Neural Quantum States

For Quantum Optimal Control of Magic States

Alberto Berni
University of Southampton

UNIVERSITY OF SOUTHAMPTON
School of Electronics and Computer Science

Introduction & Motivation

The pursuit of universal quantum computation requires a g e set including non-Clifford gates (e.g., T-gate), as purely Clifford circuits are efficiently simulable classically [11]. The dominant paradigm for implementing such gates involves preparing and consuming fragile resource states, known as *magic states*, which must be generated at exceptionally high fidelities ($\sim 82.7\%$ for T-type states) to remain below the fault-tolerance threshold [4, 5, 13].

Quantum Optimal Control (QOC) methods used to find the required control pulses are often computationally intensive and rely on heuristic searches across complex landscapes [10, 8]. This work introduces a novel framework that leverages Neural Quantum States (NQSs) and differentiable programming to transform pulse design from a heuristic art into a systematic, automated science [2].

Gate	Equation	Matrix	Transform	Notation
Identity (***I***)	$\boldsymbol{I} = \lvert 0\rangle\langle 0\rvert + \lvert 1\rangle\langle 1\rvert$	$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	$\boldsymbol{I}\lvert 0\rangle = \lvert 0\rangle$ $\boldsymbol{I}\lvert 1\rangle = \lvert 1\rangle$	I
Pauli-*X* (***X*** or **NOT**)	$\boldsymbol{X} = \lvert 0\rangle\langle 1\rvert + \lvert 1\rangle\langle 0\rvert$	$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$	$\boldsymbol{X}\lvert 0\rangle = \lvert 1\rangle$ $\boldsymbol{X}\lvert 1\rangle = \lvert 0\rangle$	X
Hadamard (***H***)	$\boldsymbol{H} = \frac{\lvert 0\rangle+\lvert 1\rangle}{\sqrt{2}}\langle 0\rvert + \frac{\lvert 0\rangle-\lvert 1\rangle}{\sqrt{2}}\langle 1\rvert$	$\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$	$\boldsymbol{H}\lvert 0\rangle = \frac{1}{\sqrt{2}}(\lvert 0\rangle+\lvert 1\rangle)$ $\boldsymbol{H}\lvert 1\rangle = \frac{1}{\sqrt{2}}(\lvert 0\rangle-\lvert 1\rangle)$	H
Controlled-NOT (**CNOT**)	$\mathbf{CNOT} = \lvert 0\rangle\langle 0\rvert \otimes \boldsymbol{I} + \lvert 1\rangle\langle 1\rvert \otimes \boldsymbol{X}$	$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$	$\mathbf{CNOT}\lvert 00\rangle = \lvert 00\rangle$ $\mathbf{CNOT}\lvert 01\rangle = \lvert 01\rangle$ $\mathbf{CNOT}\lvert 10\rangle = \lvert 11\rangle$ $\mathbf{CNOT}\lvert 11\rangle = \lvert 10\rangle$	
Toffoli (***T*** or **CCNOT**)	$\mathbf{T} = \lvert 0\rangle\langle 0\rvert \otimes \boldsymbol{I} \otimes \boldsymbol{I} + \lvert 1\rangle\langle 1\rvert \otimes \mathbf{CNOT}$	$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}$	$\mathbf{T}\lvert 000\rangle = \lvert 000\rangle, \mathbf{T}\lvert 001\rangle = \lvert 001\rangle$ $\mathbf{T}\lvert 010\rangle = \lvert 010\rangle, \mathbf{T}\lvert 011\rangle = \lvert 011\rangle$ $\mathbf{T}\lvert 100\rangle = \lvert 100\rangle, \mathbf{T}\lvert 101\rangle = \lvert 101\rangle$ $\mathbf{T}\lvert 110\rangle = \lvert 111\rangle, \mathbf{T}\lvert 111\rangle = \lvert 110\rangle$	

Figure 1. Universal gate sets require non-Clifford gates (e.g., T-gate) implemented via magic states.

The NQS Paradigm

A Neural Quantum State (NQS) is a variational ansatz where a neural network, parameterized by θ, represents the complex coefficients $\psi_\theta(\sigma)$ of a many-body quantum state [7]. Its key advantage is **expressivity**; unlike tensor networks (e.g., MPS) efficient for "area-law" entanglement, NQS can compactly represent the "volume-law" states common in 2D systems and quench dynamics [11, 15].

For fermionic systems like superconducting qubits, the NQS is defined in the occupation number basis $\{|n\rangle\}$, mapping a configuration of occupied orbitals to an amplitude $\psi_\theta(n)$. The network is trained via Variational Monte Carlo (VMC): configurations are sampled from the probability distribution $|\psi_\theta(n)|^2$ using MCMC, and the network parameters θ are optimized to minimize the stochastically estimated energy.

