

Machine Learning-Driven Anomaly Detection in Dijet Events with ATLAS

Or: *Where's Waldo?**



Tobias Golling, University of Geneva
On behalf of the ATLAS Collaboration

*

Where's Waldo? illustrations © 1987–2014 by Martin Handford. "Where's Waldo?"™ & © Classic Media Distribution Limited. All rights reserved.

Gets harder the *more* you have to scan





<http://atlas.ch>

10^{12} data events

Run: 280673

Event: 1273922482

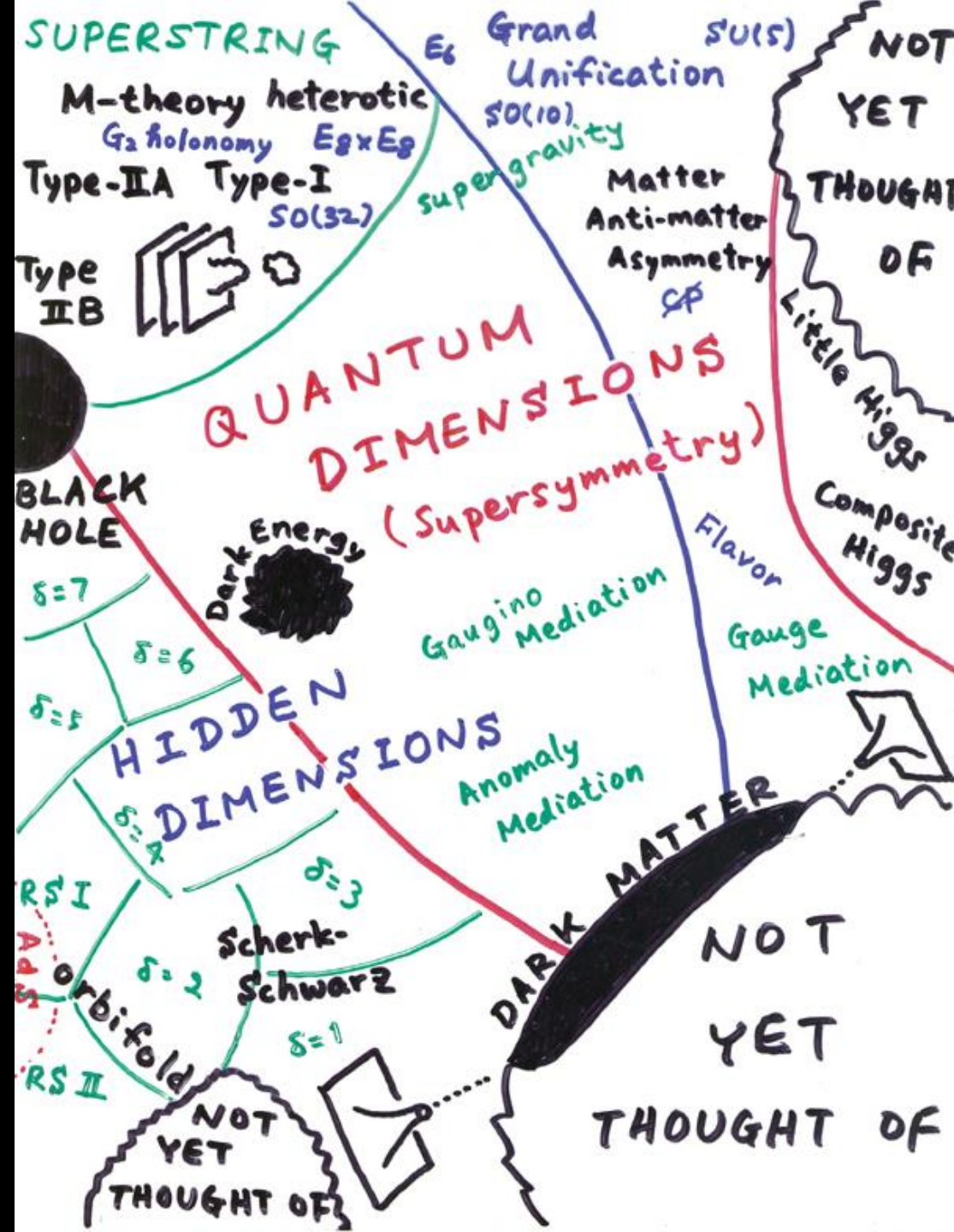
2015-09-29 15:32:53 CEST

Our data is *impenetrable* by the human brain

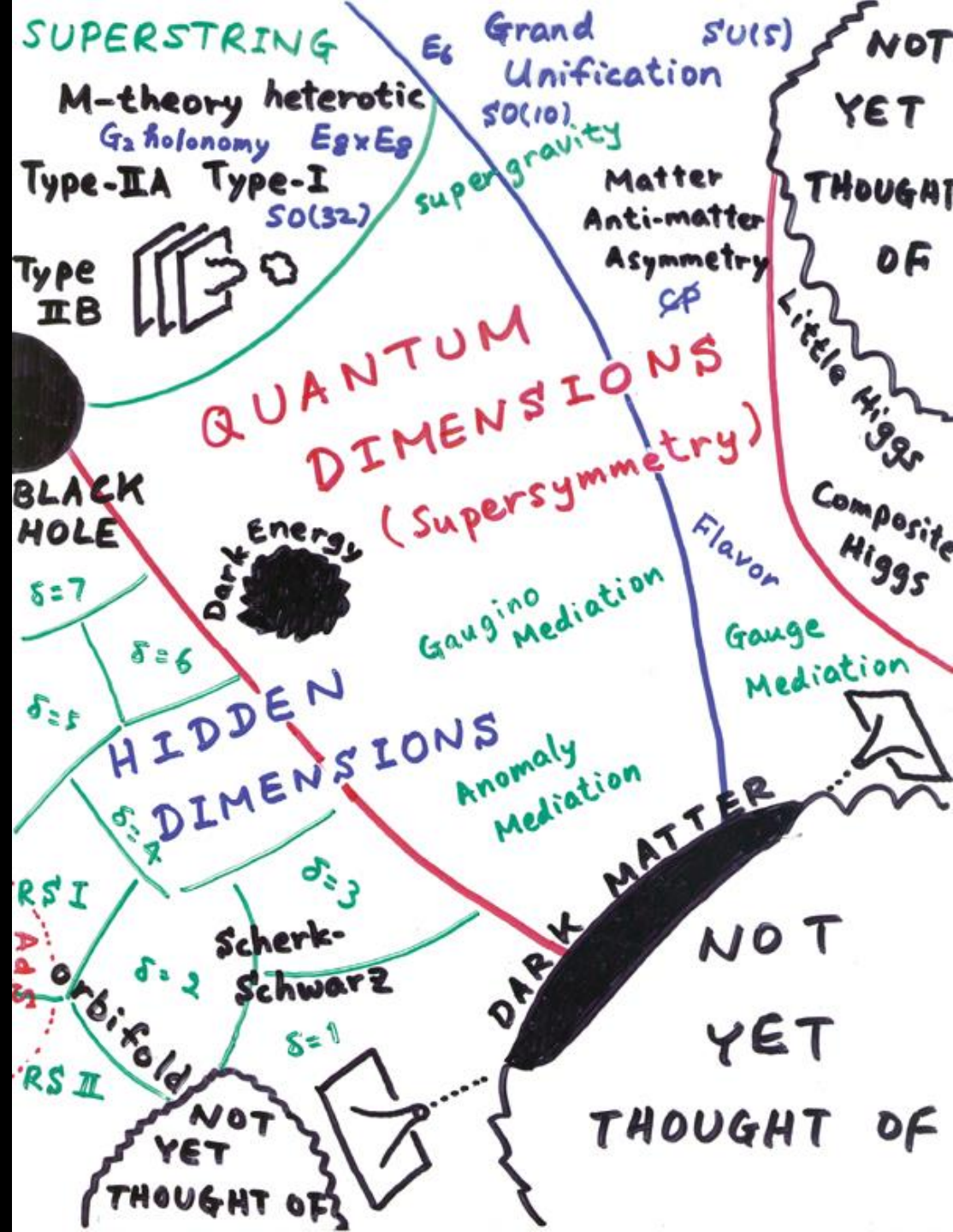
And we don't know *what* we're looking for



