Quantum Tensor Networks for Sampling
Abstract

Simon Botond

August 2025

Background and Motivation

Efficient sampling from complex probability distributions is a cornerstone of modern machine learning and artificial intelligence, enabling critical tasks such as probabilistic inference, generative modeling, and optimization. As the sophistication of models increases with richer logical structures or large graphical dependencies, the computational demands for sampling algorithms escalate, often growing exponentially in the number of variables and logical constraints. Classical sampling methods face prohibitive challenges in high-dimensional or multimodal spaces. This so-called curse of dimensionality highlights a fundamental limitation: the sample space grows exponentially, just like the time and resource cost of traditional procedures. This thesis investigates the potential of quantum algorithms - specifically, amplitude amplification - to overcome such limitations and deliver practical advantages for sampling tasks in artificial intelligence environments.

Data Representation in ML and AI

Data in Machine learning and Artificial intelligence systems can be encoded as numerical records, feature vectors, relational graphs, or tensor networks. Knowledge arises from explicit encoding of patterns, constraints, and dependencies, often through *logical formulas*, *probability distributions* or *probabilistic graphical models* (Bayesian networks, Markov Logic Networks). The choice and structure of representation have far-reaching implications for the performance and expressiveness of learning and inference algorithms. Both probabilistic and logical systems can be described by a set of properties, each called a categorical variable. A so-called *Atomic representation* of a system is described by the categorical variables X taking values x from a finite set:

$$[i] := \{0, 1, 2, \ldots, i-1\}$$

with cardinality i. A *Factored representation* of the same system is the set of categorical variables X_j where $j \in [d]$, taking values in i_j. Thus system states are assignments to a set of variables. Global properties (e.g., overall probability or formula satisfaction) are built up as combinations of local factors or constraints involving subsets of these variables.

Sampling: Concepts and Role in ML/AI

Sampling is an essential operation in machine learning and artificial intelligence workflows. Whether managing uncertainty in predictions, training deep generative models, or supporting decision-making under incomplete data, sampling is vital for probabilistic inference and uncertainty quantification. Examples for classical sampling algorithms are brute-force enumeration, rejection sampling, and Markov Chain Monte Carlo (MCMC) methods. Classical algorithms, while foundational, struggle with scalability. With increasing model/problem complexity and dimensionality, efficient sampling becomes a central bottleneck. The exponential growth in state space or the slow mixing times in complex distributions renders them intractable for practical applications. Classical solutions trade approximation accuracy for computational cost, leaving a gap open for fundamentally new approaches, with a possibility of offering speedups unreachable by classical computation.

Quantum Computing: Opportunities for Sampling

Quantum computing introduces new strategies for sampling via quantum mechanical principles, such as superposition, entanglement, and interference. Quantum algorithms can theoretically sample from complex distributions with fewer resources than their classical counterparts. Of particular interest is quantum amplitude amplification, a generalization of Grover's search algorithm, which provides quadratic speedup in identifying *target* solution states in unstructured search spaces. By encoding probability distributions into quantum states, this technique allows for efficient sampling that is intractable or even infeasible on classical hardware. Recent advances in quantum technology suggest that near-term devices could support such routines for intermediate-scale problems.

Theoretical Framework for data representation and sampling

The thesis formalizes sampling frameworks on knowledge bases through exponential family distributions and tensor networks. This formalization is one of the strengths of the algorithm, as by utilizing these 2 representations, and their techniques, any well known and used knowledge base can be mapped to quantum states and circuits. Therefore this mapping is robust and scalable.

- **Exponential Families**
  Exponential families are characterized by a canonical parameterization:

$$p(x) = \exp\{\langle \theta, \phi(x) \rangle - A(\theta)\}$$

Here, θ are the canonical parameters, $\varphi(x)$ sufficient statistics, and $A(\theta)$ the log-partition function normalizing the distribution. Many statistical models including Bernoulli, normal, and binomial, as well as complex logical connectives fall within this family, allowing for generalization of probabilistic and logical models. Mean parameterization is defined via μ, where the mean parameters are described as the expected values of the sufficient statistics under the distribution, $\mu = \mathbb{E}_p[\phi(x)]$. This representation provides dual insights for inference and learning.

- **Tensor Networks**
  Tensor networks encode high-dimensional probabilistic models into networks of lower-order tensors, capturing conditional dependencies and enabling efficient marginalization and sampling. Operations on tensor networks — contraction, decomposition, normalization — mirror the structure and computation of graphical models. Slice tensor decomposition is a method introduced to factorize joint distributions, enabling scalable representations amenable to quantum circuit mapping. It works in parallel with the factored representation of the knowledge bases we have, slicing them into smaller parts (slices), so the individual assignments and connections can be easily modelled as qubits and quantum circuits.