NQSs for Fermionic Systems

Fermionic NQS Ansatz: $|\psi_\theta\rangle = \sum_n \psi_\theta(n)|n\rangle$

Second-Quantized Hamiltonian: $\hat{H} = \sum_{ij} t_{ij} c_i^\dagger c_j + \sum_{ijkl} V_{ijkl} c_i^\dagger c_j^\dagger c_k c_l$

Local Energy Estimator: $E_{\text{loc}}(n) = \frac{1}{\psi_\theta(n)} \sum_{n'} \langle n|\hat{H}|n'\rangle \psi_\theta(n')$

Variational Energy (MC Estimate): $E_\theta = \langle E_{\text{loc}}(n) \rangle_{n\sim|\psi_\theta|^2} \geq E_{gs}$

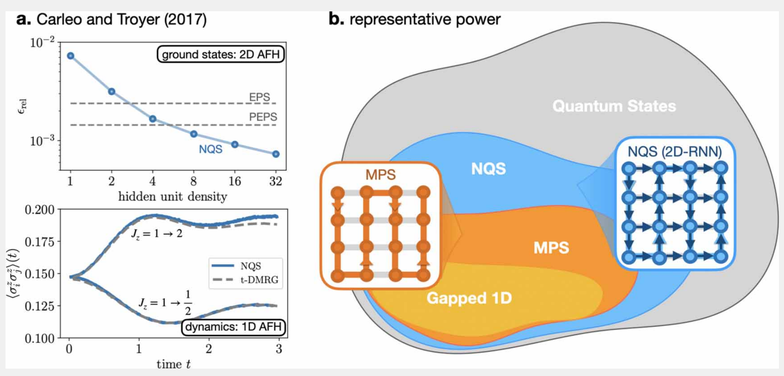

Figure 2. An NQS maps a physical configuration to its wave function coefficient. The network's high internal connectivity (right) allows it to capture complex "volume-law" entanglement, surpassing the "area-law" limitations of 1D methods like Matrix Product States (left). Adapted from Lange et al. [11, 7].

Conclusion & Future Work

This project introduces a scalable, automated pipeline for pulse engineering that bridges machine learning and quantum control. Our framework aims to significantly accelerate the development of robust, high-fidelity quantum gates essential for fault-tolerance [2].

Future Work

- Incorporate realistic hardware noise models into the simulation to discover control solutions that are inherently more robust [2].
- Deploy the discovered pulse sequences on a physical photonic quantum processor for experimental validation and tomographic verification [2].

Key References

1. Hannah Lange, Anka Van de Walle, Atiye Abedinnia, and Annabelle Bohrdt. From architectures to applications: A review of neural quantum states.
arXiv preprint, 2024
2. Giuseppe Carleo and Matthias Troyer. Solving the quantum many-body problem with artificial neural networks.
Science, 355(6325):602–606, 2017
3. Ville Bergholm et al. Pennylane: Automatic differentiation of hybrid quantum-classical computations, 2018
4. Navin Khaneja, Timo Reiss, Christoph Kehlet, Thomas Schulte-Herbrüggen, and Steffen J. Glaser. Optimal control of coupled spin dynamics: design of nmr pulse sequences by gradient ascent algorithms.
Journal of Magnetic Resonance, 172:296–305, 2005

A full bibliography is available via the QR code below.

NQS Architectures & Methods

The choice of NQS architecture imparts an *inductive bias*; a set of assumptions about the quantum state it can most efficiently represent [11]. A survey of prominent architectures includes:

Restricted Boltzmann Machines (RBMs): The pioneering NQS architecture, these are energy-based models analogous to an Ising model with visible (physical) and hidden (learned) units. Their all-to-all connectivity between layers allows them to efficiently capture the long-range correlations found in volume-law entangled states [7, 11].

Convolutional Neural Networks (CNNs): Ideal for systems on regular lattices. The convolution operation inherently respects the system's translational symmetry, significantly reducing the number of required variational parameters while capturing local correlations effectively [11].

Autoregressive Networks (RNNs, Transformers): A strong candidate for this project. These models represent the wave function's probability distribution as a product of conditional probabilities, $p(\sigma) = \prod_i p(\sigma_i|\sigma_{<i})$. This structure has a key advantage: it allows for **perfect and efficient sampling** of configurations, bypassing the computationally expensive Markov Chain Monte Carlo (MCMC) methods required by other architectures. Transformers, in particular, use a **self-attention mechanism** to capture global correlations across the entire system, making them exceptionally well-suited for highly entangled states [11].

