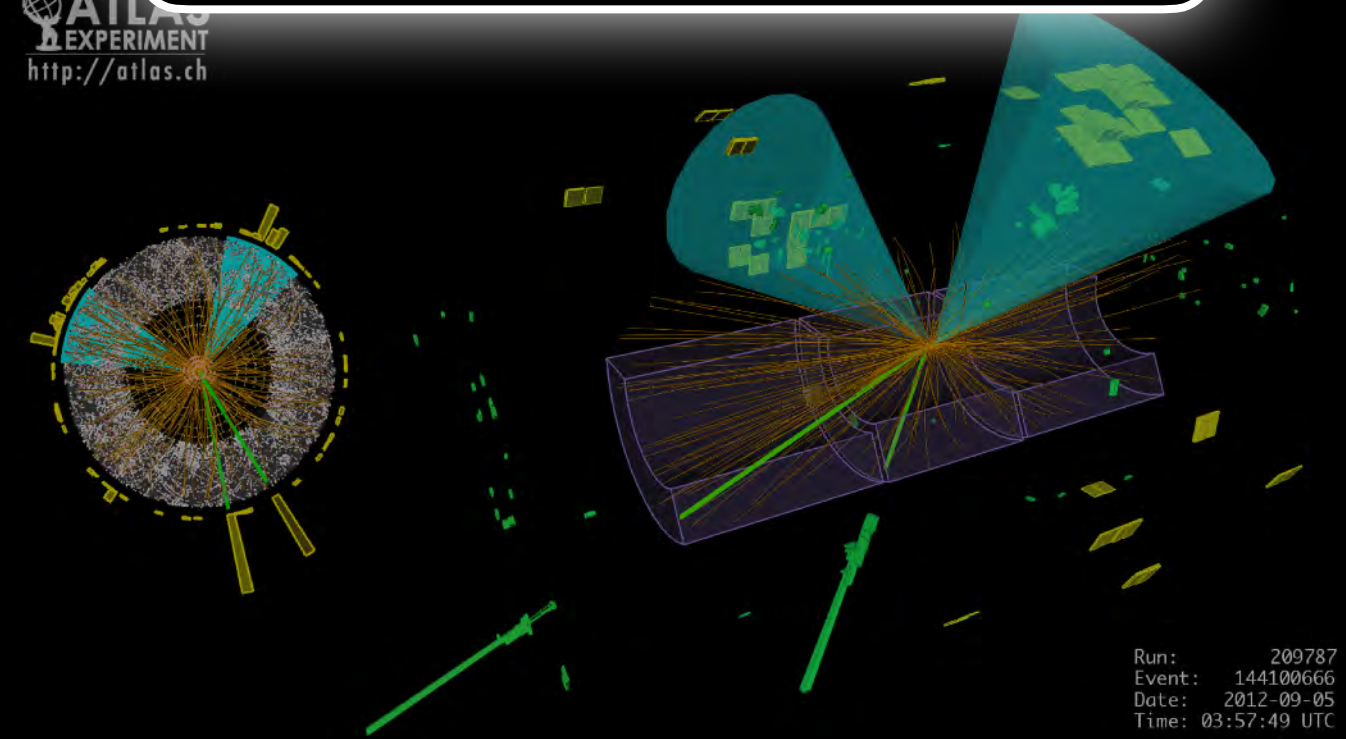
CMS

CMS Experiment at LHC, CERN
Data recorded: Sat Apr 23 08:05:38 2011 EDT
Run/Event: 163332 / 19637106



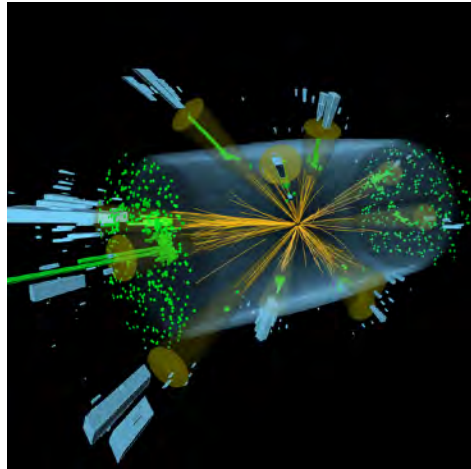
 Strong

Jets: Manifestation of underlying Quarks & Gluons
 Theory: Quantum Chromodynamics (QCD) in the Standard Model of Particle Physics

ATLAS EXPERIMENT
<http://atlas.ch>



Run: 209787
Event: 144100666
Date: 2012-09-05
Time: 03:57:49 UTC

A Generative Model to Mimic the LHC?



What if you could generate **synthetic LHC collisions** that were indistinguishable from **real LHC data**?

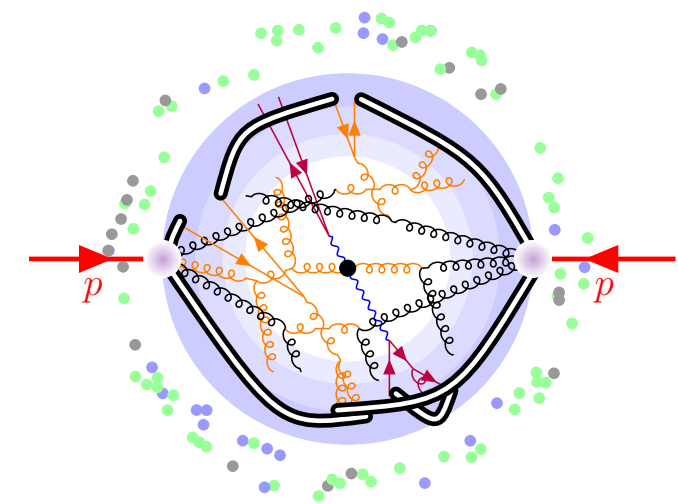
[extreme extrapolation of Butter, Diefenbacher, Kasieczka, Nachman, Plehn, Shih, Winterhalder, [SciPost 2022](#)]

According to Hogg/Villar, this conflicts with the epistemology of natural science...

What if you could generate **synthetic LHC collisions** based entirely on **first-principles QFT calculations**?

[aspirational goal in Campbell, Diefenthaler, Hobbs, Hoeche, Isaacson, Kling, Mrenna, Reuter, et al., [Snowmass 2022](#)]

Amazing if you had it! But we are nowhere near being able to do this...



Centaur Science:

Combine **models of LHC data** for the full phase space with **first-principles predictions** whenever available

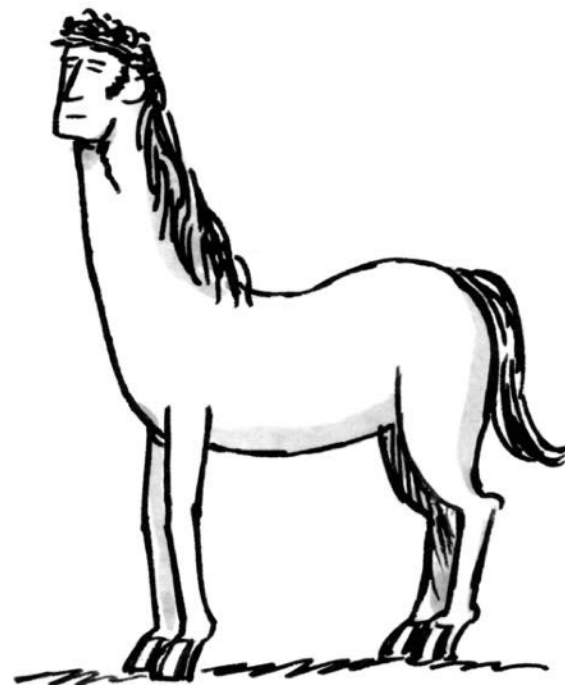
Parton Shower + Matrix Element through Optimization?



General purpose event generator
that can make arbitrary LHC
predictions but at low accuracy

(could be a parton shower based on QCD... or an AI chatbot)

Specialized QCD calculations
that can make high-accuracy LHC
predictions for specific observables



General purpose event generator
for the LHC that will...



...provably reproduce any
specialized QCD calculation?

Parton Shower + Matrix Element through Optimization?

Training Data:

Precision QCD Calculations (*in what form?*)

Loss Function(al):

Information-Theoretic Measure of Similarity

Learnable Function:

Custom Boltzmann Factor

Optimizer:

Efficient Non-Linear Solver (L-BFGS)

General purpose event generator
for the LHC that will...



...provably reproduce any
specialized QCD calculation?

Result of Information-Theoretic Optimization

See backup slides
for derivation

Boltzmann Factor: $p \propto e^{-\beta E}$ Maximum entropy solution subject to **constraint** on **average energy**

Generalization: $p(x) = q(x) \exp \left[-\beta_0 - \beta_1 f_1(x) - \beta_2 f_2(x) - \dots \right]$

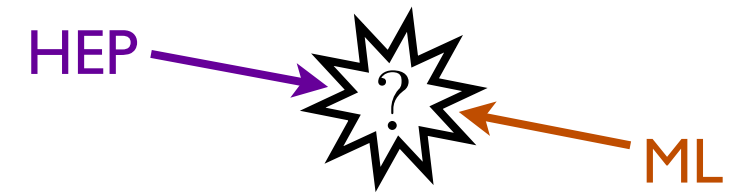
Prior \downarrow Sets Partition Function \downarrow

Lagrange Multiplier \uparrow Constrained Quantity \uparrow

If you can perform precision QCD calculations of **moment constraints**,
then you can improve **prior predictions** with learned **Lagrange multipliers**

[Boltzmann 1877]

QCD Theory meets Information Theory



Plucking a random
QCD calculation:

[e.g. Larkoski, Salam, JDT, JHEP 2013;
based on Banfi, Salam, Zanderighi, JHEP 2005]

$$\Sigma(e^{-L}) = N \frac{e^{-\gamma_E R'}}{\Gamma(1 + R')} e^{-R} \quad R = \frac{\alpha_s}{\pi} \frac{C}{\beta} (L + B)^2 \quad L \equiv \ln \frac{R_0^\beta}{C_1^{(\beta)}}$$

↑
Sudakov Form Factor

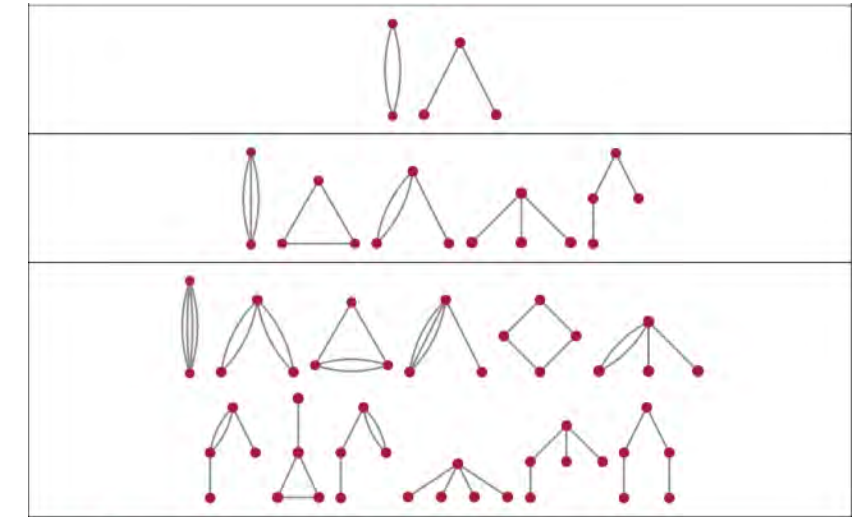
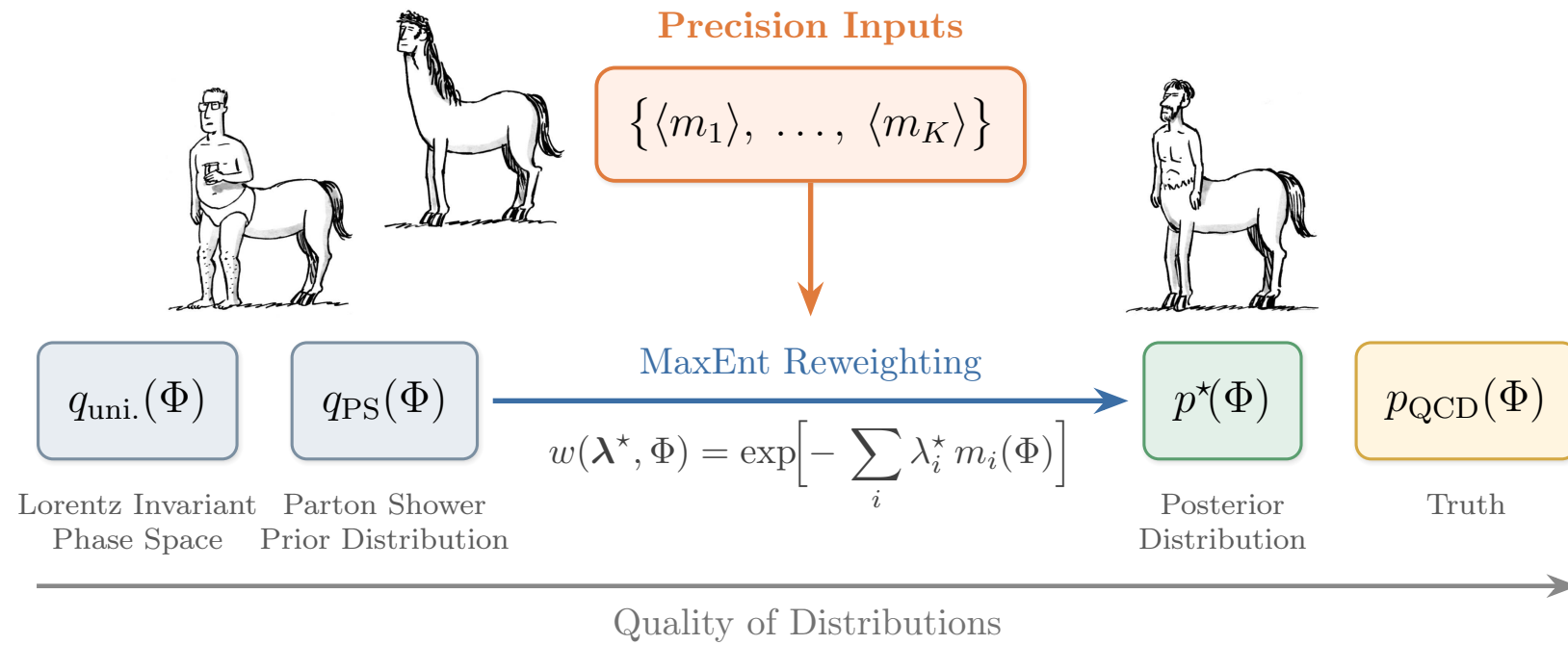
To a **QCD theorist**, this is a (simplified) **next-to-leading-logarithmic calculation**

