Introduction to Quantum Machine Learning and Neural Networks (Part 2)

Nathan Thompson

Department of Mathematics, Statistics, and Physics
Wichita State University

Quantum Technologies for Fundamental Physics: Global Networks
Lawrence, KS May 28, 2026

WSU
Wichita State University

Warm-up:
Quantum Encoding

WSU
Wichita State University

Basis Encoding

- Encode data as computational basis states
- Examples:
  - Bit string:
    $x = (1, 0, 1) \rightarrow |101\rangle$
  - Number as binary:
    $6 = 4 + 2 + 0 = 1 \cdot 2^2 + 1 \cdot 2^1 + 0 \cdot 2^0 = 110_2 = |110\rangle$
- Typically one qubit per feature

Pros

- Simple
- Implementation is straightforward

Cons

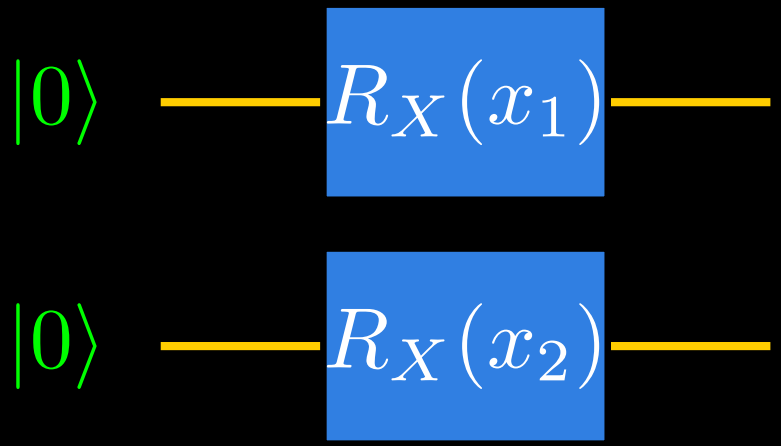
- Poor for continuous data
- No natural notion of similarity

WSU
Wichita State University

Angle Encoding

- Suppose $x = (x_1, x_2)$
- Angle encoding:

$$x \to R_X(x_1)\,|0\rangle \otimes R_X(x_2)\,|0\rangle$$



Pros

- Natural for continuous data
- Efficient and hardware friendly
- Effective with shallow circuits

Cons

- One rotation per feature
- Limited expressivity without entangling layers
- Capturing global correlations requires deep circuits

WSU
Wichita State University

Amplitude Encoding

- Encode features as state amplitudes
- Example:
  - $x \rightarrow \sum_i \frac{x_i}{\|x\|} |i\rangle$

Pros

- Most compact
- Full vector stored as single state

Cons

- State preparation is expensive
- Difficult to implement in practice

WSU
Wichita State University

Example: Amplitude Encoding

- Scheme:
  - $x \to \sum_i \frac{x_i}{\|x\|} |i\rangle$
  - $|i\rangle$ is i in binary
- Suppose $x = (x_0, x_1, x_2, x_3) = (1, 2, -1, 1)$
- Normalize x:

$$\|x\| = \sqrt{1^2 + 2^2 + (-1)^2 + 1^2} = \sqrt{7}$$

- Encode:

$$|\psi(x)\rangle = \tfrac{1}{\sqrt{7}} |00\rangle + \tfrac{2}{\sqrt{7}} |01\rangle - \tfrac{1}{\sqrt{7}} |10\rangle + \tfrac{1}{\sqrt{7}} |11\rangle$$

- Cost: $\lfloor \log_2(n) \rfloor + 1$ qubits (the bit length of n)