Mapping to Quantum Circuits

Quantum circuits are constructed from gates representing unitary operations on qubits - such as Hadamard ($\hat{H}$), Pauli($\hat{X}$, $\hat{Y}$, $\hat{Z}$), CNOT ($C\hat{X}$), Toffoli ($CC\hat{X}$) and rotation gates ($\hat{R}_P(\theta)$). Each variable from the sliced knowledge base is encoded as a qubit, with the structure translating into gate sets. This state preparation involves a Hadamard gate to create superpositions on variable qubits and single/double qubit gates to encode the relevant probabilities as amplitudes, with logical structure represented by a network of entangling multi qubit gates. The preparation is followed by oracle ($\mathbf{O}$) and diffusion ($\mathbf{D}$) operator sequences for amplitude amplification. The procedures leverage tensor-to-circuit compilation mapping, creating a systematic workflow from classical probabilistic models to scalable quantum sampling routines.

Amplitude Amplification—Quadratic Speedup

Grover's algorithm and its generalization, amplitude amplification, provide the essential quadratic speedup by rotating quantum states in a two-dimensional subspace defined by the good (*target*) and bad components. Oracle circuits mark solution states (e.g., those satisfying indicator function of), while the diffusion operator reflects amplitudes about the mean — inverting the quantum state toward heightened probability of good states. The process can be visualized geometrically: the initial quantum state forms an angle α with the good state subspace, where

$$P_g^0 = \sin^2\left(\frac{\alpha}{2}\right)$$

so $\alpha = 2\arcsin\left(\sqrt{P_g^0}\right)$. Each application of the amplitude amplification operator $\mathbf{Q} = \mathbf{OD}$ advances the state by a rotation of 2α within this subspace, increasing the amplitude on the target (*good*) states. After k applications, the probability of measuring a good state is

$$P_g^k = \sin^2\left((2k+1)\frac{\alpha}{2}\right)$$

This means the amplification process follows a sinusoidal pattern, with each iteration continuously boosting the probability for the target state.

The optimal number of amplification rounds k is found by maximizing P_g^k, reaching its peak when $(2k+1)\frac{\alpha}{2} \approx \frac{\pi}{2}$, i.e., when the rotated state aligns closest to the good subspace. Solving for k yields

$$k \approx \lfloor \frac{\pi}{4\sqrt{P_g^0}} \rfloor$$

demonstrating the core quantum advantage: the required number of oracle queries scales as $\mathcal{O}(1/\sqrt{P_g^0})$, compared to $\mathcal{O}(1/P_g^0)$ classically. This quadratic speedup is directly visible in the number of iterations needed to amplify a rare target state's probability to near unity.

Experimental Evaluation and Scalability

Simulations were undertaken using *Qiskit Aer* on up to 25 qubits, encoding logical formulas into quantum circuits. The slice tensor decomposition enabled scalable mapping, with amplification observed across rounds of Grover iteration, rapidly increasing the probability of sampling satisfying states. Empirical results confirmed that, for representative logical formulas encoded in exponential family form, only the canonical parameter evolves, while the sufficient statistics remain stable, demonstrating preservation of the exponential family and alignment with theoretical expectations.

The amplification process was benchmarked against classical sampling algorithms. Classical algorithms scale in $O(1/P)$ or exponentially with the number of variables, rendering them impractical at high dimensions or for rare-event sampling. Quantum amplification, by contrast, achieves quadratic improvement, with the turnover point for quantum advantage projected at $n \approx 50$ qubits for unstructured problems, and lower for structured or highly parallelizable connectives.

Resource Estimation and Hardware Implications

A detailed analysis of quantum and classical runtime scaling was performed, with hardware constraints evaluated for leading platforms. Trapped ion quantum computers were identified as especially suitable: all-to-all connectivity, high fidelity, deep circuit support, and scalability to dozens of qubits make them ideal for complex sampling routines. The required circuit depth per amplitude amplification iteration, overhead from state preparation and diffusion, and error mitigation strategies were quantified. Projected roadmap includes practical demonstrations in the 32–50 qubit regime within the next several years.

Discussion and Future Outlook

The thesis bridges quantum computing and machine learning by formalizing a framework for integrating quantum sampling into classical AI workflows. Symbolic and probabilistic reasoning are unified through tensor networks and exponential families, enabling direct mapping of logical connectives to quantum circuits and models. The empirical demonstration of quadratic speedup, benchmarking against classical methods, and resource estimation for current hardware provide actionable pathways toward practical deployment in scalable probabilistic inference.

Several open challenges remain: noise and decoherence in hardware, complexity in oracle circuit construction for nontrivial logical formulas, and limitations in quadratic speedup bounds. Future research directions include development of more robust tensor decompositions, automatic construction of efficient quantum oracles, and benchmarking on near-term quantum devices. Applications in high-dimensional sampling problems for finance, logistics, scientific computing, and hybrid quantum-classical AI pipelines are envisioned as areas of impact.

Conclusion

This thesis advances the frontier of quantum-enhanced sampling in machine learning, providing a systematic method for mapping classical probabilistic and logical models onto quantum states and circuits. By quantifying the resource demands, benchmarking against classical baselines, and analyzing the hardware landscape, it lays the groundwork for scalable, robust, and expressive sampling architectures where quantum computation becomes a core enabler for future artificial intelligence.