To an **ML practitioner**, this is an opportunity for **neurosymbolic learning**

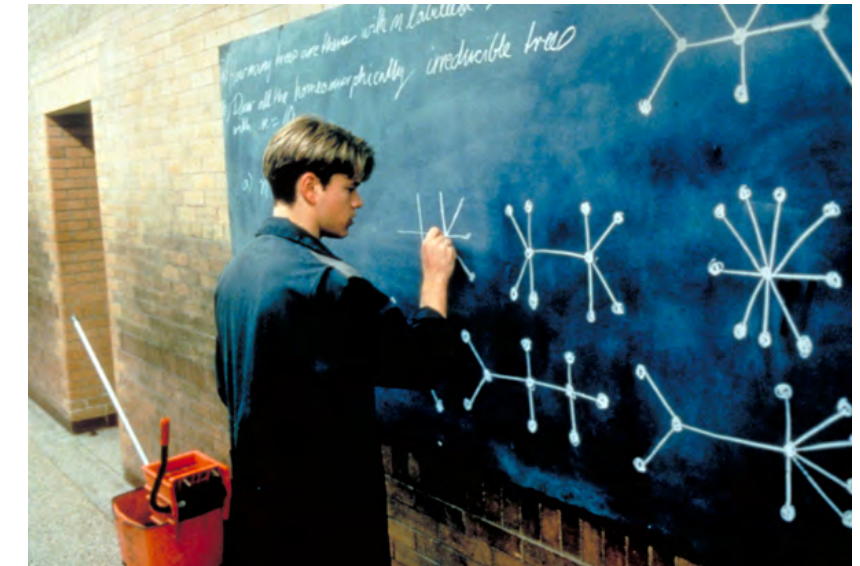
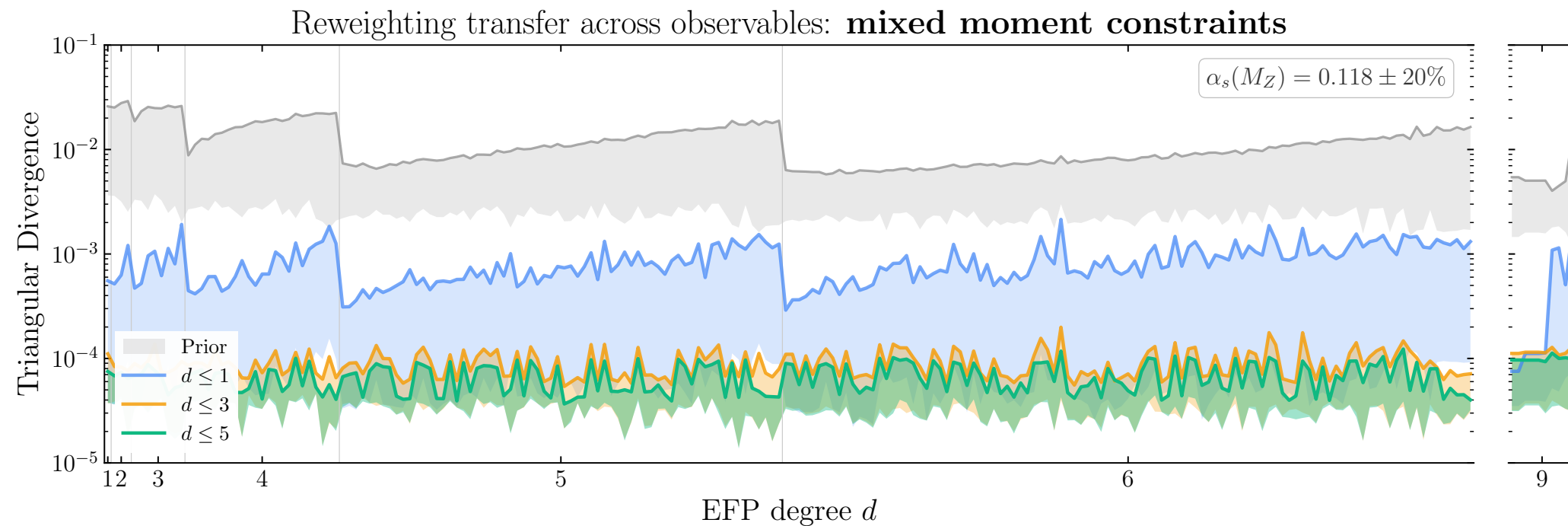
To a **Centaur Scientist**, the **Sudakov form factor** is a kind of **Boltzmann factor**,
where the **cusp anomalous dimension** is like a **Lagrange multiplier** that
enforces a constraint on the **second logarithmic moment** of the distribution

↳ *To my knowledge, logarithmic moments have never
been measured or calculated before in QCD!*

Towards a Trustable Foundation Model for Particle Theory

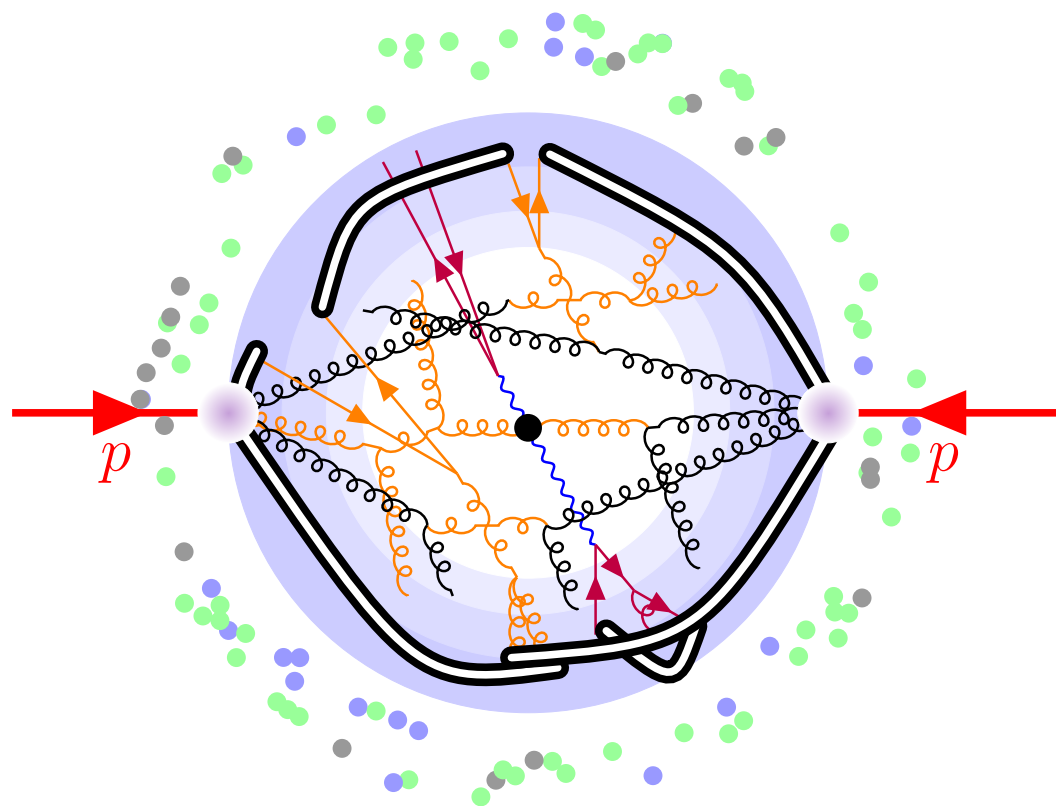


Systematic expansion in moments of Energy Flow Polynomials



[Assi, Lee, JDT, arXiv 2026; Assi, Lee, Höche, JDT, PRL 2025; using Komiske, Metodiev, JDT, JHEP 2018]



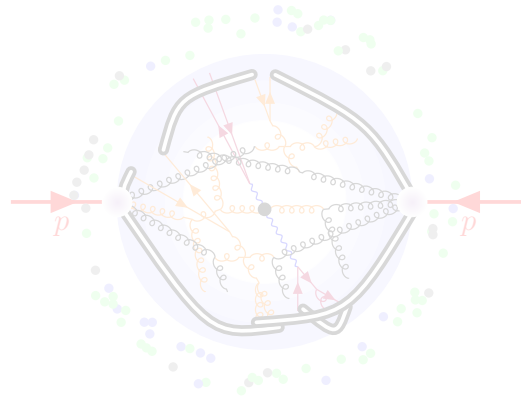


Centaur science continuing the tradition of relating properties of *physical systems* to concepts in *computation/information theory*



Machine Learning through the Lens of Physics

*AI is transforming physics research, but many foundational aspects of AI can be translated into the **language of physics***



QCD Theory meets Information Theory

*To confront thorny challenges in theoretical physics, we can reframe them as well-defined **optimization problems***



Scenes from my Sabbatical

*Being a “**centaur scientist**” means capitalizing on emerging AI technologies while insisting on scientific rigor and robustness*

Question I've been wrestling with on sabbatical: What does it mean to do robust **discovery science** with **AI**?

Exactness guarantees?

e.g. lattice sampling

Explicit verification?

e.g. symbolic engines

Statistical verification?

e.g. calibrated confidence intervals



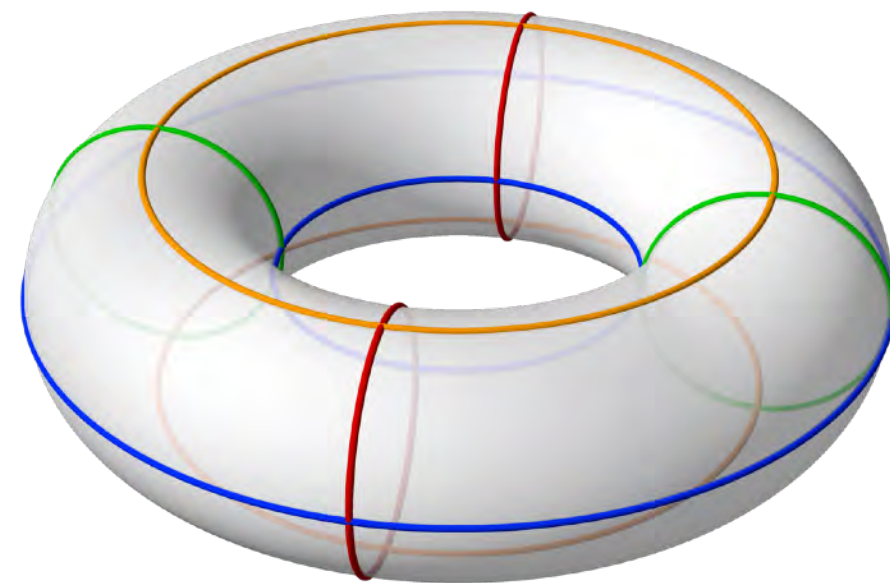
Sean Benevedes, PhD defense April 14, 2026!
[e.g. wifi ensembles in Benevedes, JDT, [PRD 2025](#);
TAMMs in Alvarez, Benevedes, Szewc, JDT, [arXiv 2026](#)]

Case study to show how my thinking has evolved:
Searching for “**new theoretical physics**” through **optimization**

Case Study: Two-Dimensional Conformal Field Theories

*Highly constrained mathematical framework that we know a lot about
(though not my area of expertise...)*

Posing a research question to  Claude



Please identify a two-dimensional conformal field theory with central charge in the range of 1.00 to 1.15.

The Z_5 parafermion CFT (equivalently, the $SU(2)_5/U(1)$ coset) has central charge $c = 8/7 \approx 1.143$.

[Fateev, Zamolodchikov, [Sov. Phys. JETP](#) 1985]

Only four decades too late...

[h/t to Siddharth Mishra-Sharma for the Claude credits!]

Case Study: Two-Dimensional Conformal Field Theories

Ok, great. So let me revise my question...

Please identify a two-dimensional conformal field theory with central charge in the range of 1.00 to 1.14.

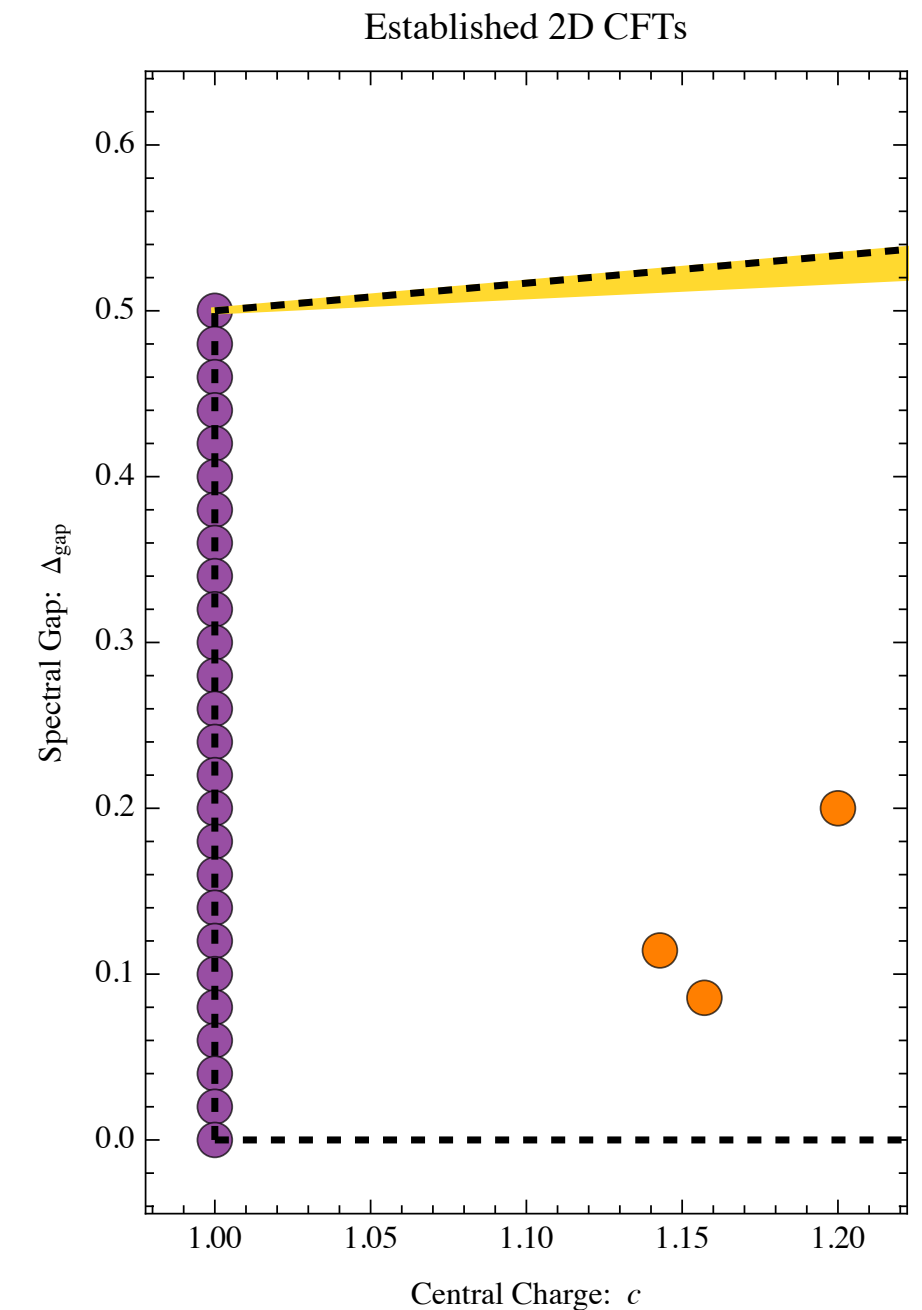
Ah, you've narrowed the window to exclude the Z_5 parafermion ($c = 8/7 \approx 1.1429$).

To my knowledge, there is no known unitary 2D CFT with central charge strictly in the range (1.00, 1.14). The theories cluster at the boundaries:

- **Below:** $c = 1$ theories (free boson, orbifolds)
- **Above:** Z_5 parafermion at $c = 8/7 \approx 1.143$, $N=1$ minimal model ($m=5$) at $c = 81/70 \approx 1.157$

Are there any other 2D CFTs in this range?

(And could I convince the experts if I found one? Could I convince myself?)



[Benjamin, Fitzpatrick, Li, JDT, arXiv 2026;

bounds from Hellerman, JHEP 2011; Friedan, Keller, JHEP 2013; Collier, Lin, Yin, JHEP 2018; Fitzpatrick, Li, JHEP 2024]



Strategies to Study/Discover CFTs

Ruling out theories (“dual”) is a lot easier than ruling them in (“primal”)

Indirect AI/ML?

Could you think about it a bit more and come up with a new theory that doesn't exist in the literature?

The honest answer: I can't construct a specific theory in (1.00, 1.14). The known RCFT constructions—cosets, W -algebra minimal models, orbifolds—all seem to have discrete spectra of central charges that skip this interval.

This could mean:

1. There's a deep reason (modular bootstrap + unitarity) forbidding unitary RCFTs here
2. There's an irrational CFT we haven't characterized
3. There's an exotic construction nobody has tried

At least it refuses to hallucinate...

Direct AI/ML?