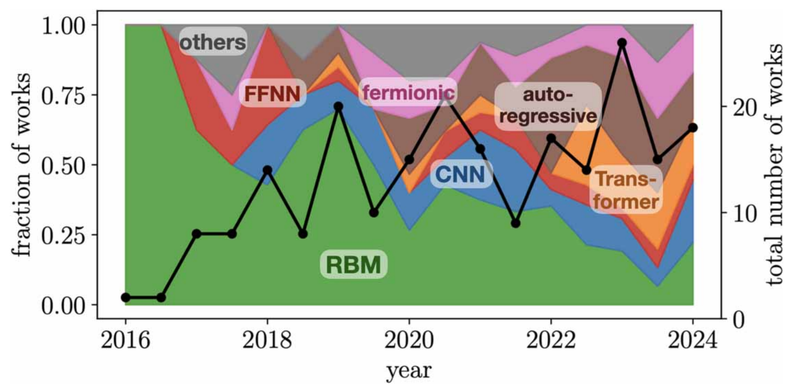

Figure 3. Popularity of NQS architectures over time, showing a recent rise in autoregressive models like Transformers. Adapted from Lange et al. [11].

A Differentiable Framework for QOC

We propose a closed-loop optimization framework that recasts the QOC problem into an end-to-end differentiable computational graph, solvable using modern deep learning machinery [2, 14].

State Representations

The framework utilizes two distinct NQSs:

- **Target NQS** $|\psi_{\theta_{target}}\rangle$: A static, pre-trained network that serves as a high-fidelity, differentiable model of the desired magic state [2].
- **Evolving NQS** $|\psi_{\theta(t)}\rangle$: A dynamic network whose parameter evolution is governed by the time-dependent Schrödinger equation, simulated via a time-dependent Variational Monte Carlo (t-VMC) method [6].

The Differentiable Cost Function

The core objective is to minimize the infidelity between the final evolved state and the target state at time T. This cost function is defined with respect to the classical pulse parameters $\{\alpha_k\}$:

$$\mathcal{L}(\{\alpha_k\}) = 1 - |\langle\psi_{\theta_{target}}|\psi_{\theta(T)}\rangle|^2$$

Automated Pulse Discovery via Backpropagation

The control pulse, defined by parameters $\{\alpha_k\}$, shapes a time-varying Hamiltonian $H(\{\alpha_k\}, t)$. Because the entire simulation is a single computational graph built in a framework like PennyLane [1], we can use backpropagation to efficiently compute the exact physical gradient of the cost function with respect to the pulse parameters, $\nabla_{\{\alpha_k\}}\mathcal{L}$. An optimizer (e.g., Adam) then automatically updates $\{\alpha_k\}$ to systematically minimize the infidelity [2].

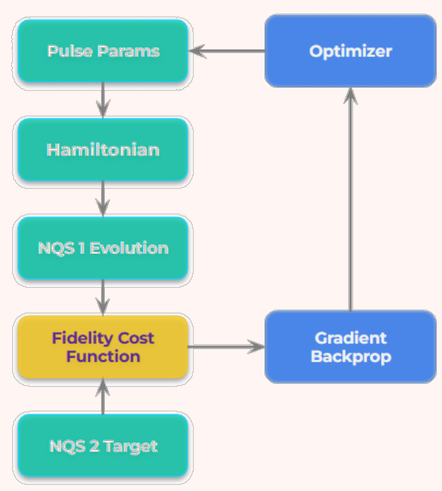

Figure 4. The closed-loop differentiable framework. Gradients of the infidelity are back-propagated through the time evolution to automatically update the classical pulse parameters.

Implementation Plan & Validation

Technology Stack

Our framework is built using **PennyLane**, a library for quantum machine learning and automatic differentiation, with **PyTorch** providing GPU/TPU acceleration and NQS model construction [1, 12, 3].

Validation Strategy

We employ a rigorous, multi-step validation process:

1. **Benchmarking**: The framework is first validated on simple problems with known analytical solutions, such as discovering a $\pi/2$ pulse for a single qubit and a CNOT gate for a two-qubit system [3].
2. **Cross-Validation**: Optimized pulses are benchmarked against established QOC algorithms like GRAPE using an independent, industry-standard physics simulator (**QuTiP**) to ensure an unbiased, apples-to-apples comparison [9, 3].

PyTorch PENNYLANE

Contact: ab3u21@soton.ac.uk

Workshop BIDW02 — Isaac Newton Institute, October 2025