

Multi-Excitation Projective Simulation with a Many-Body Physics-Inspired Inductive Bias (Abstract Reprint)

Philip A. LeMaitre, **Marius Krumm** and **Hans J. Briegel**
University of Innsbruck, Institute for Theoretical Physics

Abstract Reprint. This is an abstract reprint of a journal article by [LeMaitre *et al.*, 2026].

Abstract

With the impressive progress of deep learning, applications relying on machine learning are increasingly being integrated into daily life. However, most deep learning models have an opaque, oracle-like nature making it difficult to interpret and understand their decisions. This problem led to the development of the field known as eXplainable Artificial Intelligence (XAI). One method in this field known as Projective Simulation (PS) models a chain-of-thought as a random walk of a particle on a graph with vertices that have concepts attached to them. While this description has various benefits, including the possibility of quantization, it cannot be naturally used to model thoughts that combine several concepts simultaneously. To overcome this limitation, we introduce Multi-Excitation Projective Simulation (mePS), a generalization that considers a chain-of-thought to be a random walk of several particles on a hypergraph. A definition for a dynamic hypergraph is put forward to describe the agent's training history along with applications to AI and hypergraph visualization. An inductive bias inspired by the remarkably successful few-body interaction models used in quantum many-body physics is formalized for our classical mePS framework and employed to tackle the exponential complexity associated with naive implementations of hypergraphs. We prove that our inductive bias reduces the complexity from exponential to polynomial, with the exponent representing the cutoff on how many particles can interact. We numerically apply our method to two toy environments and a more complex scenario modelling the diagnosis of a broken computer. These environments demonstrate the resource savings provided by an appropriate choice of inductive bias, as well as showcasing aspects of interpretability. A quantum model for mePS is also briefly outlined and some future directions for it are discussed.

References

[LeMaitre *et al.*, 2026] Philip A. LeMaitre, Marius Krumm, and Hans J. Briegel. Multi-excitation projective simulation with a many-body physics-inspired inductive bias. *Artificial Intelligence*, 352:104489, 2026.