Classification

[Chen, He, Lal, Zaz, [arXiv 2020](#); Kuo, Seif, Lundgren, Whitsitt, Hafezi, [PRR 2022](#)]

Reinforcement Learning

[Kántor, Niarchos, Papageorgakis, [PRL 2022](#), [PRD 2022](#); + Richmond, [PRD 2023](#)]

Generative Modeling

[Laio, Valenzuela, Serone, [PRD 2022](#)]

Gradient-Based Methods

[Reehorst, Rychkov, Simmons-Duffin, Sirois, Su, van Rees, [SciPost 2021](#)]
[Afkhami-Jeddi, [JHEP 2022](#)]

Genetic Algorithms

[Huang, Lee, Liao, Rumbutis, [arXiv 2025](#)]

Spectral Bias Studies

[Ghosh, Kumar, Niarchos, Stergiou, [arXiv 2026a](#), [arXiv 2026b](#)]

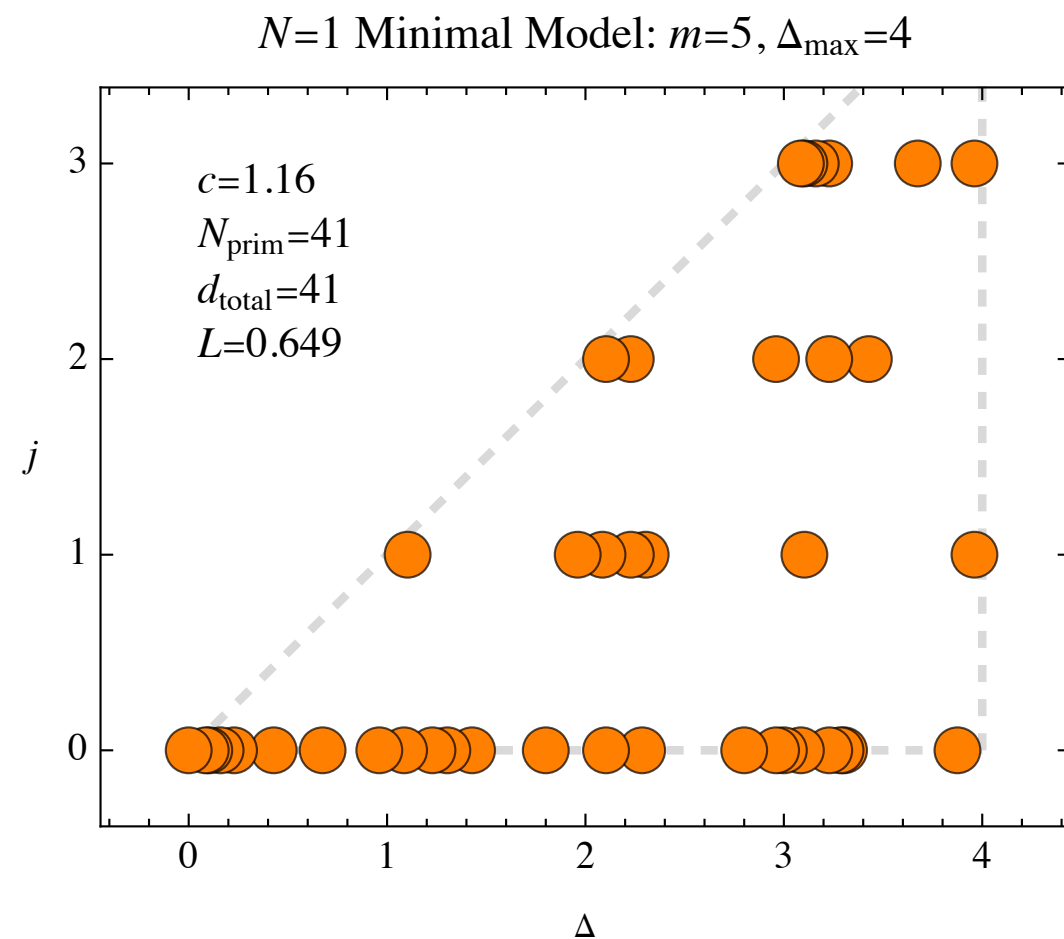
[Please let me know if I missed your ML-related CFT paper!]

[see reviews of dual approach to conformal bootstrap in Poland, Rychkov, Vichi, [RMP 2019](#); Rychov, Su, [RMP 2024](#); alternative numerical approach in Gliozzi, [PRL 2013](#), [JHEP 2016](#)]

Warm-Up Problem: Modular Bootstrap

Necessary but not sufficient condition for unitary 2D CFT to exist

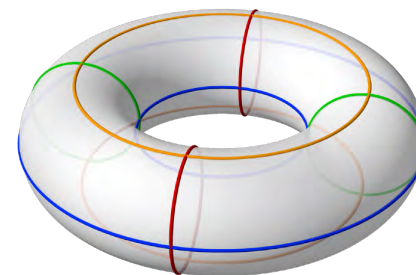
Example of more general **bootstrap philosophy** of defining theoretical objects by consistency conditions



Spectrum of (Virasoro Primary) Operators:

Dimension: Δ	$\Delta \geq j \geq 0$	assuming parity for $-j$
Spin: j	$d > 0$	
Degeneracy: d	j, d are integers	

(Euclidean) Partition Function: $Z(\tau; c, \{\Delta_a, j_a, d_a\})$



Modular Invariance:
 $Z(\tau) = Z(-1/\tau)$

[Cardy, [NPB 1991](#); Hellerman, [JHEP 2011](#); Friedan, Keller, [JHEP 2013](#); Collier, Lin, Yin, [JHEP 2018](#); ...]

“Thinking Like a Machine”

Modular bootstrap as “self-supervised learning”

✓ Training Data:
MC Integration over τ

$$L = \int_{\mathcal{F}} \frac{d\tau d\bar{\tau}}{\text{Im } \tau} \mathcal{L}$$

✗ Loss Function(al):
Modular Invariance

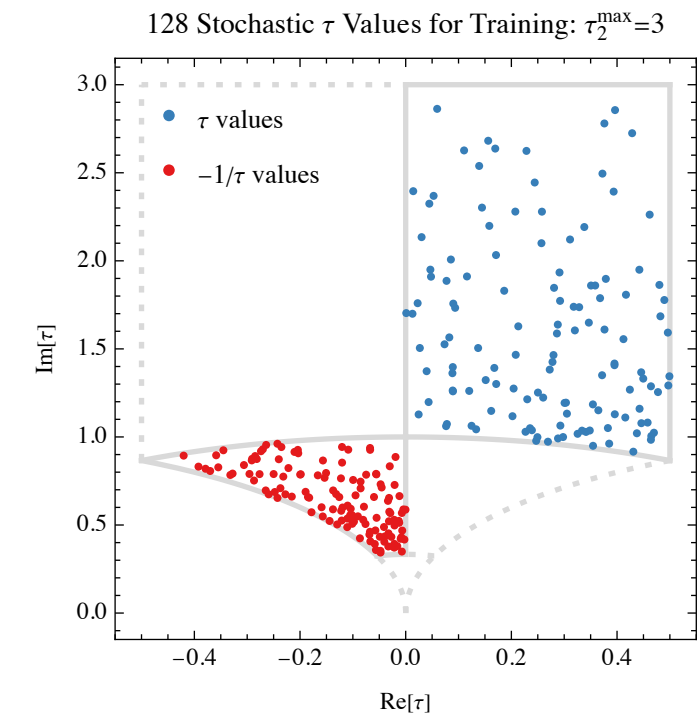
$$\mathcal{L} = (Z(\tau) - Z(-1/\tau))^2$$

✓ Learnable Function:
Partition Function

$Z(\tau; c, \{\Delta_a, j_a, d_a\})$ with $\Delta_a < \Delta_{\max}$ as parameters

✗ Optimizer:
Gradient Descent

$$\Delta_a^{(i)} = \Delta_a^{(i-1)} - \eta \frac{dL}{d\Delta_a}$$



Unfortunately, this approach fails dramatically at finding anything close to physical...

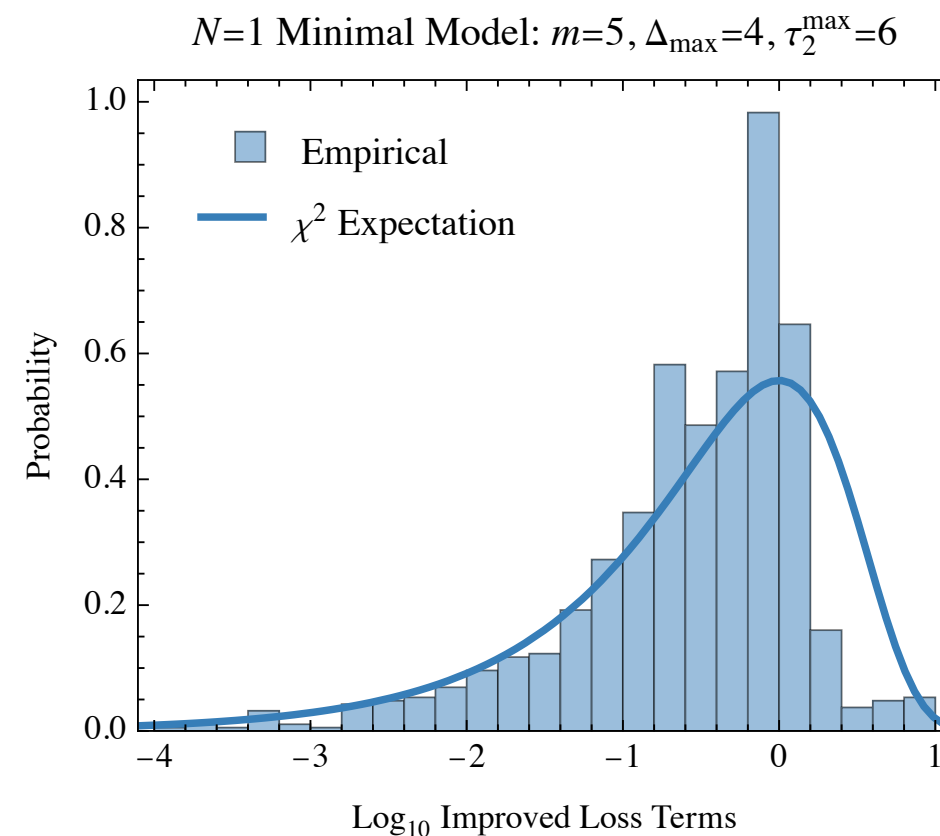
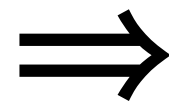
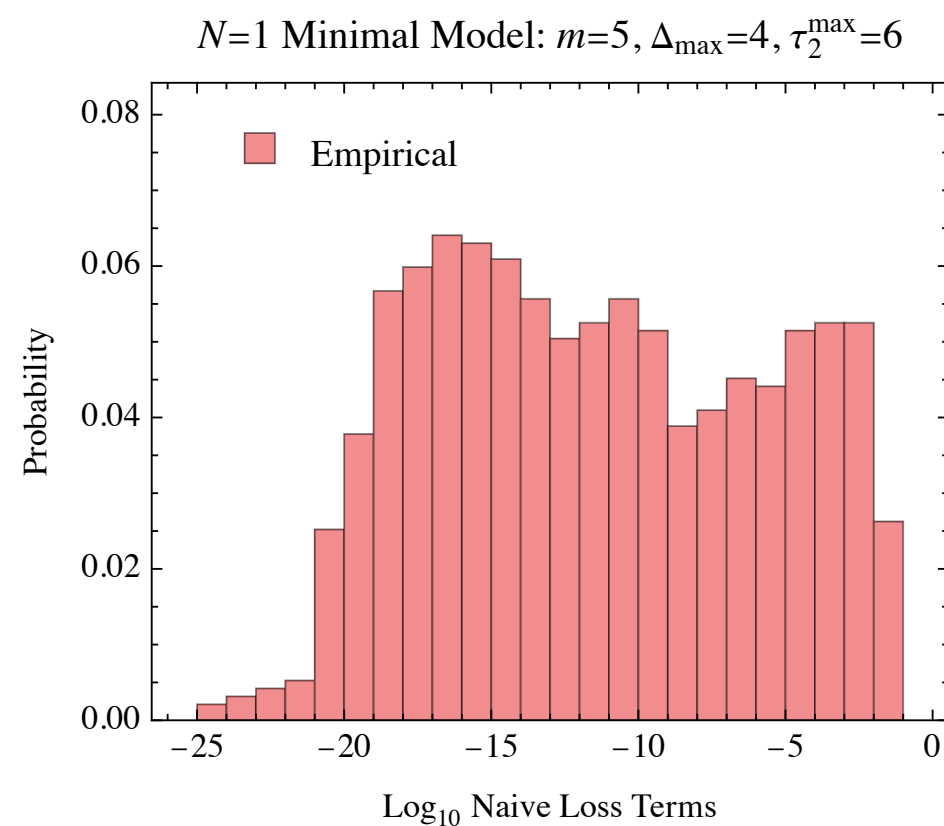
[similar philosophy to physics-informed machine learning for PDEs: Karniadakis, Kevrekidis, Lu, Perdikaris, Wang, Yang, [Nature Reviews 2021](#)]

[see ML approach to S-matrix bootstrap in Dersy, Schwartz, Zhiboedov, [JHEP 2024](#); Niarchos Papageorgakis, [PRD 2024](#); Gumus, Leflot, Tourkine, Zhiboedov, [JHEP 2025](#), [arXiv 2026](#)]

“Thinking Like an Experimental Physicist”

What are the **Uncertainties** on your Loss Function?

Use “ χ^2 ” instead of MSE:
$$\mathcal{L} = \left(\frac{Z(\tau) - Z(-1/\tau)}{\sigma(\tau)} \right)^2$$



Remarkable that this distribution comes from purely **theoretical data**

[Benjamin, Fitzpatrick, Li, JDT, [arXiv 2026](#)]

“Thinking Like an Experimental Physicist”

Needed a robust method to **quantify uncertainties**

Collaborating with Claude helped me **identify connections** in the literature

We previously discussed a strategy for taking a known CFT with $c = 1$ and deforming its partition function to mimic the Cardy scaling behavior of a CFT with $c > 1$, in order to estimate uncertainties associated with truncating the primary spectrum. Can you remind me of that formula and where it first appeared in the literature?

To shift a partition function from central charge c to a target c' , use:

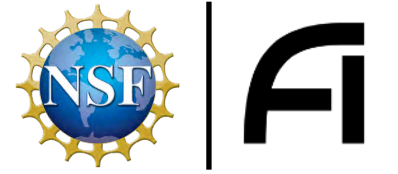
$$\tilde{Z}_{c'}(\tau) = (\text{Im } \tau)^{(c'-c)/2} \cdot |\eta(\tau)|^{-(c'-c)} \cdot Z_c(\tau)$$

The $|\eta|$ factor shifts the vacuum energy to give the correct $q^{-c'/24}$ scaling, while the $(\text{Im } \tau)^{1/2}$ factor preserves modular covariance under $S : \tau \rightarrow -1/\tau$.

A similar construction appears in **Benjamin, Collier, Fitzpatrick, Maloney, and Perlmutter, "Harmonic analysis of 2d CFT partition functions" (arXiv:2107.10744)**, where they define the primary partition function $Z_p = (y^{1/2}/|\eta|^2)^c Z$ to enable spectral decomposition on the modular fundamental domain.

[Benjamin, Fitzpatrick, Li, JDT, arXiv 2026]

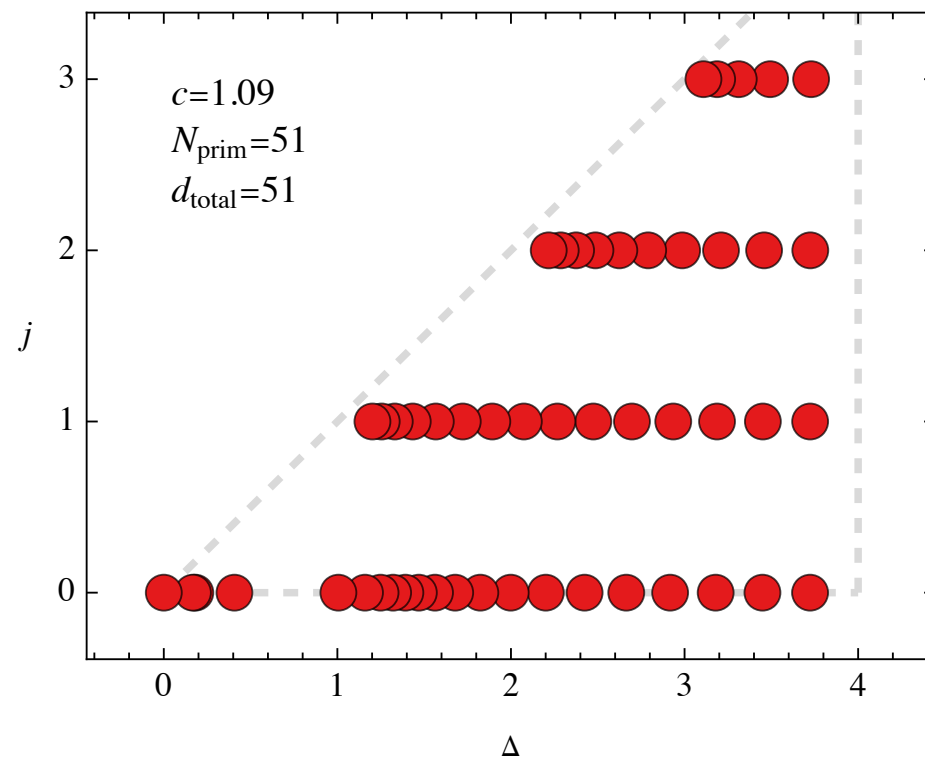
“Thinking Like an Mathematical Physicist”



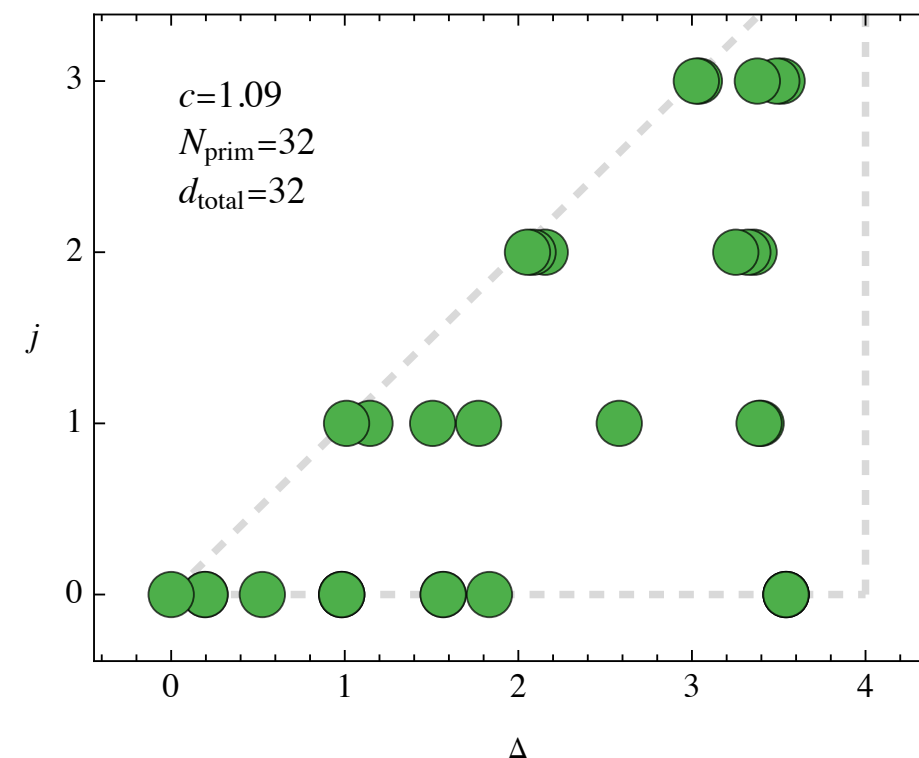
What is the **Geometry** of your Loss Landscape?