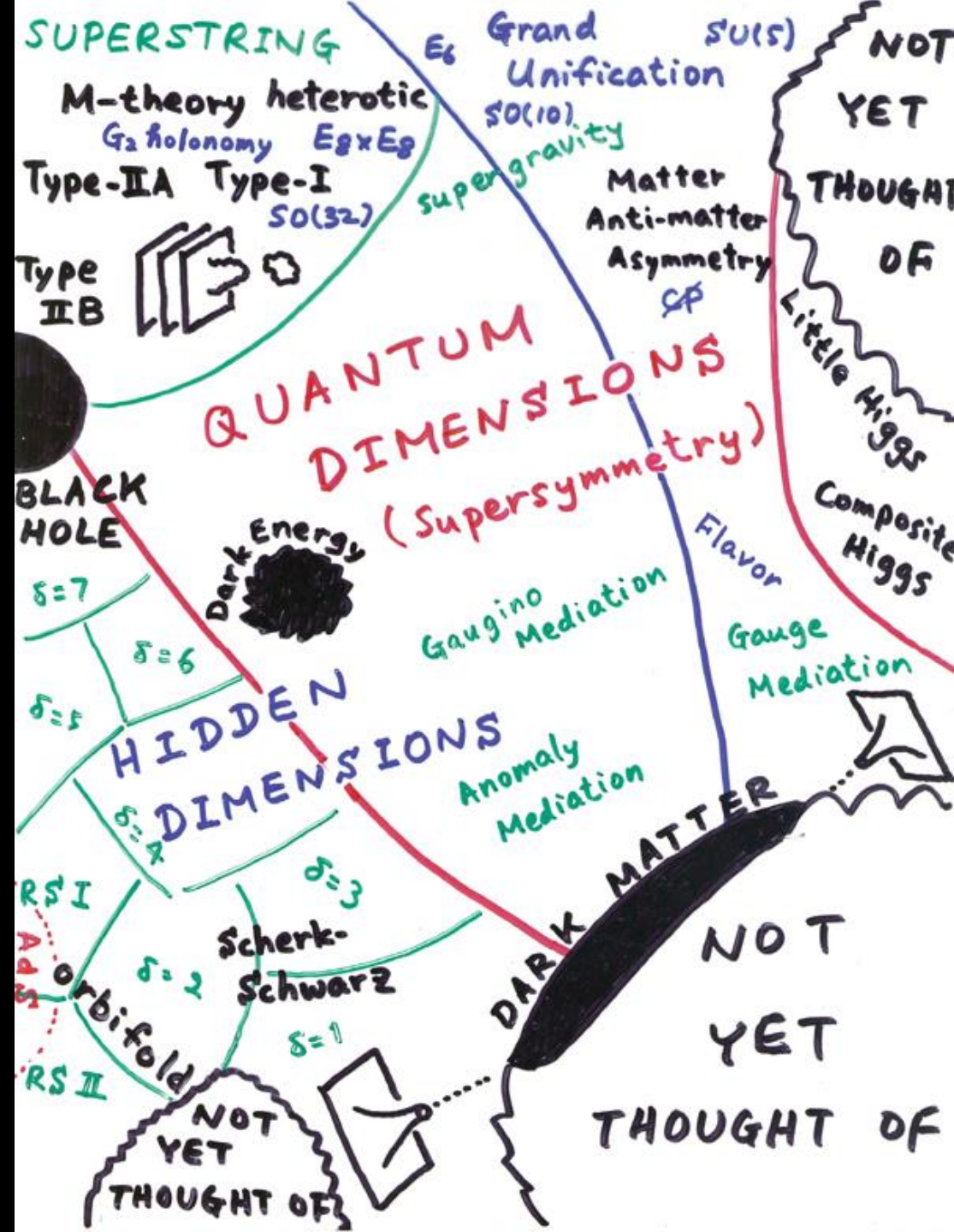
[ChatGPT]