WSU
Wichita State University

XOR Example

WSU
Wichita State University

XOR

Classic Machine Learning Example

- Requires a non-linear separator
- Not solvable by a (single-layer) perceptron

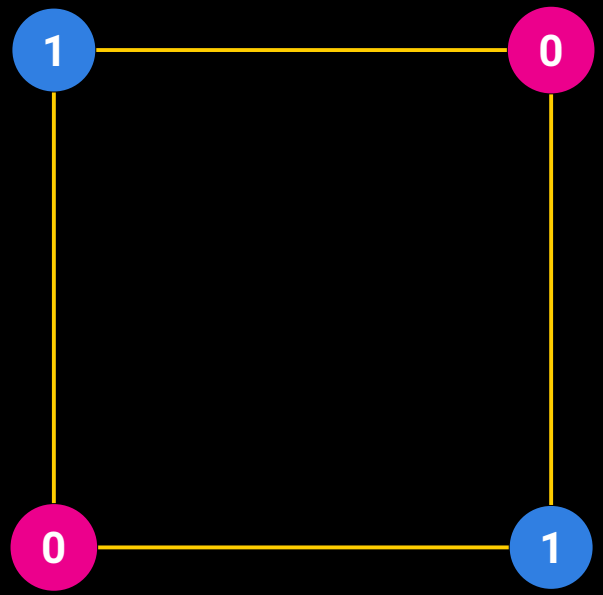
Input	Output
0 & 0	0
0 & 1	1
1 & 0	1
1 & 1	0

WSU
Wichita State University

XOR - Visualization

Classic Machine Learning Example

- Requires a non-linear separator
- Not solvable by a (single-layer) perceptron

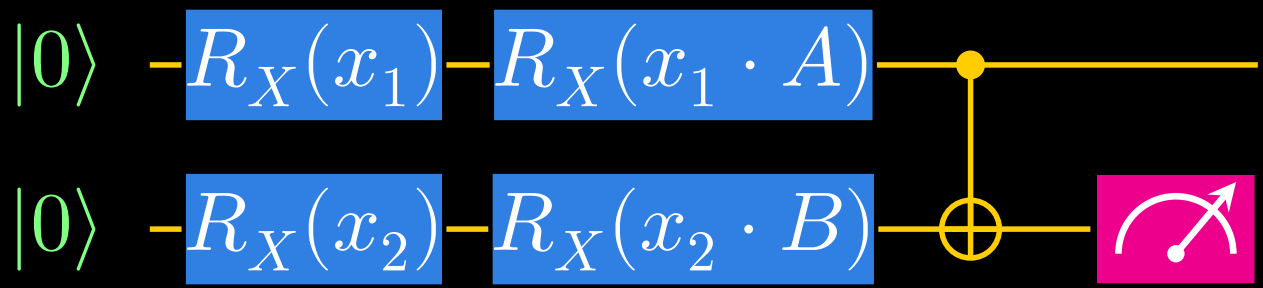


WSU
Wichita State University

XOR - MATLAB Implementation

Adapted from

`https://www.mathworks.com/help/matlab/math/solve-xor-problem-using-quantum-neural-network.html`



WSU
Wichita State University

Completing the Model

A Quantum Machine Learning model has three components:

- Encoding (already covered)
- Circuit (transforms the state)
- Measurement (produces output)

Our Next Question:

What do the circuit and measurement add to the model?

WSU
Wichita State University

Quantum Machine Learning Pipeline

$$x \to |\psi(x)\rangle \to U(\theta)\,|\psi(x)\rangle \to \text{measurement} \to f(x, \theta)$$

- Encoding defines the representation
- Circuit transforms that representation
- Measurement extracts information

Together, these define the model $f(x, \theta)$

WSU
Wichita State University

Parametrized Quantum Circuits

WSU
Wichita State University

Parametrized Quantum Circuits

- A quantum circuit with tunable parameters θ
- Applies a unitary transformation:

$$U(\theta)\,|\psi(x)\rangle$$

- Parameters control gate operations
- The circuit structure determines what transformations are possible