Use **quasi-second order** methods instead of SGD: $\Delta_a^{(i)} = \Delta_a^{(i-1)} - g_{ab} \frac{dL}{d\Delta_b}$

Spectrum After Optimization: Gradient Descent



Spectrum After Optimization: Sven: $N_{SV}=16$



Code name: “Sven”
Singular Value dEsceNt

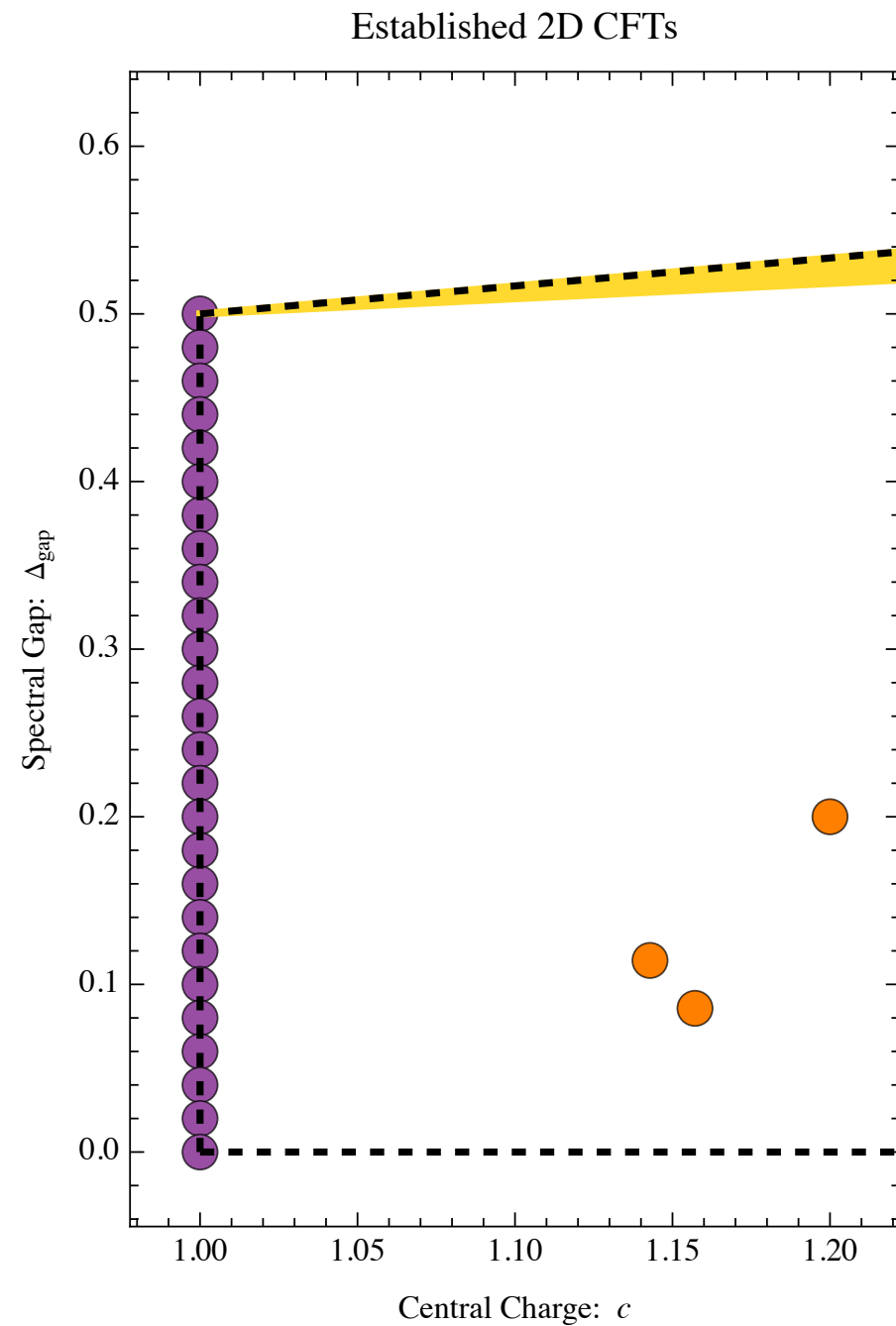
Now available in [GitHub!](#)

[Bright-Thonney, Harvey, Lukas, JDT, arXiv 2026]



Updating the Space of Candidate 2D CFTs

Remember: Modular invariance is necessary but not sufficient



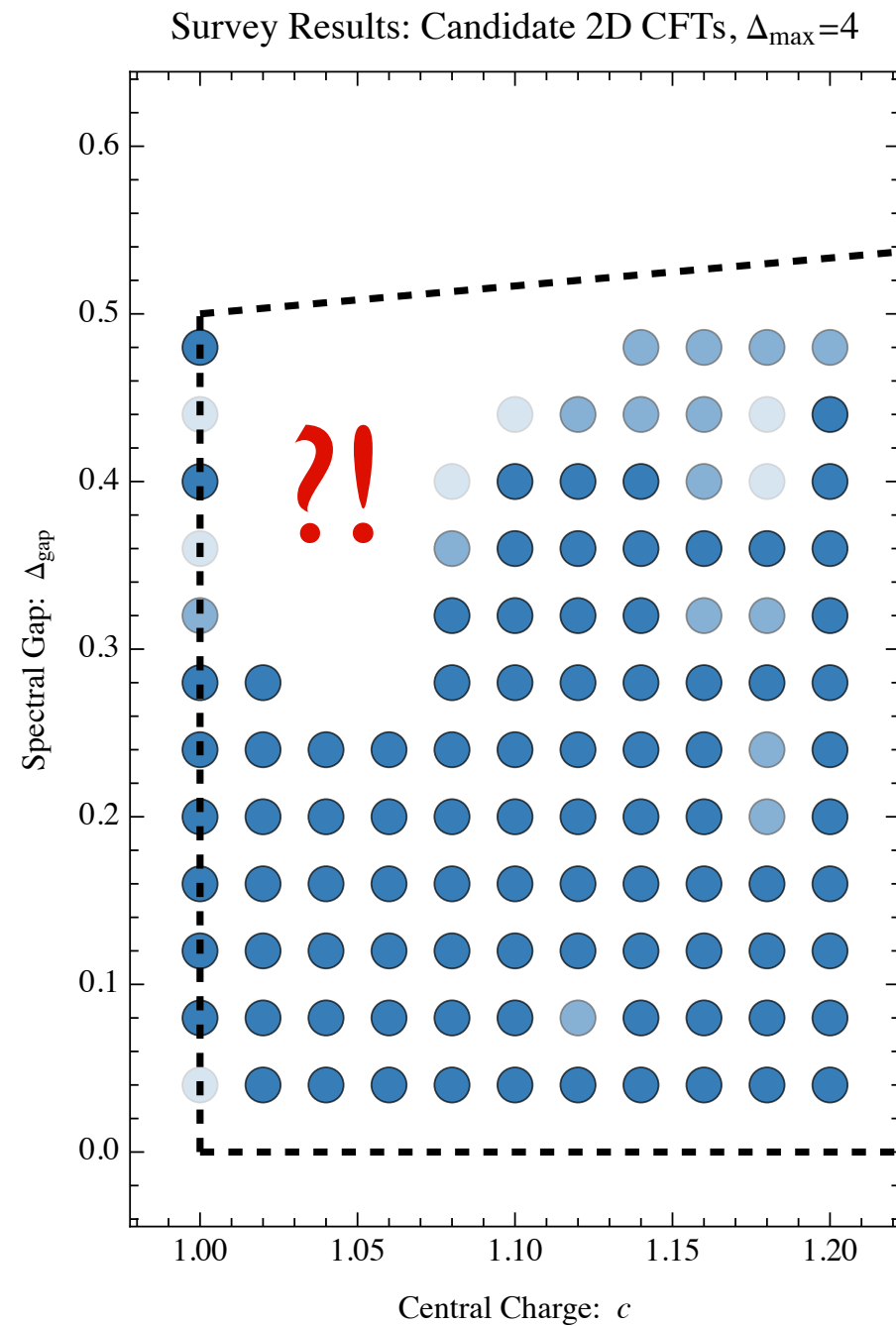
[Benjamin, Fitzpatrick, Li, JDT, [arXiv 2026](#);
Bright-Thonney, Harvey, Lukas, JDT, [arXiv 2026](#)]



Updating the Space of Candidate 2D CFTs



Remember: Modular invariance is necessary but not sufficient



Highly suggestive of an **obstruction**...

I don't know how we would have gotten this result without AI/ML methods!

Direct AI/ML: Reframe modular bootstrap in **optimization** language

Indirect AI/ML: **Reason** through conceptual issue of estimating uncertainties

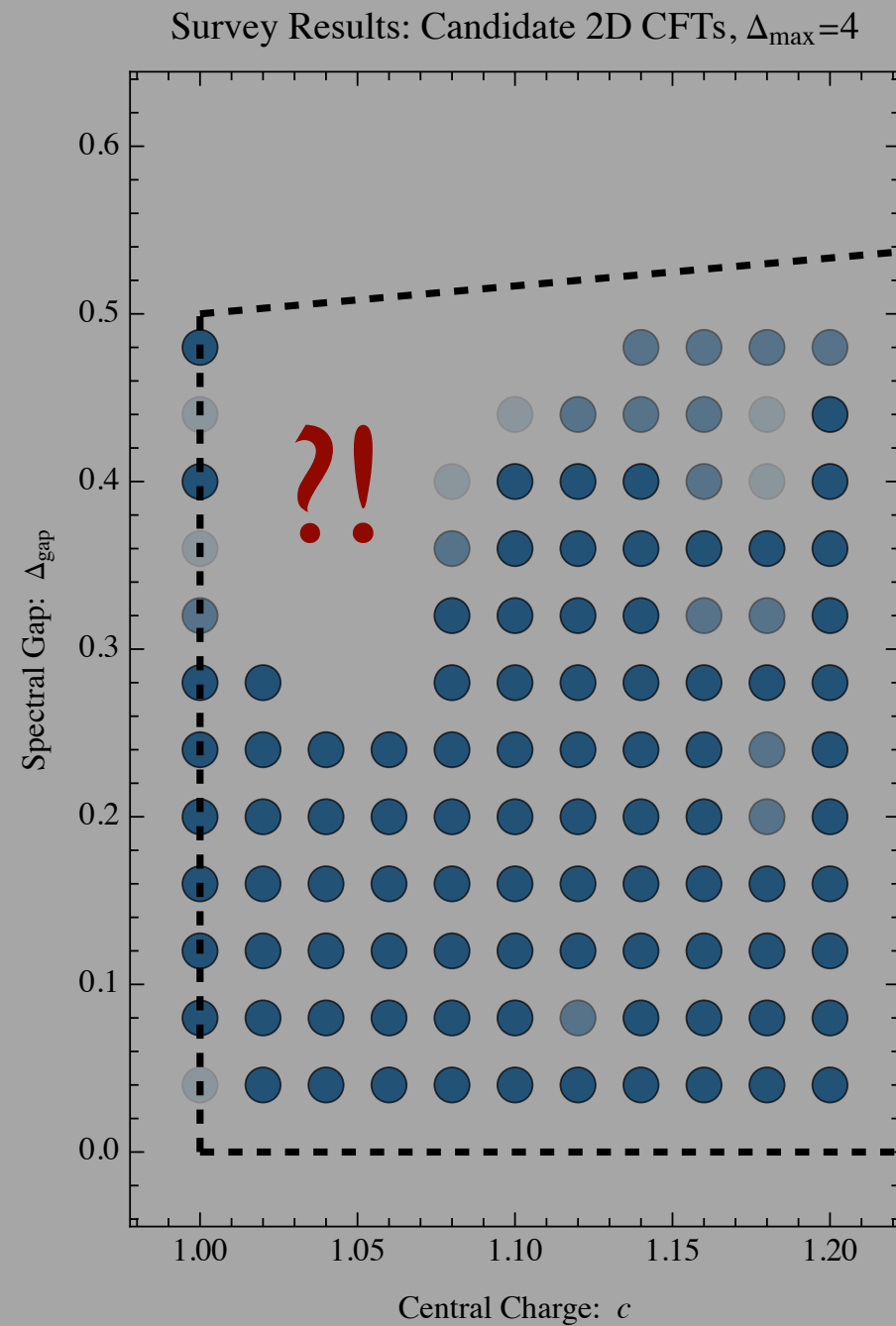
[Benjamin, Fitzpatrick, Li, JDT, [arXiv 2026](#);
Bright-Thonney, Harvey, Lukas, JDT, [arXiv 2026](#)]



Updating the Space of Candidate 2D CFTs



Remember: Modular invariance is necessary but not sufficient



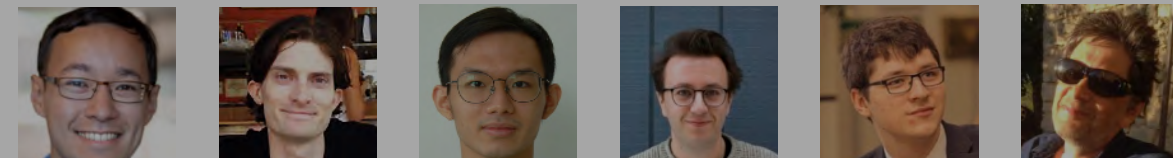
But I also don't know how we would have gotten this result without physics input!

