



Contribution ID: 825

Type: **Parallel**

QCD in Language Models: What they really know about QCD?

This study presents an analysis of modern open-source large language models (LLMs)—including Llama, Qwen, and Gemma—to evaluate their encoded knowledge of Quantum Chromodynamics (QCD). Through reverse engineering of these models’ representations, we uncover the naturally idiosyncratic patterns in how foundational QCD concepts are embedded within their parameter spaces. Our methodology combines targeted probing techniques and knowledge extraction protocols to assess the models’ understanding of critical QCD principles like color confinement, asymptotic freedom, and the running coupling constant. We further demonstrate how these latent representations can be leveraged for “connecting dots” between important QCD concepts, potentially enabling novel insights since LLMs allow us to understand the problem holistically limited only by memory. This work provides a tool for utilizing LLMs as an assistant in theoretical physics research, while also highlighting current limitations in their representation of advanced quantum field theory concepts that future model development should address.

Secondary track

Author: SULC, Antonin

Session Classification: T16

Track Classification: T16 - AI for HEP (special topic 2025)