↓



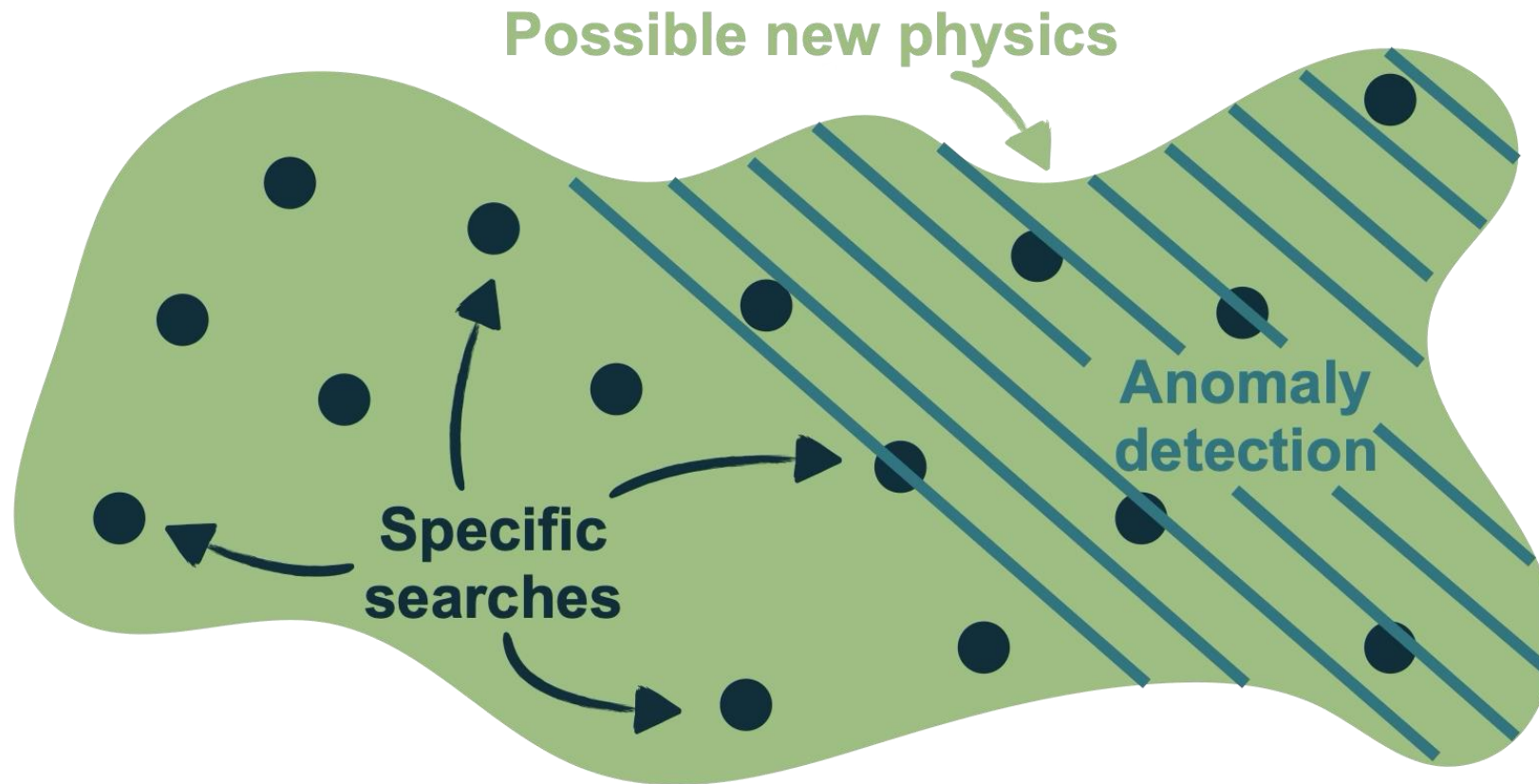
Nature
↓



*There is no
universally best
goodness of fit
test*

[Cousins 2016](#)

Trading power of test for coverage

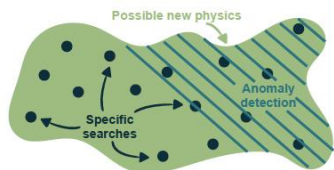


Machine Learning-Driven Anomaly Detection in Dijet Events with ATLAS

based on [1] (submitted to Phys. Review D)
Tobias Golling* and Dennis Noll** for the ATLAS Collaboration

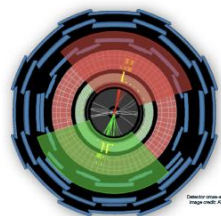
Motivation

- Want to find new physics in LHC collider data
- Many more BSM models than possible analyses
- Anomaly detection for resonant signals using machine learning (ML)
- Technique: Signal agnostic bump-hunt using many jet features
- Related analyses: [2, 3]