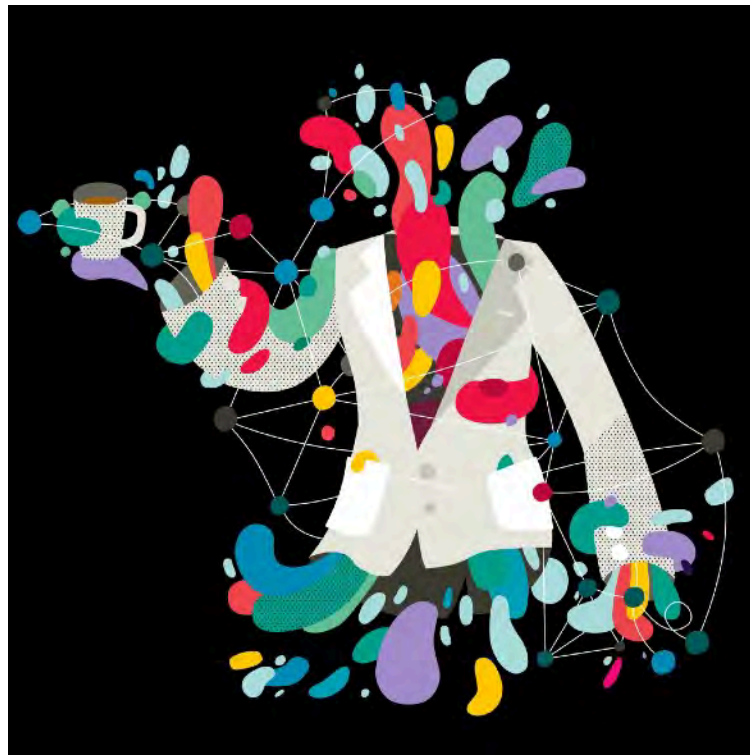
Identify tractable problem of **relevance to experts** in the field

Develop new optimization algorithms based on **physics intuition**

Pursue **statistical robustness** even in a purely theoretical setting

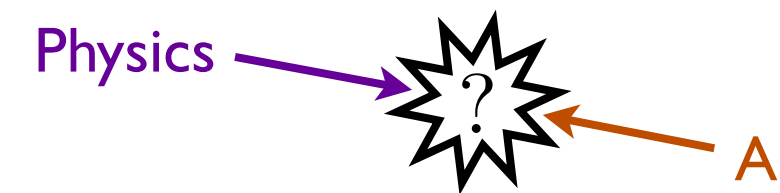
[Benjamin, Fitzpatrick, Li, JDT, [arXiv 2026](#);
Bright-Thonney, Harvey, Lukas, JDT, [arXiv 2026](#)]





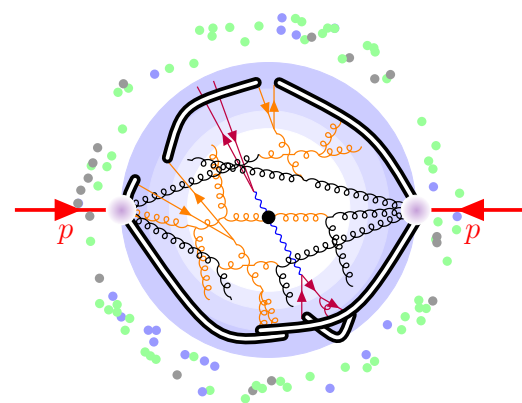
*Advancing (theoretical) physics through a
centaur-style merging of physical reasoning
and computational algorithms*

Centaur Science: Adventures in AI+Physics



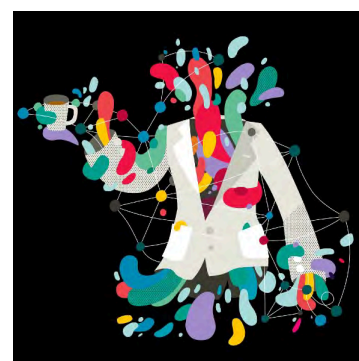
Machine Learning through the Lens of Physics

*AI is transforming physics research, but many foundational aspects of AI can be translated into the **language of physics***



QCD Theory meets Information Theory

*To confront thorny challenges in theoretical physics, we can reframe them as well-defined **optimization problems***



Scenes from my Sabbatical

*Being a “**centaur scientist**” means capitalizing on emerging AI technologies while insisting on scientific rigor and robustness*



The NSF Institute for Artificial Intelligence and Fundamental Interactions (IAIFI /aI-faI/ iaifi.org)



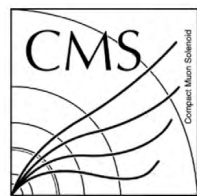
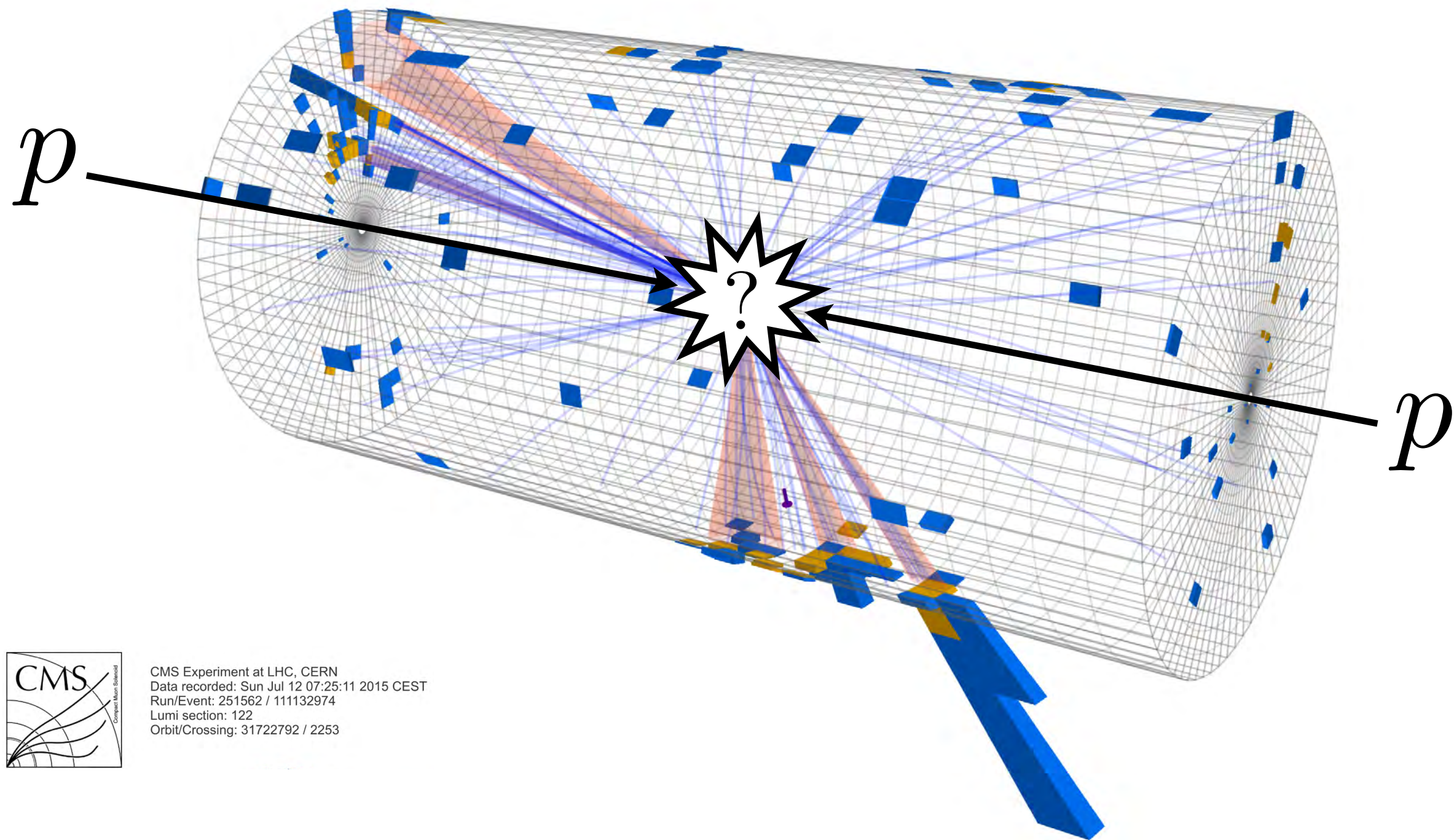
*Artificial intelligence
as a pathway to
scientific insight*

AI

*Physics intelligence
as a pathway to
AI innovation*

*Progress in AI+Physics driven by
early career talent with interdisciplinary training*

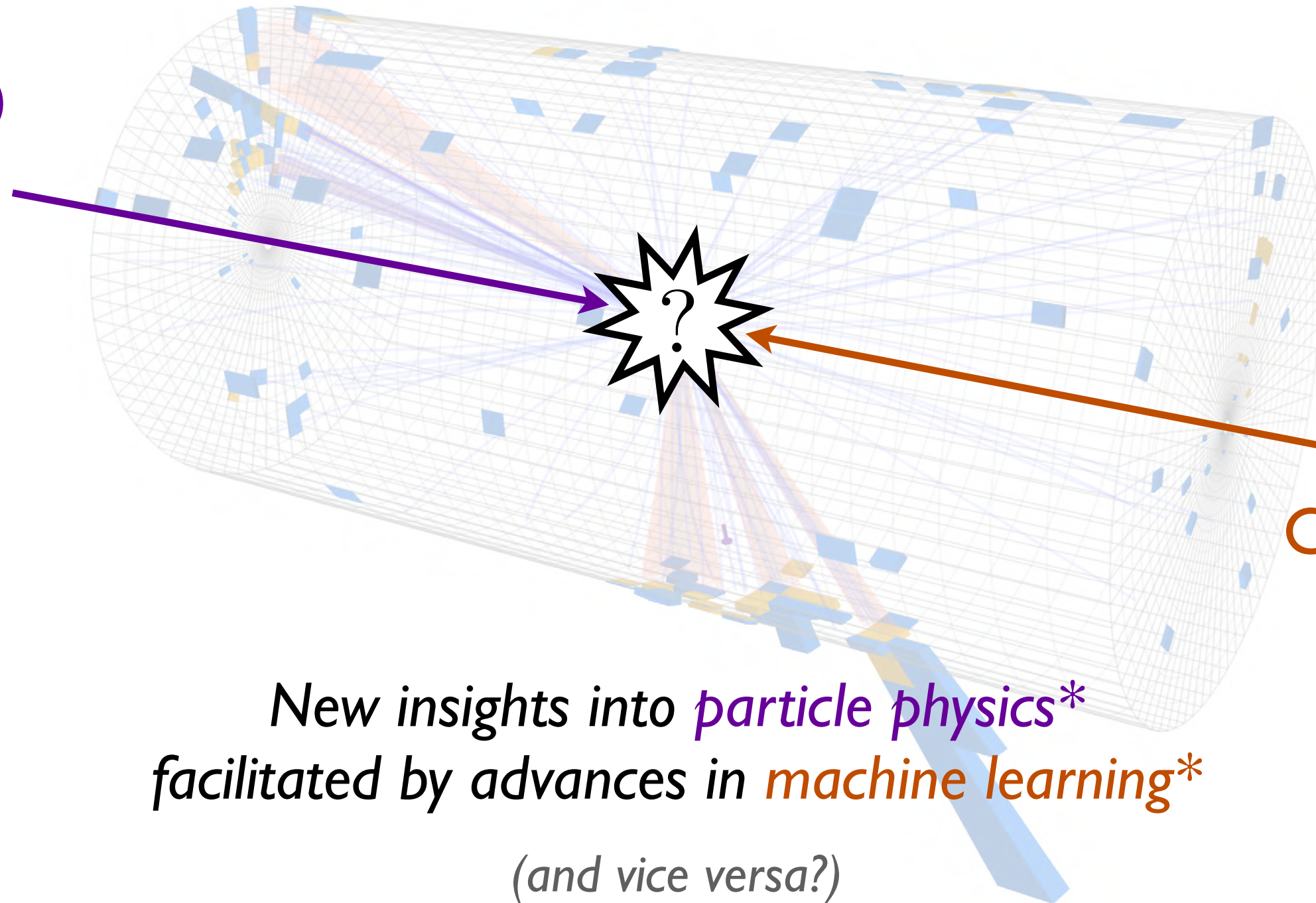
Backup Slides



CMS Experiment at LHC, CERN
Data recorded: Sun Jul 12 07:25:11 2015 CEST
Run/Event: 251562 / 111132974
Lumi section: 122
Orbit/Crossing: 31722792 / 2253

Collision of HEP and ML

(Theoretical)
High Energy
Physics



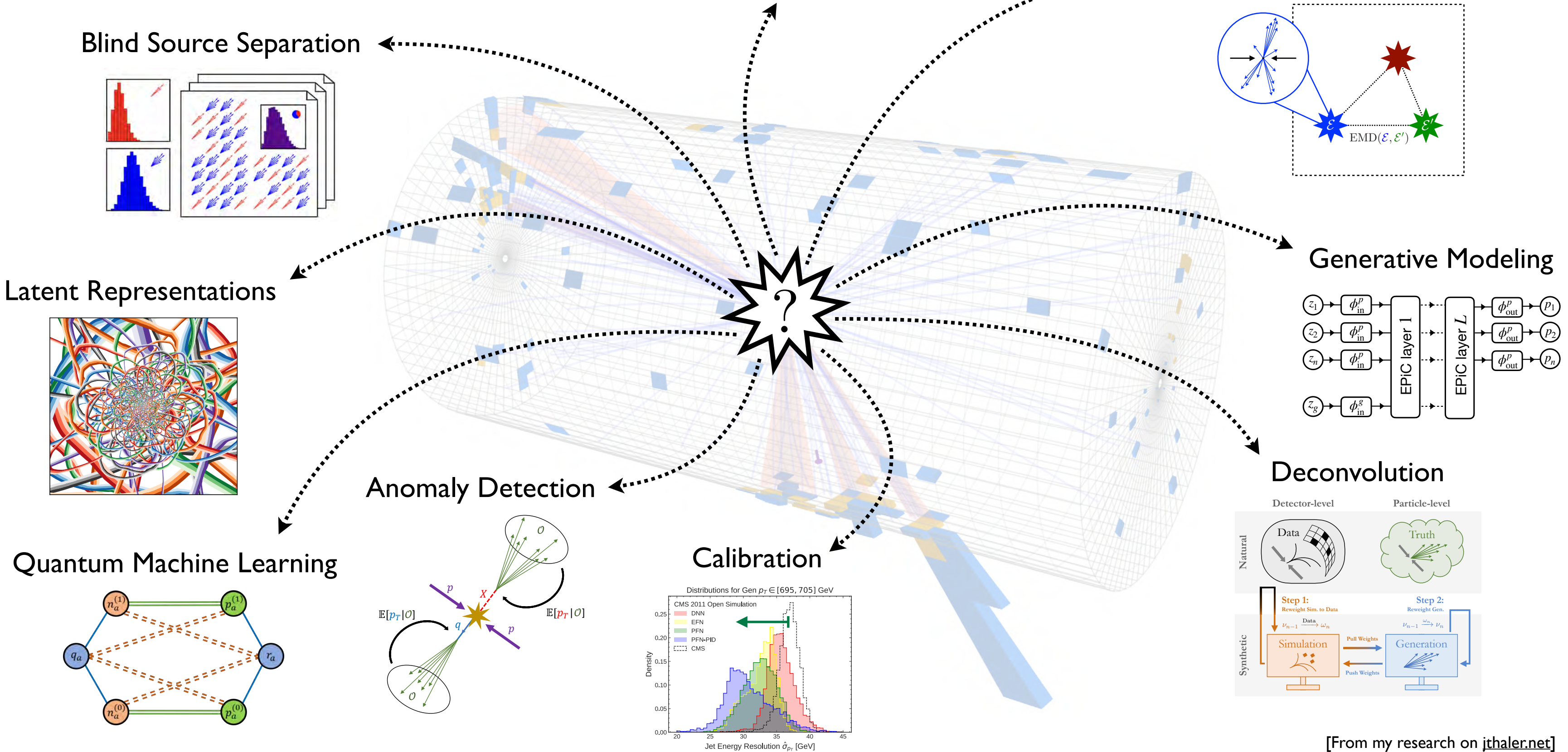
Mathematics,
Statistics,
Computer Science

*New insights into **particle physics***
facilitated by advances in **machine learning****

(and vice versa?)

A Smashing Success!

The Next Big Thing?



[From my research on jthaler.net]



The Future of Artificial Intelligence and the Mathematical and Physical Sciences (AI+MPS)

*Community Paper from the NSF Future of AI+MPS Workshop
Cambridge, Massachusetts — March 24–26, 2025*



Cross-Disciplinary Opportunities:

Advocate for Diverse Funding Streams

Pursue the Science of AI →

Establish Scalable AI Infrastructures

Facilitate Interdisciplinary Collaborations

Cultivate Key AI Techniques for Science

Leverage AI for Conducting Research

Educate and Train an AI+MPS Workforce

Empower AI Innovation

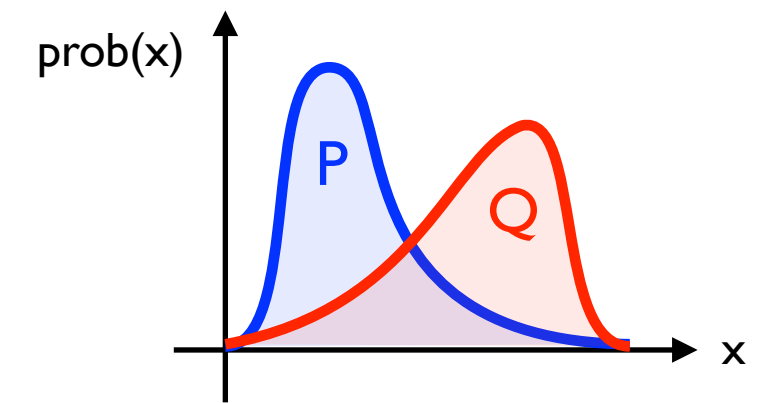
AI Innovations from Science
Understanding AI Behaviors
Robust and Reproducible AI

Machine Learning I01: Likelihood Ratio Trick

Key tool for *simulation-based inference*

Many HEP problems can be expressed in this form!

Goal:	Estimate $p(x) / q(x)$
Training Data:	Finite samples P and Q
Learnable Function:	$f(x)$ parametrized by, e.g., neural networks
Loss Function(al):	$L = -\langle \log f(x) \rangle_P + \langle f(x) - 1 \rangle_Q$



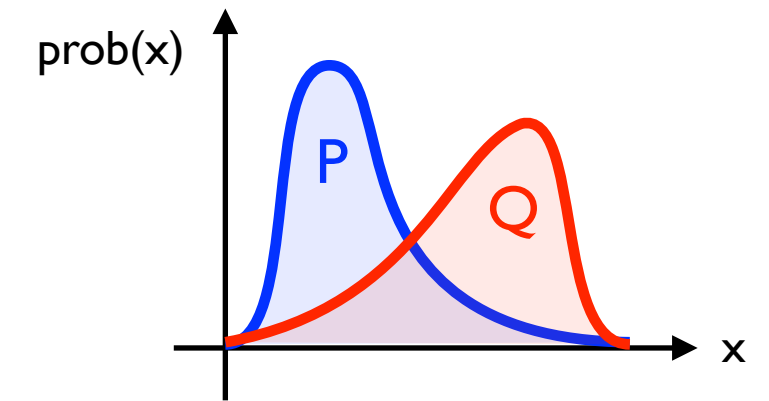
[see e.g. Cranmer, Pavez, Louppe, [arXiv 2015](#); D'Agnolo, Wulzer, [PRD 2019](#);
simulation-based inference in Cranmer, Brehmer, Louppe, [PNAS 2020](#);
relation to f-divergences in Nguyen, Wainwright, Jordan, [AoS 2009](#); Nachman, [JDT, PRD 2021](#)]

Machine Learning 101: Likelihood Ratio Trick

Key tool for *simulation-based inference*

Many HEP problems can be expressed in this form!