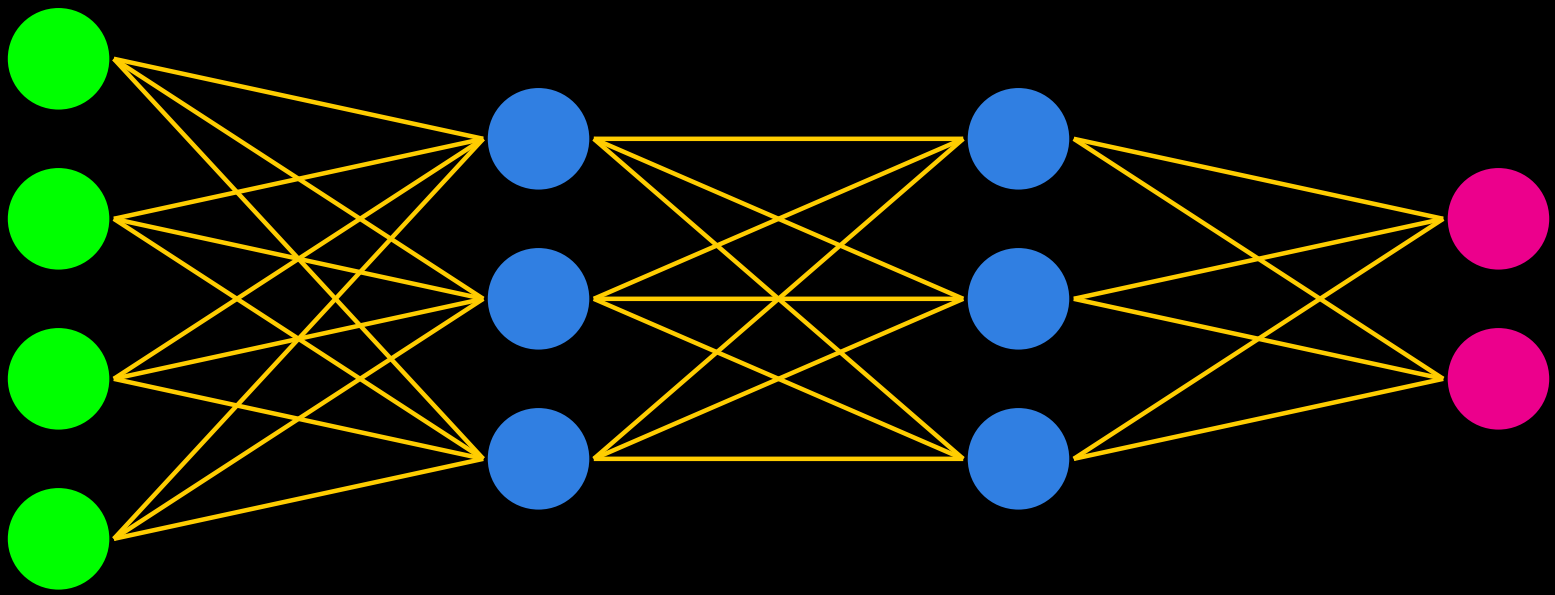
Connection to CML

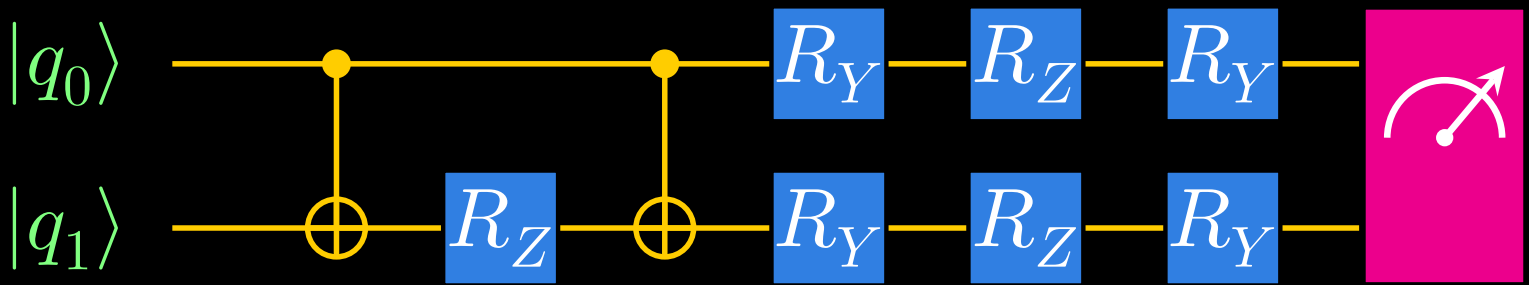
Parameters act like the weights in a neural network