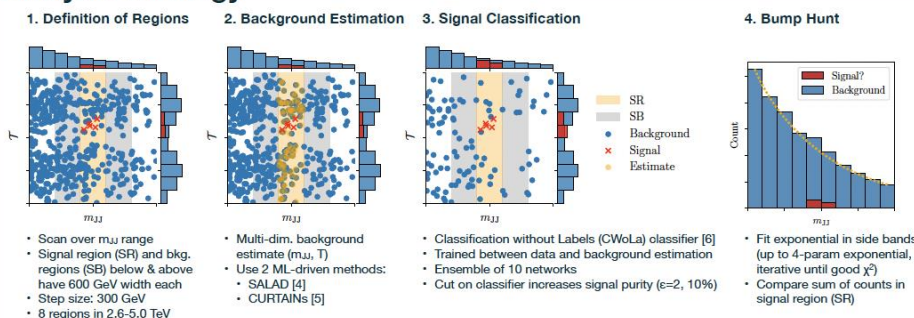


Dataset

- Using ATLAS data recorded 2015-2018 (139 fb⁻¹)
- Target events with ≥ 2 large radius jets (anti-kt, R=1) with low $\Delta\eta$
- Used jet features (T): Masses (m), Substructure (τ_{21} , τ_{32})



Analysis Strategy



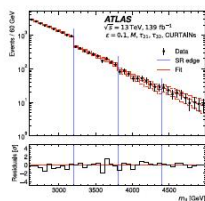
Inference & Results

Uncertainties (in order):

- Background fit
- Statistics
- Network ensemble

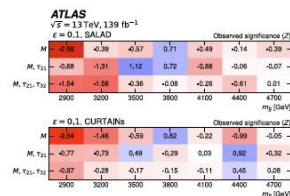
Validation

- Performed on ML-generated background-only dataset
- Successful for $m_{\text{LL}} > 2900$ GeV



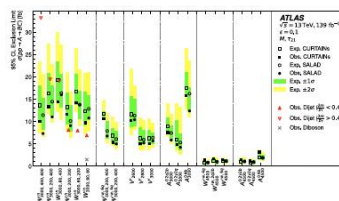
Signal Agnostic

- Observed significance of: 2 methods, 2 efficiencies, 7 m_{JJ} regions, 3 feature sets
- Largest significance is 1.24σ (1.26σ) & local deficit of -2.98σ (-2.54σ) for SALAD (CURTAINS)



Signal Specific

- Set limits @ 95% CLs to 20 investigated signal models
- Analysis has a broad performance on many models
- Similar performance for SALAD and CURTAINS
- Different feature sets have different sensitivity (more not always better) - scan over feature sets is one of the strengths of this analysis



Looking forward to
discussing with you in
Poster Session B
tomorrow !

[1] ATLAS Collaboration, Weekly supervised anomaly detection for resonant new physics in the dijet final state using proton-proton collisions at $\sqrt{s}=13$ TeV with the ATLAS detector, (2025), arXiv:2502.09770 [hep-ex].

[2] ATLAS Collaboration, Dijet Resonance Search with Deep Supervision Using $\sqrt{s}=13$ TeV pp Collisions in the ATLAS Detector, Phys. Rev. Lett. 125 (2020) 231801, arXiv:2005.02093 [hep-ex].

[3] CMS Collaboration, Model-agnostic search for dark resonances with anomalous jet substructure in proton-proton collisions at $\sqrt{s}=13$ TeV, (2024), arXiv:2412.01747 [hep-ex].

[4] A. Andreassen, B. Nachman, and D. B. Kaplan, Simulation Assisted Likelihood Free Anomaly Detection, Phys. Rev. D 101 (2020) 085004, arXiv:2001.06001 [hep-ph].

[5] A. Andreassen, B. Nachman, and D. B. Kaplan, Simulating the Flow of New Physics, Phys. Rev. D 102 (2020) 044004, arXiv:2006.04044 [hep-ph].

[6] M. Hasegawa, R. Nisikawa, and J. Tani, New Physics without Likelihood: Learning from Quantum Ensembles, JHEP 10 (2023) 134, arXiv:1709.05044 [hep-th].