Goal: Estimate $p(x) / q(x)$
Training Data: Finite samples P and Q
Learnable Function: $f(x)$ parametrized by, e.g., neural networks



Action/Lagrangian:
assuming enough data

$$L \approx \int dx \left(-p(x) \log f(x) + q(x) (f(x) - 1) \right) !!$$

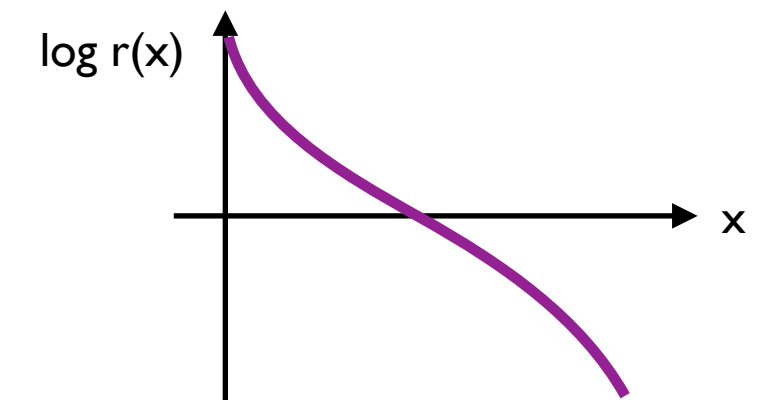
Euler-Lagrange:

$$\arg \min_{f(x)} L = \frac{p(x)}{q(x)}$$

Likelihood ratio

$$-\min_{f(x)} L = \int dx p(x) \log \frac{p(x)}{q(x)}$$

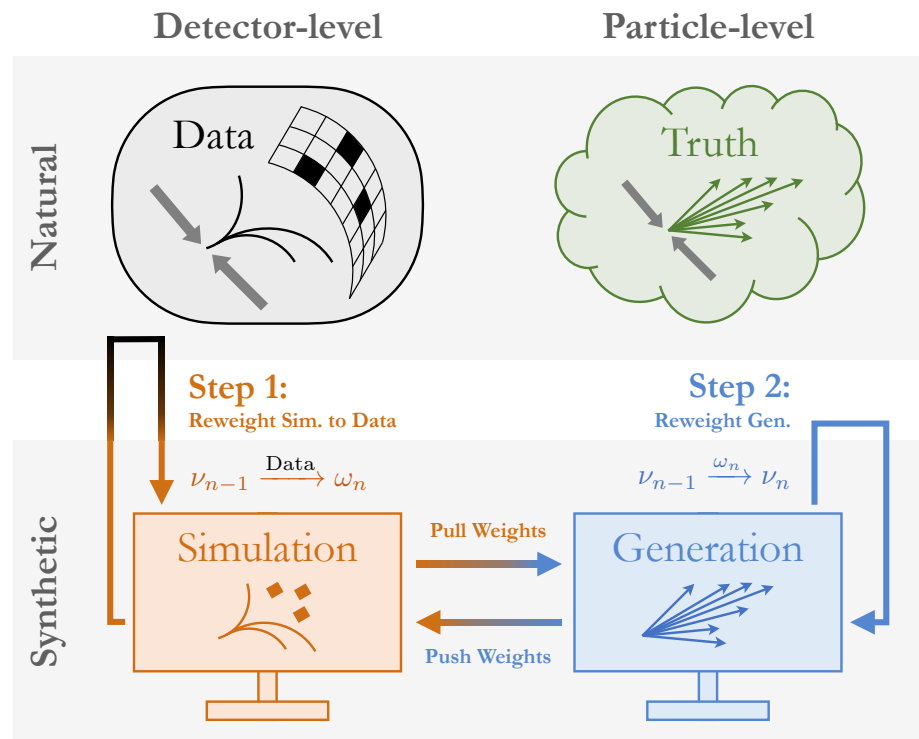
Kullback–Leibler divergence



[see e.g. Cranmer, Pavez, Louppe, [arXiv 2015](#); D’Agnolo, Wulzer, [PRD 2019](#);
simulation-based inference in Cranmer, Brehmer, Louppe, [PNAS 2020](#);
relation to f-divergences in Nguyen, Wainwright, Jordan, [AoS 2009](#); Nachman, [JDT, PRD 2021](#)]

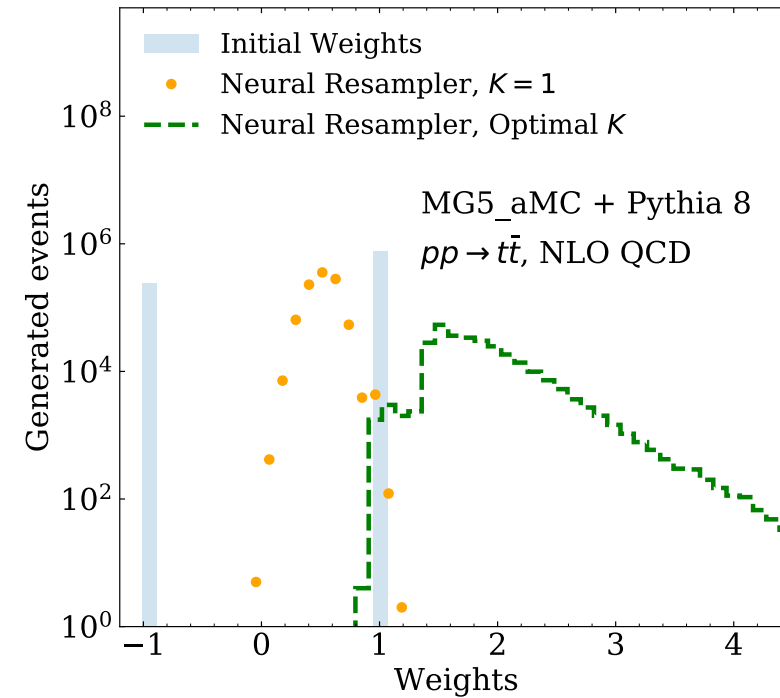
Applications of Likelihood Ratio Trick

Detector Unfolding



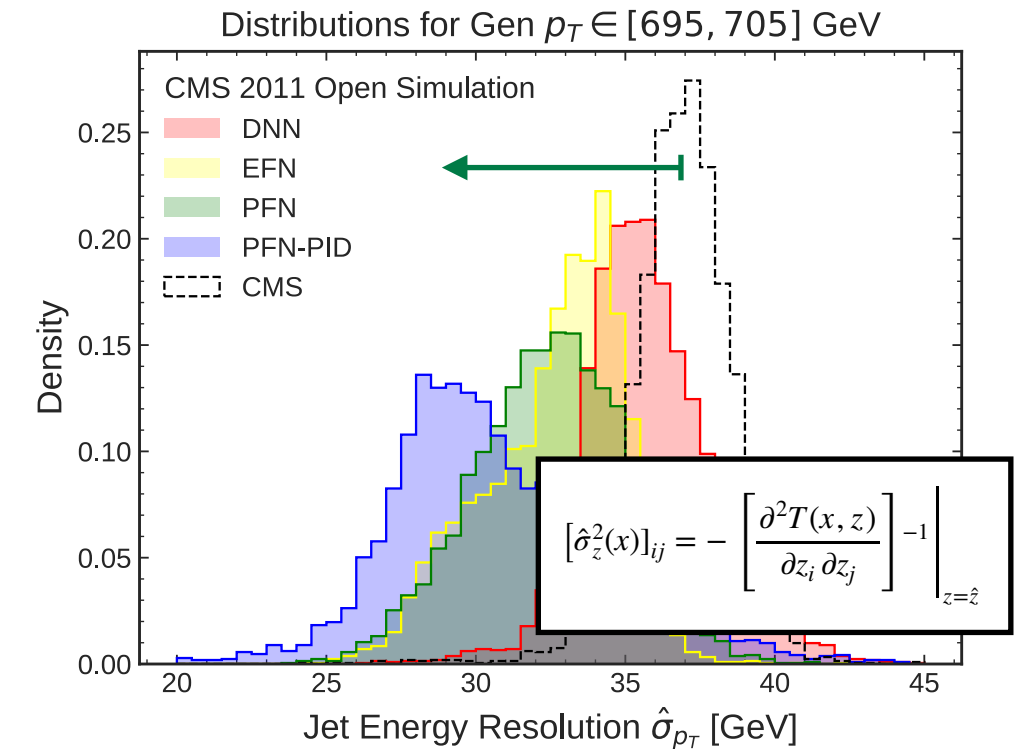
[Andreassen, Komiske, Metodiev, Nachman, JDT, PRL 2020; + Suresh, ICLR SimDL 2021]

Monte Carlo Reweighting



[Nachman, JDT, PRD 2020; inspired by Andersen, Gutschow, Maier, Prestel, EPJC 2020]

Resolution Estimation



[Gambhir, Nachman, JDT, PRL 2022, PRD 2022]

Theme: Convert sampled data into usable function approximation

And you can extract frequentist uncertainties on the likelihood ratio!
[see wifi ensembles in Benevedes, JDT, PRD 2025]



Indirect AI/ML: Natural Language Processing

ChatGPT with clever prompt engineering

From April Fools' Sendup to Public Engagement Event!



marAI curie

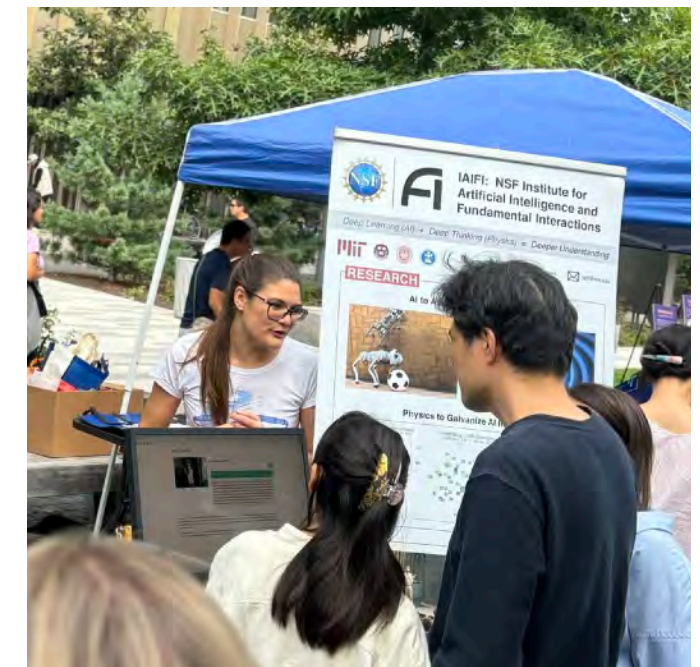


Please, pose your question below.

Ask

Why is it important to fund AI and Science?

While I am not familiar with the concept of "AI", or artificial intelligence, as it is beyond my time, I can certainly speak to the importance of funding science. Science is the key to understanding the world around us. It is through scientific research that we have been able to make significant advancements in medicine, technology, and various other fields. In my own experience, it was through the generous funding of our research that my husband, Pierre, and I were able to discover the elements polonium and radium. This discovery not only expanded our understanding of atomic structure and radioactivity, but also led to the development of new treatments for cancer. Therefore, funding science is not just an investment in knowledge, but also an investment in the future of humanity. It allows us to push the boundaries of what we know and to improve the quality of life for all. As for AI, if it is a scientific field, I would presume that funding it would have similar benefits, allowing for advancements that could potentially benefit society in ways we cannot yet imagine.

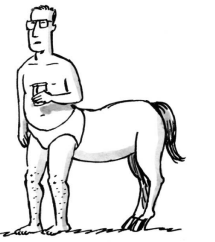


[Grosso, CSF 2024]

[Mishra-Sharma, Gambhir, and other troublemakers, chatjesset.com]

Information Theory and Optimization

“Predicting” a distribution by maximizing entropy



Shannon Entropy:
$$H(X) = \sum_{i=1}^n p(x_i) \log \frac{1}{p(x_i)}$$

Maximum:
$$p(x_i) = \frac{1}{n}$$

Continuum Version

Relative Entropy:
(KL divergence we saw before)

$$-D_{\text{KL}}(P||Q) = \int dx p(x) \log \frac{q(x)}{p(x)} + (\beta_0 - 1) \left(1 - \int dx p(x)\right)$$

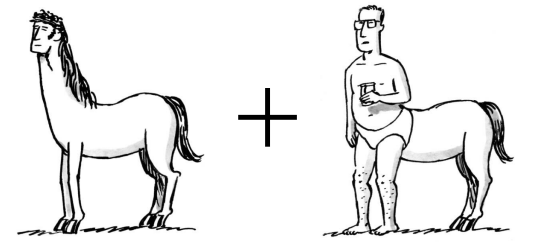
↑
Lagrange Multiplier to enforce normalization

Extremum:
$$p(x) = e^{-\beta_0} q(x) \quad \beta_0 = 0 \quad \Rightarrow \quad p(x) = q(x)$$

Maximizing (unconstrained) entropy is equivalent to **choosing a prior**

Information Theory and Constrained Optimization

“Predicting” a distribution by maximizing entropy with *fixed moments*



Relative Entropy:
$$-D_{\text{KL}}(P\|Q) = \int dx p(x) \log \frac{q(x)}{p(x)} + (\beta_0 - 1) \left(1 - \int dx p(x)\right)$$

Plus **Moment** Constraints:
$$+ \sum_j \beta_j \left(c_j - \int dx p(x) f_j(x)\right)$$

Extremum:
$$p(x) = q(x) \exp \left[-\beta_0 - \sum_j \beta_j f_j(x) \right]$$

Same manipulation as Boltzmann’s approach to statistical mechanics!

$e^{-\beta E}$: *Lagrange multipliers* set to values that *satisfy constraints*

[Assi, Lee, Höche, JDT, PRL 2025]



Information Theory and Numeric Constrained Optimization

“Predicting” a distribution by maximizing entropy with *fixed moments* and auto-differentiation



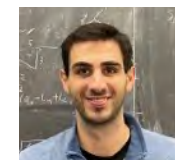
Extremum: $p(x) = q(x) \exp \left[-\beta_0 - \sum_j \beta_j f_j(x) \right]$

Target **Moment**: c_j Current **Moment**: $d_j = \int dx p(x) f_j(x)$

Loss Function: $\mathcal{L} = \sum_j \left(\frac{c_j - d_j}{c_j + d_j} \right)^2$

Numerically minimize loss to estimate *Lagrange multipliers*
As long as they aren't strictly incompatible, works for any set of *basis functions*

[Assi, Lee, Höche, JDT, PRL 2025]



Simplified Calculation for Thrust

Leading log at fixed coupling

$$p(\tau) = \frac{-2\alpha_s C_F \ln \tau}{\pi} \frac{1}{\tau} \exp \left[- \frac{\alpha_s C_F}{\pi} \ln^2 \tau \right]$$

Sudakov Form Factor

Ordinary Moments

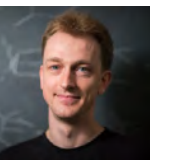
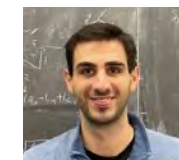
$$\langle \tau^m \rangle = \frac{2\alpha_s C_F}{\pi} \frac{1}{m^2} + \mathcal{O}(\alpha_s^2)$$

Logarithmic Moments

$$\langle \ln^n \tau \rangle = (-1)^n \left(\frac{\pi}{\alpha_s C_F} \right)^{n/2} \Gamma \left[1 + \frac{n}{2} \right]$$

Example of “Sudakov Safe” Observable

[Assi, Lee, Höche, JDT, PRL 2025;
Sudakov safety in Larkoski, Marzani, JDT, PRD 2015]



Simplified Calculation for Thrust

Leading log at fixed coupling

Sudakov Form Factor

$$p(\tau) = \frac{-2\alpha_s C_F \ln \tau}{\pi} \frac{1}{\tau} \exp \left[-\frac{\alpha_s C_F}{\pi} \ln^2 \tau \right]$$

Mean \Rightarrow Strong Coupling

$$\langle \tau \rangle = \frac{2\alpha_s C_F}{\pi} + \mathcal{O}(\alpha_s^2)$$

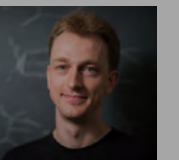
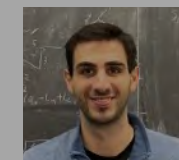
Characterizes **Fixed-Order** Information