WSU
Wichita State University

Classical & Quantum Neural Nets

Parametrized transformations of data





Wichita State University

Output from Quantum Model

- After applying the circuit, the state is:

$$U(\theta)\,|\psi(x)\rangle$$

- We measure an observable O
- Output is an expectation value:

$$f(x,\theta) = \langle\psi(x)|U^{\dagger}(\theta)OU(\theta)|\psi(x)\rangle$$

Connection to Classical

This is the quantum version of $f(x, W)$

WSU
Wichita State University

Quantum ML Models

- A QML model is defined by:
  - Encoding $x \to |\psi(x)\rangle$
  - Circuit $U(\theta)$
  - Measurement O

Key Point

All three together define the hypothesis class

WSU
Wichita State University

Quick Check-in

If we fix the encoding but change $U(\theta)$,
are we using the same model?

No!

Changing the circuit changes the transformation and thus the hypothesis class.

WSU
Wichita State University

Variational Quantum Algorithms

WSU
Wichita State University

Shifting to Algorithms

- We've discussed:
  - How to encode data
  - A parametrized quantum model $f(x, \theta)$
- What kinds of computations are possible?

Variational Quantum Algorithms

WSU
Wichita State University

Variational Quantum Algorithms

Elements of a VQA:

- Hybrid of quantum and classical methods
- Parametrized quantum circuits
- Circuit parameters are optimized using a classical algorithm

Hybrid System

Quantum computer evaluates the function,
classical computer optimizes the function

WSU
Wichita State University

Variational Computation Loop

- Initialize parameters θ
- Prepare quantum state
- Apply circuit $U(\theta)$
- Measure to compute cost $C(\theta)$
- Update θ using a classical optimizer
- Repeat

WSU
Wichita State University

Quantum Learning

Recall

"Learning" is optimization

- Define a cost function

$$C(f(x, \theta), y)$$

- $C(\theta)$ is computed via measurements on a quantum device
- Find

$$\min_{\theta} C(\theta)$$

- Circuit structure completes the model design
- Optimization determines the best parameters

WSU
Wichita State University

Quick Check-in

What role does the quantum computer play in the learning loop?

Answer

The quantum computer evaluates the model by computing $C(\theta)$

WSU
Wichita State University

Why a Hybrid Approach?

- Current quantum hardware is limited:
  - Noisy
  - Small number of qubits (∼100-1000)
  - Noisy Intermediate-Scale Quantum (NISQ) era
- Fully quantum algorithms are difficult to run on hardware
- This approach enables learning on near-term quantum devices

WSU
Wichita State University

Variational Quantum Eigensolvers (VQE)

WSU
Wichita State University

Variational Quantum Eigensolver

- Standard NISQ routine
- **Goal**: find the lowest energy state of a quantum system
- Corresponds to finding the ground state of a Hamiltonian H

Main Benefit

Many problems in physics, chemistry, and materials science can be expressed as finding the ground state of a Hamiltonian

WSU
Wichita State University

Setting Up a VQE

- We want to compute the minimum energy:

$$E_0 = \min_{\theta} \langle \psi(\theta) | H | \psi(\theta) \rangle$$

  - $|\psi(\theta)\rangle$ is prepared using a parametrized circuit
  - H is the Hamiltonian of the system

- Optimal parameter values $\hat{\theta}$ approximate the ground state of H

WSU
Wichita State University

The VQE Loop

- Prepare an initial state $|\psi(\theta)\rangle$
- Measure the expectation value