Logarithmic Mean \Rightarrow Sudakov Peak

$$\langle \ln \tau \rangle = -\frac{\pi}{2\sqrt{\alpha_s C_F}}$$

Characterizes **Resummed** Information

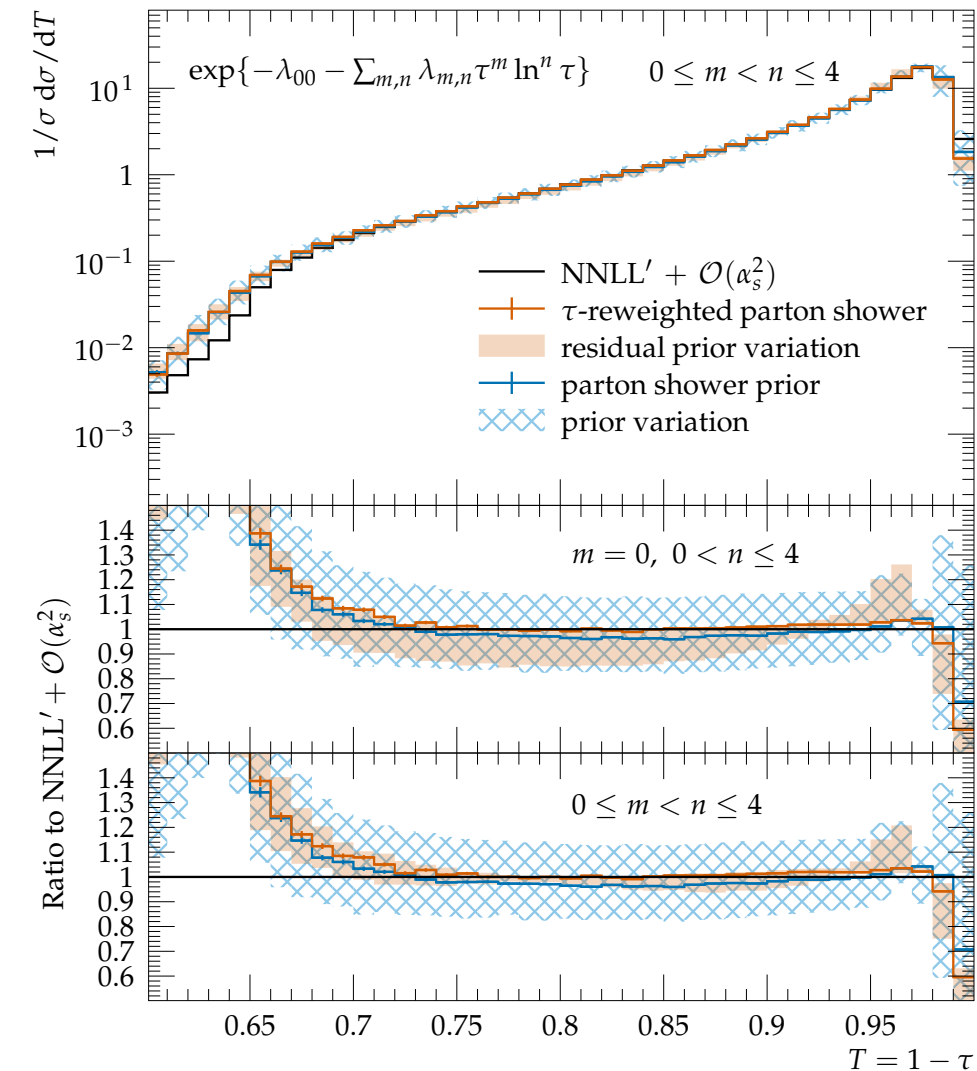
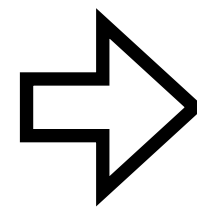
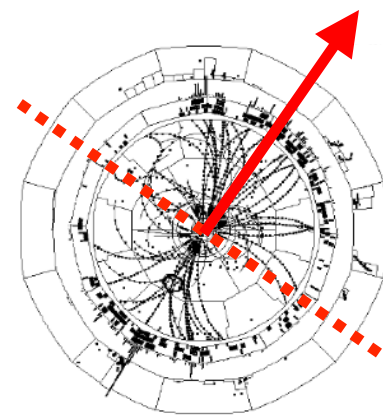
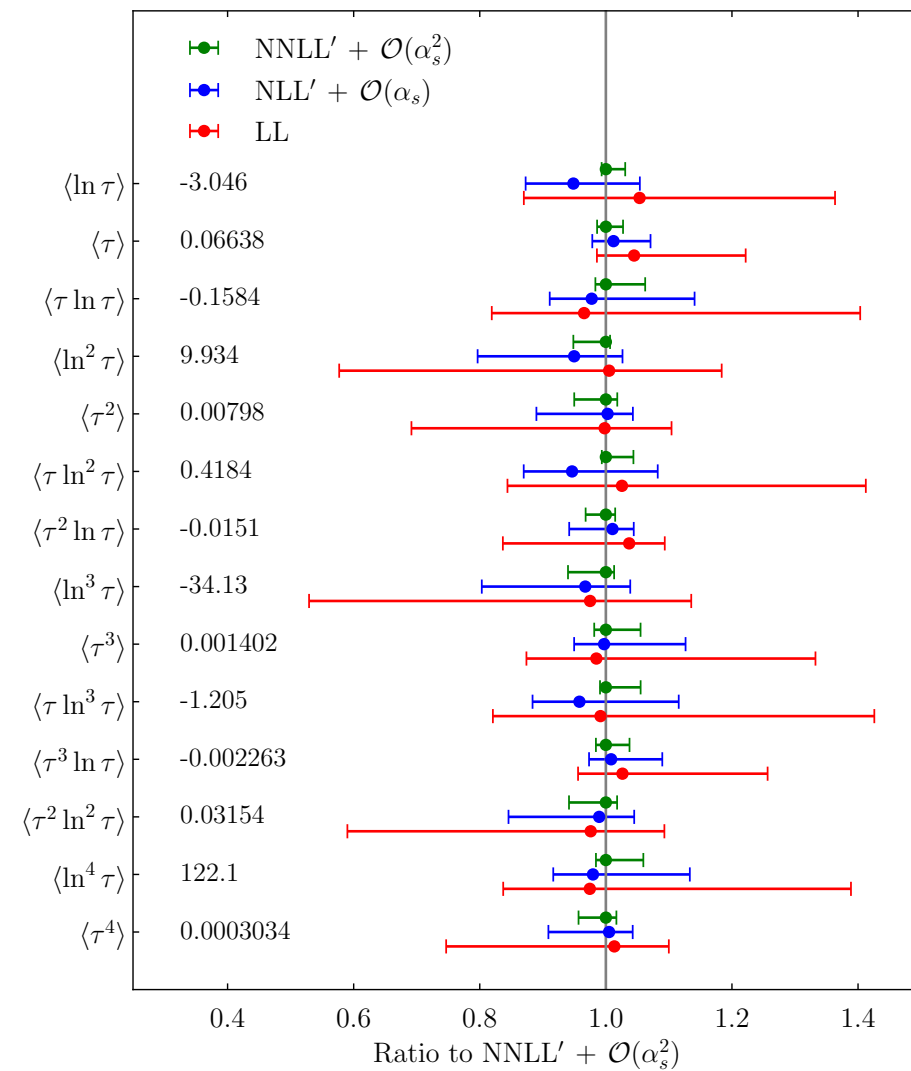
[Assi, Lee, Höche, JDT, PRL 2025;
Sudakov safety in Larkoski, Marzani, JDT, PRD 2015]



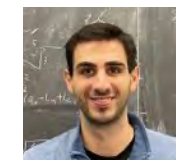
Proof-of-Concept Study with “Thrust”

First **QCD Calculation** of Thrust Logarithmic Moments!

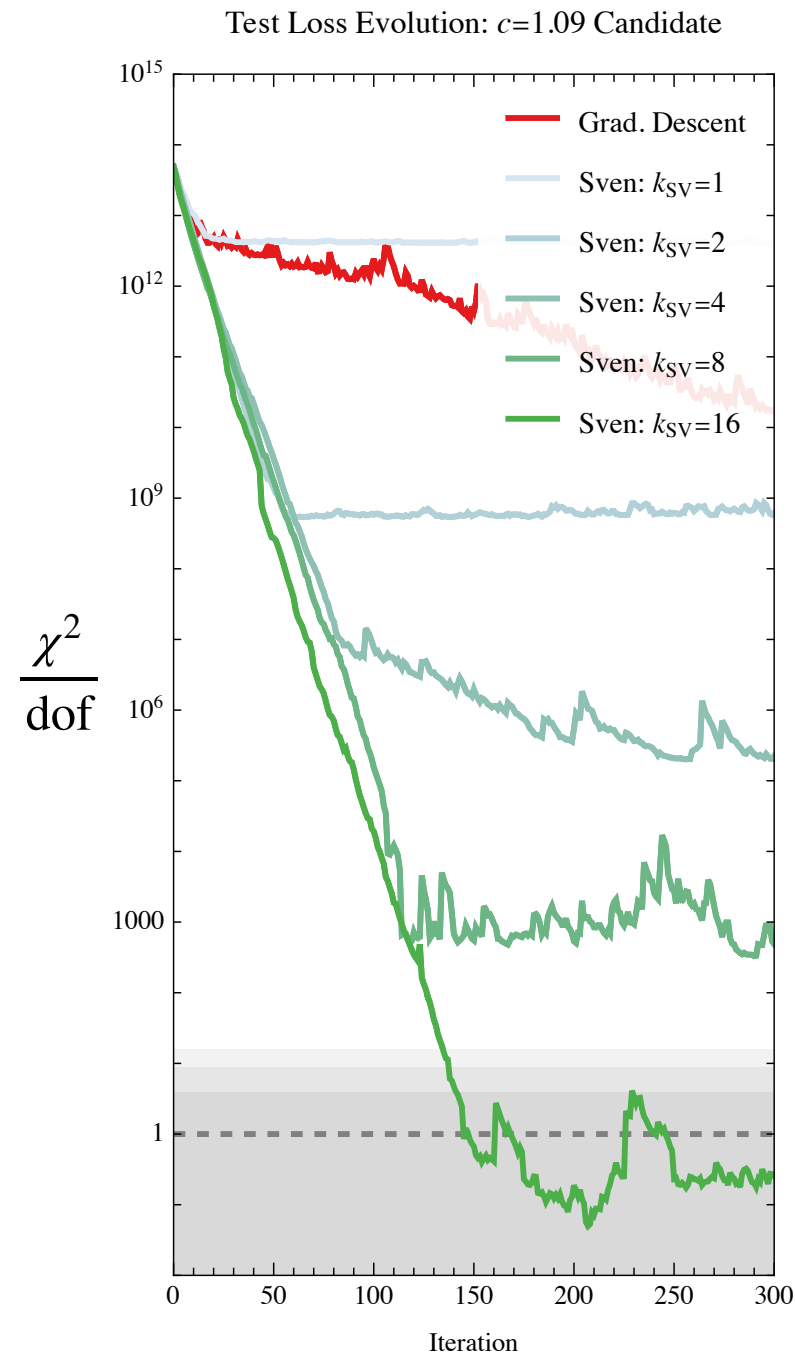
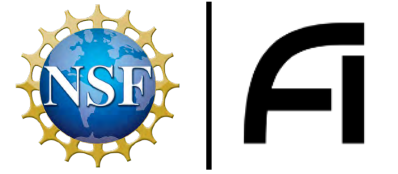
Priors from imperfect generators yield consistent results after **moment reweighting!**



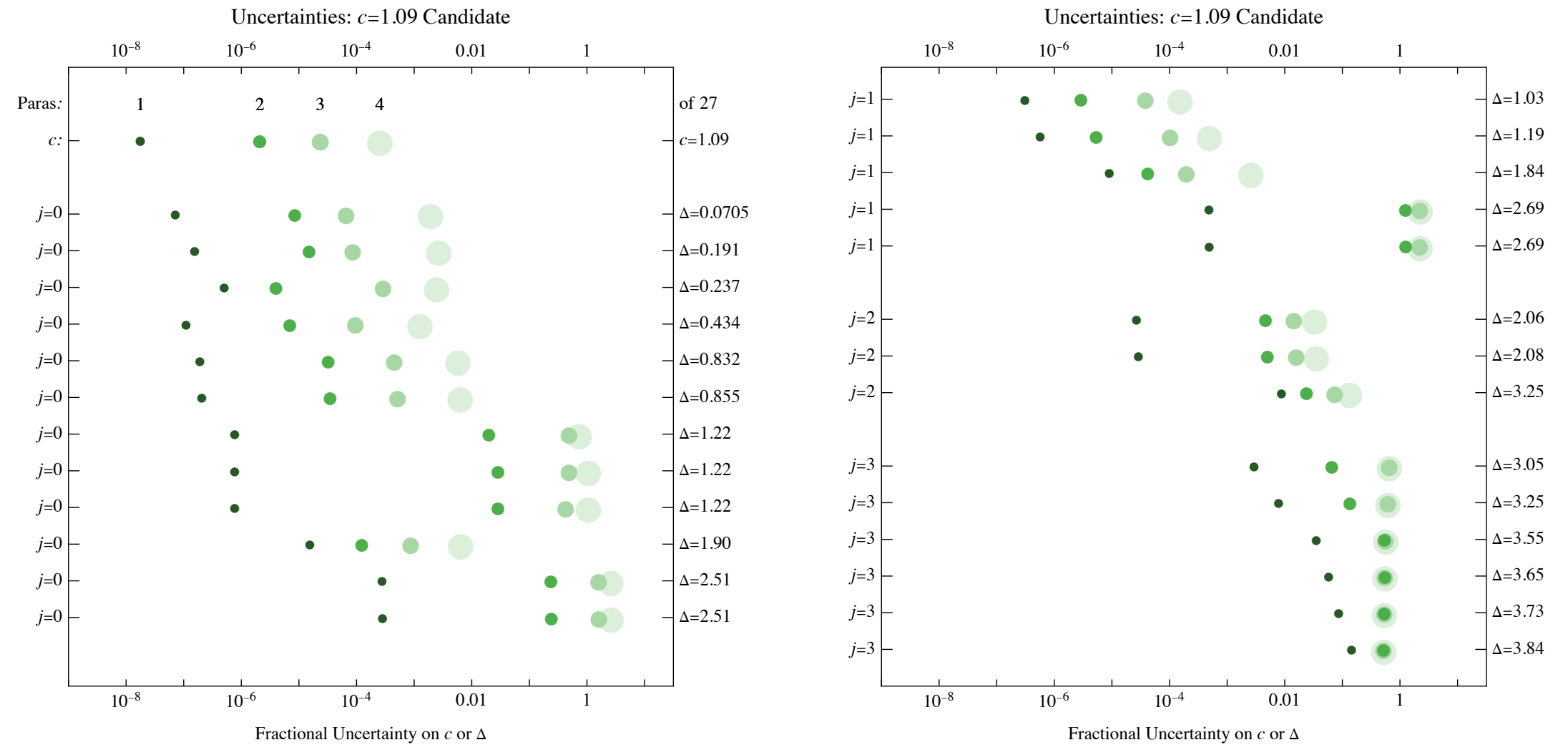
[Assi, Lee, Höche, JDT, [PRL 2025](#);
see Sudakov safety in Larkoski, Marzani, JDT, [PRD 2015](#);
see related moment construction in Desai, Nachman, JDT, [PRD 2024](#)]



Descending into the Modular Bootstrap



Tight constraints on individual operators...



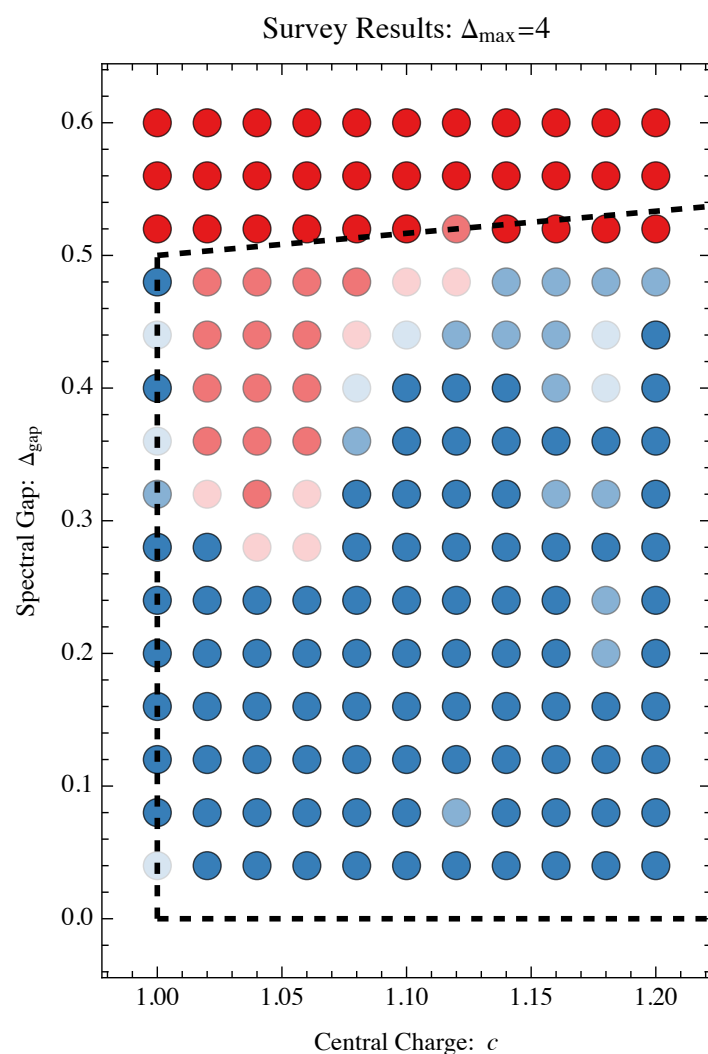
...almost certainly spanning a continuous space of solutions!