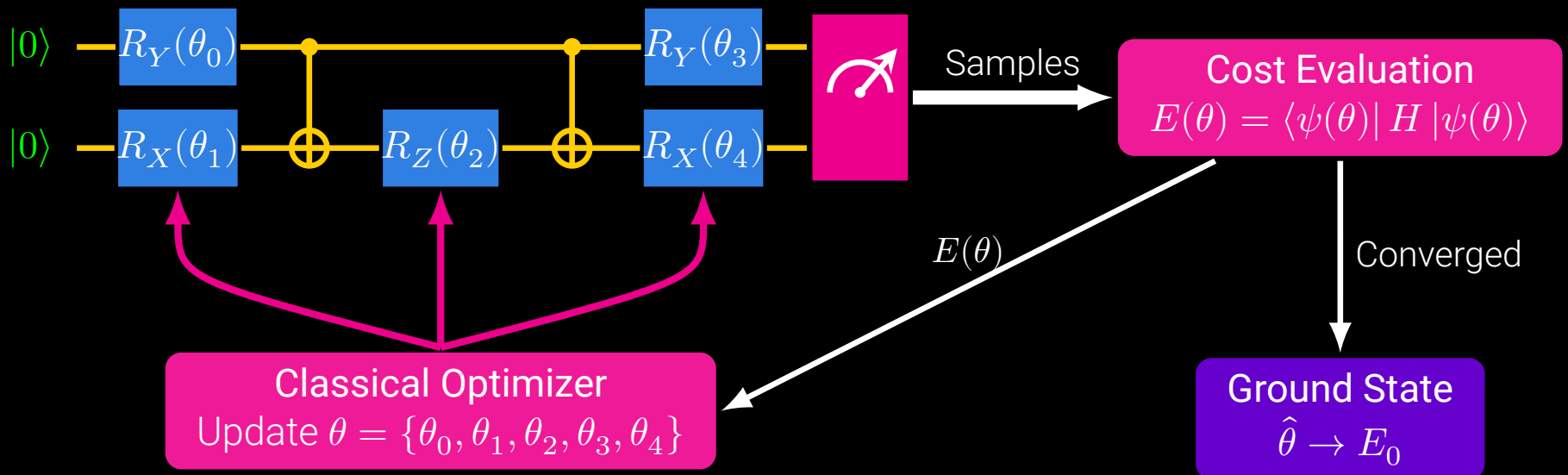
$$\langle\psi(\theta)|H|\psi(\theta)\rangle$$

- Use a classical optimizer to update θ
- Repeat until the system converges

WSU
Wichita State University

The VQE Loop - Diagram

1. State Preparation & Measurement



2. Classical Optimization Loop

WSU
Wichita State University

Quick Check-in

What is the cost function in VQE?

Answer

The expected energy:

$$\langle \psi(\theta) | H | \psi(\theta) \rangle$$

WSU
Wichita State University

VQE - MATLAB Implementation

- Pauli operators

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

$$Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \qquad Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

- Construct H as a combination

$$H = 0.2I \otimes I + 0.05I \otimes Z + 0.05Z \otimes I + 0.7Z \otimes Z + 0.15X \otimes X - 0.05Y \otimes Y$$

WSU
Wichita State University

Quantum Approximate Optimization Algorithm (QAOA)

Wichita State University

From VQE to QAOA

- VQE uses a general-purpose quantum circuit
- Designed for problems like finding ground states

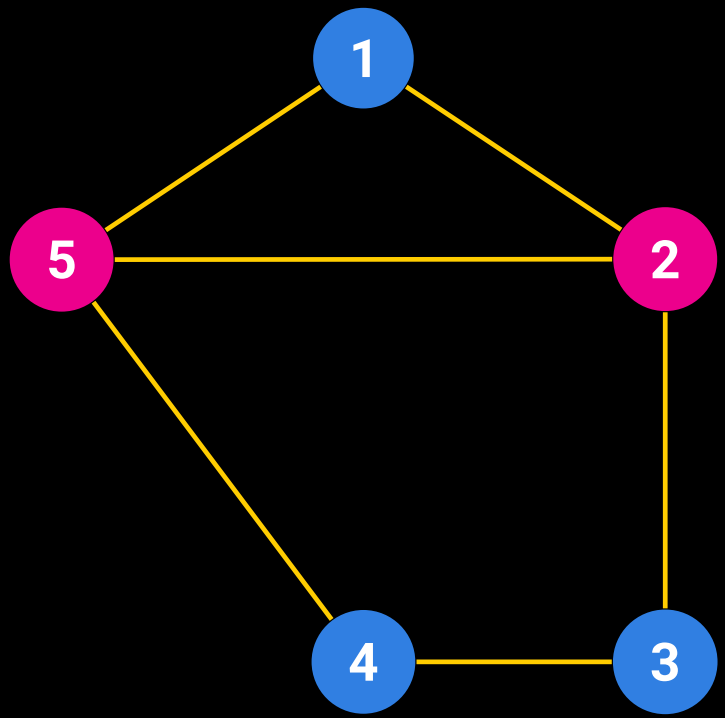
Next Question

What if we design the circuit specifically for an optimization problem?

WSU
Wichita State University

Quantum Approximate Optimization Algorithm

- A variational algorithm for combinatorial optimization
- Example: MaxCut for graph partitioning
- Uses a structured, problem-specific circuit



Key Idea

Build the circuit from the structure of the problem

WSU
Wichita State University

The QAOA Circuit

- Alternates between two operators:
  - Problem Hamiltonian H_C
  - Mixer Hamiltonian H_B

$$U(\gamma, \beta) = e^{-i\beta H_B} e^{-i\gamma H_C}$$

- Parameters: (γ, β)

WSU
Wichita State University

How QAOA Works

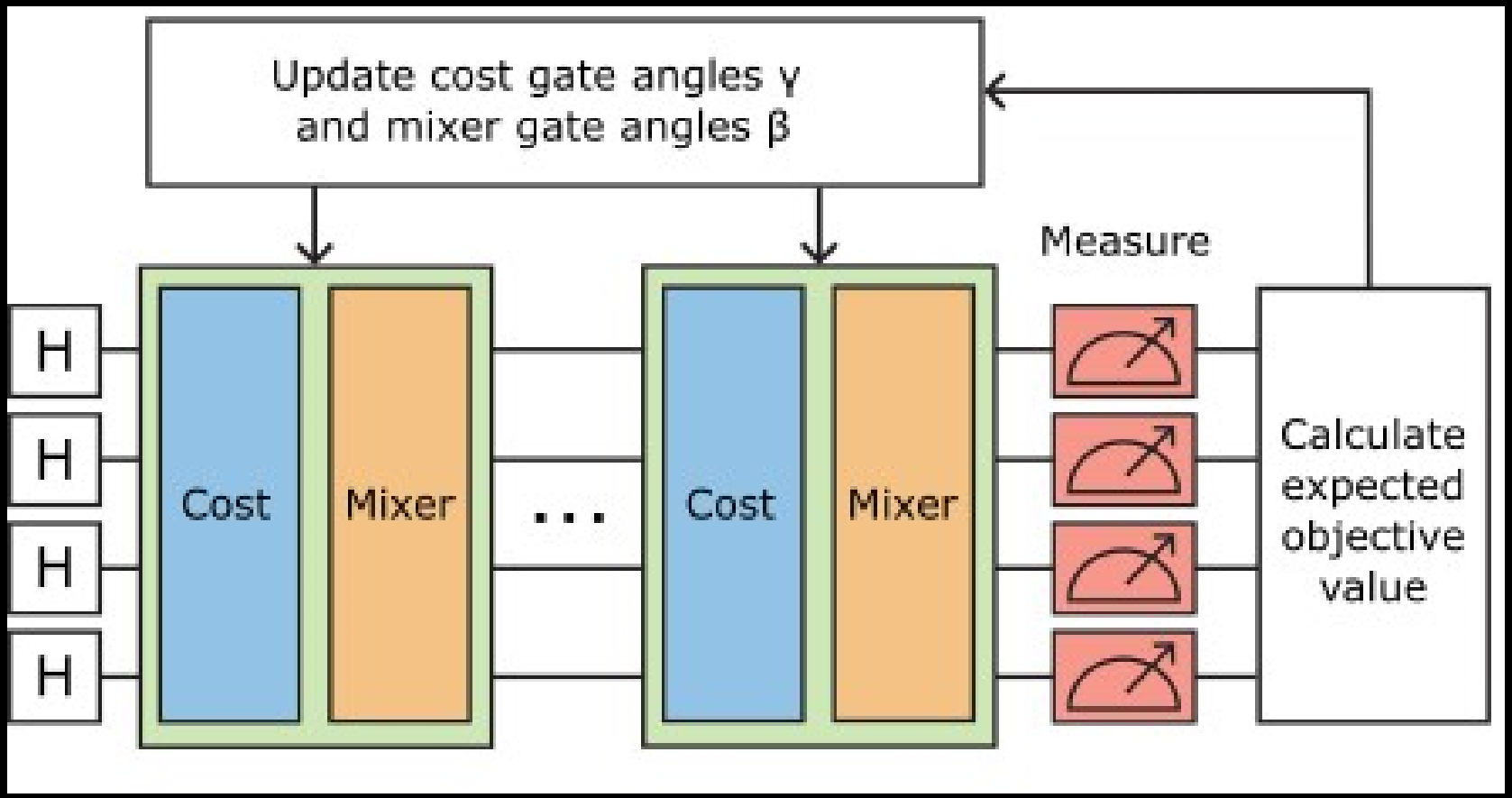
- Start from an initial state

$$H^{\otimes n} \left|0\right\rangle^{\otimes n}$$

- Apply alternating operators:
  - $e^{-i\gamma H_C}$ encodes the problem
  - $e^{-i\beta H_B}$ explores possible solutions
- Measure to estimate solution quality
- Optimize parameters (γ, β)

WSU
Wichita State University

QAOA Circuit



MathWorks:
https://www.mathworks.com/help/matlab/math/max-cut-with-qaoa.html

WSU
Wichita State University

Quick Check-in

What determines the structure of the QAOA circuit?

Answer

The optimization problem itself (via H_C and H_B)

WSU
Wichita State University

QAOA Example: Max-Cut

WSU
Wichita State University

VQE vs QAOA

VQE	QAOA
General-purpose circuit	Problem-specific circuit
Physics-based problems	Combinatorial optimization problems
Flexible ansatz	Structured, layered ansatz
Minimize energy $\langle\psi\|H\|\psi\rangle$	Maximize objective $\langle H_C\rangle$

WSU
Wichita State University

The Training Problem

WSU
Wichita State University

Computing Gradients

- Training requires gradients of the cost function:

$$\frac{\partial C(\theta)}{\partial \theta_i}$$

- We cannot perform backpropagation through a quantum circuit