[Benjamin, Fitzpatrick, Li, JDT, [arXiv 2026](#); Bright-Thonney, Harvey, Lukas, JDT, [arXiv 2026](#)]

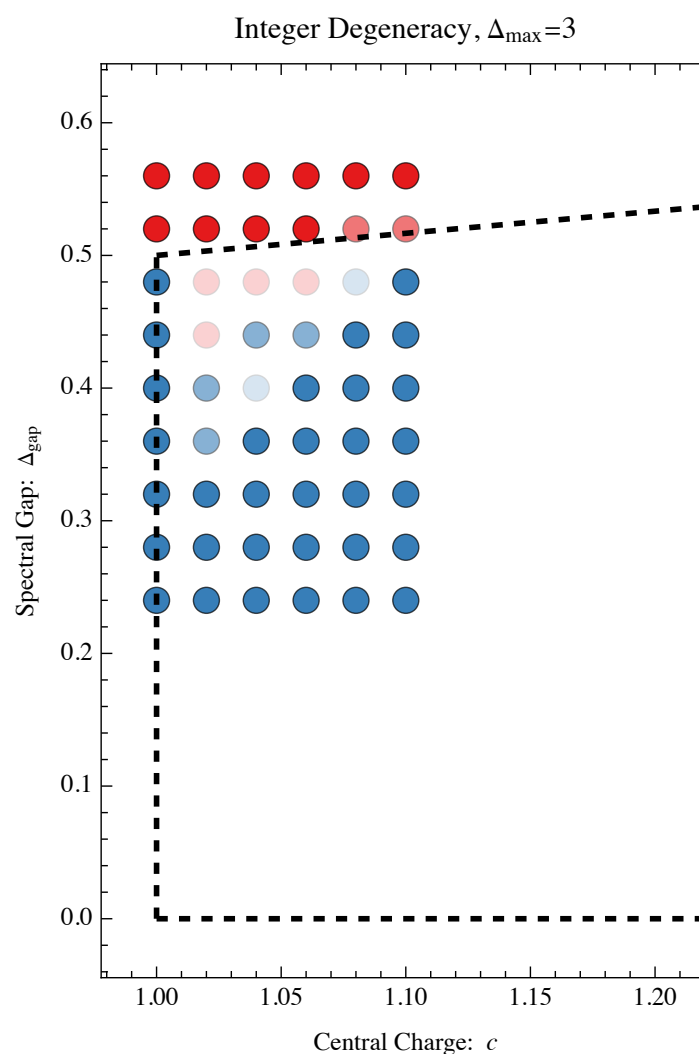
Cross Checking Modular Bootstrap Results (1 of 2)



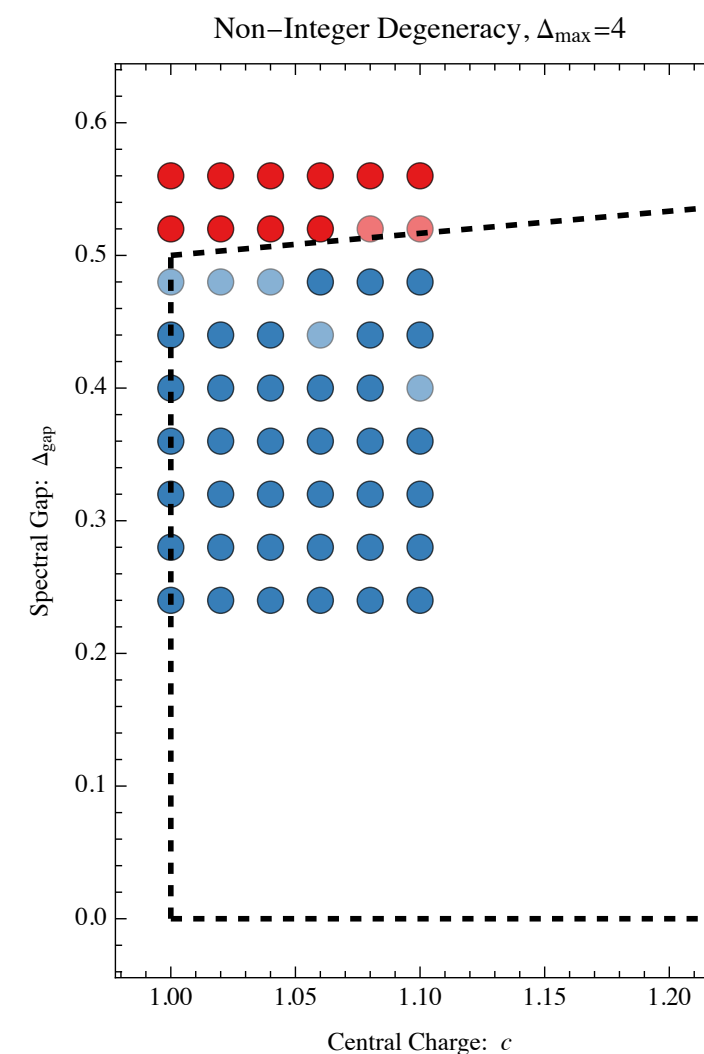
Main Result
with $\Delta_{\max} = 4$



Loosening Constraints
to $\Delta_{\max} = 4$

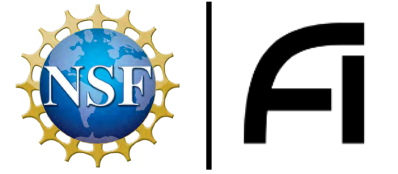


Allowing Non-Integer
Degeneracy with $\Delta_{\max} = 4$

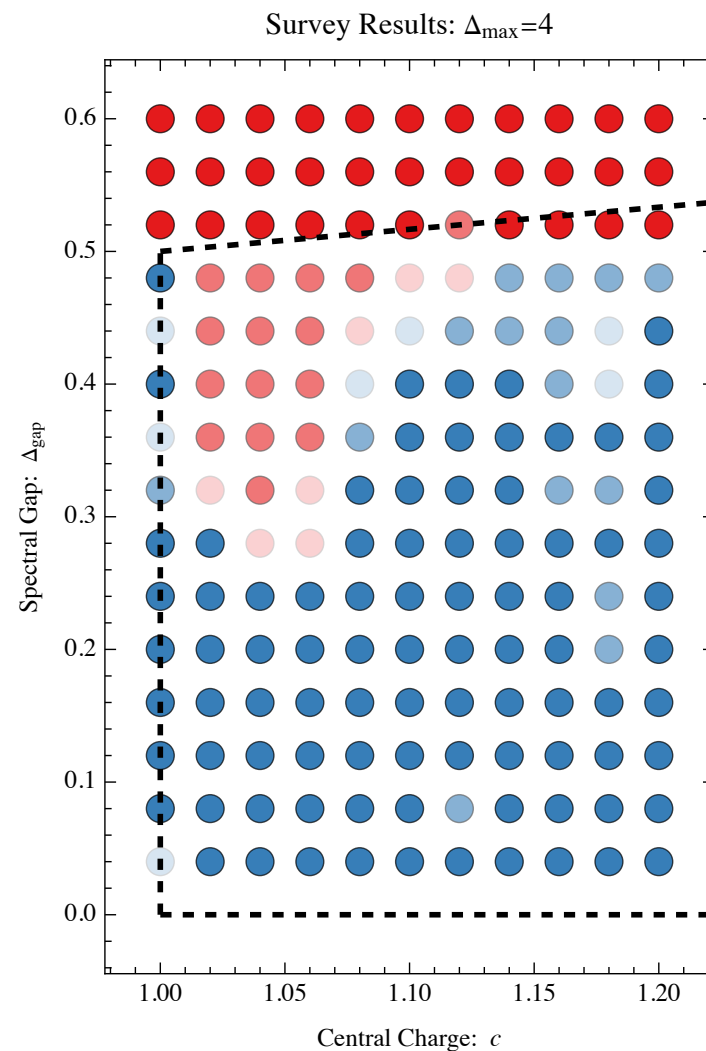


[Benjamin, Fitzpatrick, Li, JDT, [arXiv 2026](#); Bright-Thonney, Harvey, Lukas, JDT, [arXiv 2026](#)]

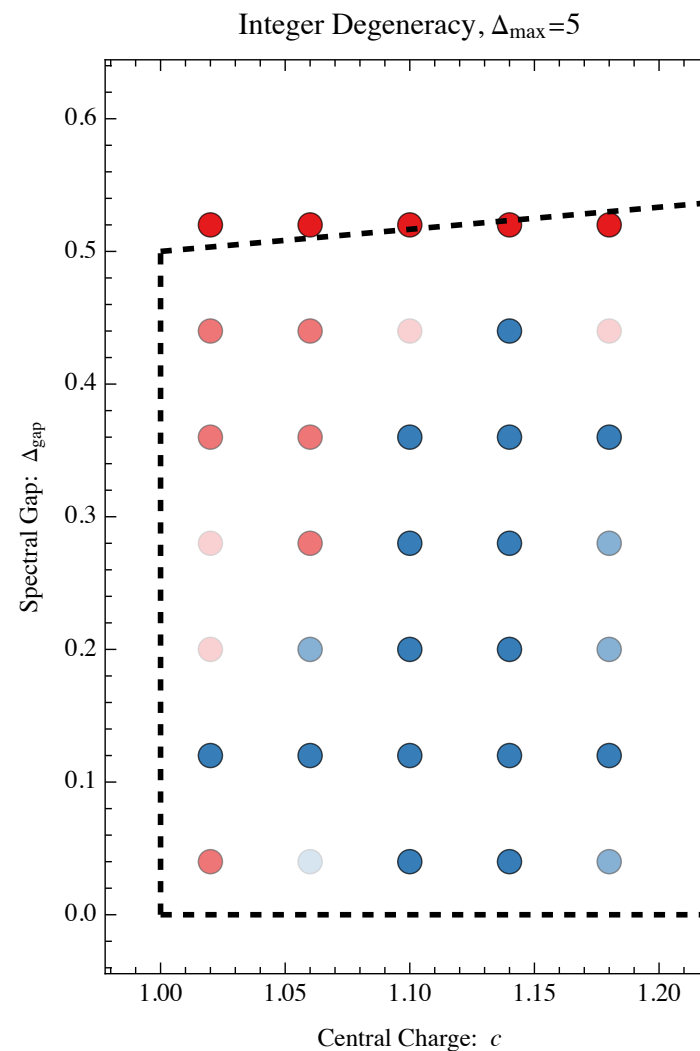
Cross Checking Modular Bootstrap Results (2 of 2)



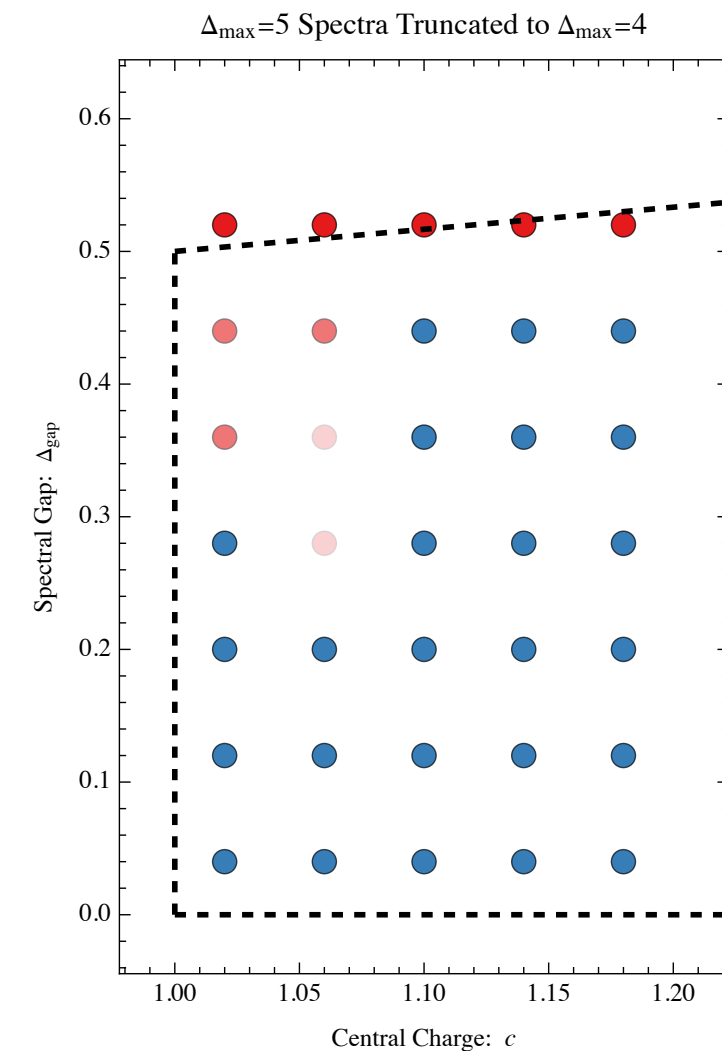
Main Result
with $\Delta_{\max} = 4$



Increasing Constraints
to $\Delta_{\max} = 5$



Evaluating $\Delta_{\max} = 5$
Spectra at $\Delta_{\max} = 4$



[Benjamin, Fitzpatrick, Li, JDT, [arXiv 2026](#); Bright-Thonney, Harvey, Lukas, JDT, [arXiv 2026](#)]

From AI Curmudgeon... to AI Evangelist!



MIT News
ON CAMPUS AND AROUND THE WORLD

Laboratory for Nuclear Science
August 26, 2020

National Science Foundation announces MIT-led Institute for Artificial Intelligence and Fundamental Interactions

IAIFI will advance physics knowledge — from the smallest building blocks of nature to the largest structures in the universe — and galvanize AI research innovation.

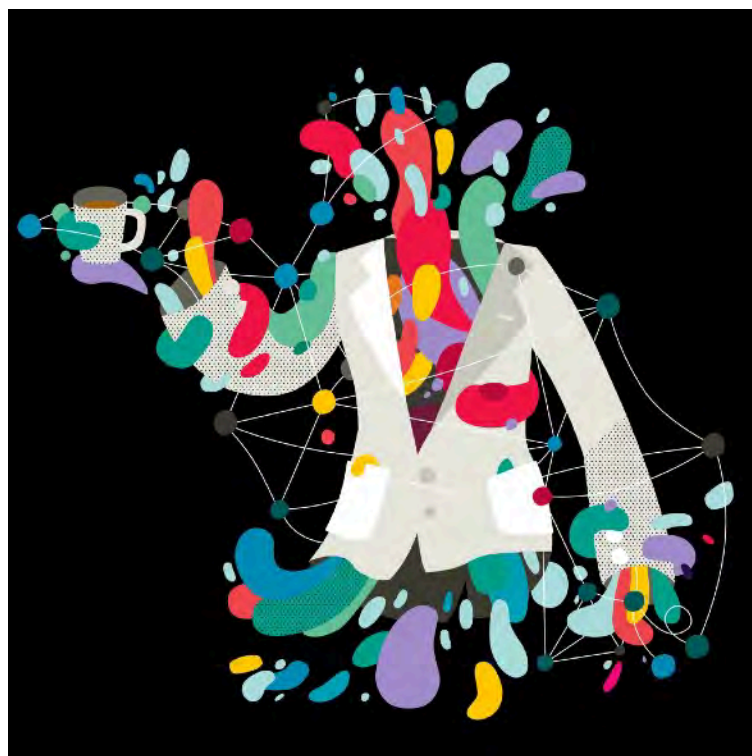
The New York Times



By **Dennis Overbye**

Nov. 23, 2020

Can a Computer Devise a Theory of Everything?



“In five to 10 years from now, I’m going to want to do exactly what you’re getting at: Here’s the data, here’s a very rough tool kit; find the equation I could put on a T-shirt, the equation that replaces the Standard Model of particle physics. What’s the equation that replaces Einstein’s general relativity?”

N.B. This was November 2020. ChatGPT was released November 2022.