- Instead, we use the parameter-shift rule

Parameter-Shift Rule

$$\frac{\partial C}{\partial \theta_i} = \frac{1}{2}[C(\theta_i + s) - C(\theta_i - s)]$$

for a fixed shift s

WSU
Wichita State University

Parameter-Shift Rule

- Most quantum circuits use rotation gates

$$R_X(\theta) = e^{-i\frac{\theta}{2}X} \qquad R_Y(\theta) = e^{-i\frac{\theta}{2}Y} \qquad R_Z(\theta) = e^{-i\frac{\theta}{2}Z}$$

- Choosing $s = \frac{\pi}{2}$ gives a special result

Exact Derivative for Rotation Gates

$$\frac{\partial C}{\partial \theta} = \frac{1}{2}\left[C\left(\theta + \frac{\pi}{2}\right) - C\left(\theta - \frac{\pi}{2}\right)\right]$$

WSU
Wichita State University

Cost of Gradient Computation

- Each gradient requires multiple circuit evaluations
- For p parameters:

$$\text{Number of evaluations} \sim O(p)$$

- Each evaluation requires repeated measurements (sampling)

WSU
Wichita State University

Barren Plateaus

- In some circuits, gradients become very small
- As system size increases:

$$\nabla_\theta C(\theta) \approx 0$$

- Optimization becomes extremely difficult

Key Idea

In high-dimensional quantum systems, the landscape can become nearly flat

Why This Matters

Even though the model is expressive, it may not be trainable

WSU
Wichita State University

How Do You Get Started?

WSU
Wichita State University

Getting Started

- Learn the language
  - Linear algebra
  - Bra-ket (Dirac) notation
  - Quantum circuits
- Python - most common programming language for quantum computing
- Machine learning resources
  - MATLAB Tutorials
  - IBM Qiskit
  - PennyLane

WSU
Wichita State University

Any Questions?

WSU
Wichita State University

Thank you!

Wichita State University