Quantum Hopfield Networks

Rajarsi Pal

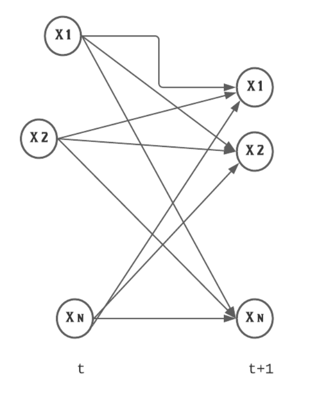
Indian Institute of Technology Madras

Term Project
ID5841 Quantum Lab

But what is a Hopfield network ? [1]

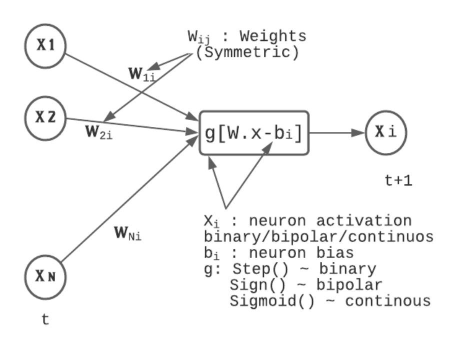
A Hopfield Network is a type of recurrent neural network. They were introduced by J.Hopfield in 1982 as a model for associative memory.

Network architecture :



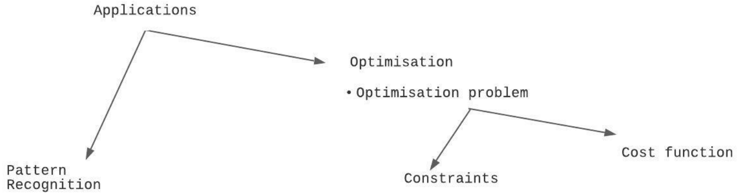
How does it work ?

- define **Energy** as, $E(X) = -\frac{1}{2N}\sum_{i,j} X_i W_{i,j} X_j + \sum_i b_i X_i$
- Neuron activation ;
  $X_i(t+1) = g_\beta(-\frac{\partial E}{\partial X_i}) = g_\beta(\frac{1}{2N}\sum_j W_{i,j}Xj - b_i))$, g_β is the activation function



- *Claim:* Energy decreases monotonically with iterations !

Applications



- Regenerate pattern from corrupted pattern
- Encode a given pattern into the weight matrix, such that stored patterns are at the Energy minima
- Since Energy decreseas monotonically to the minima, arbitrary initial configuration must converge to the stored pattern

- Reformulate the Cost function and Constraints into the Energy function
- By iteratively updating configurations Energy is minimised as Energy => Cost function, Cost function is minimised!

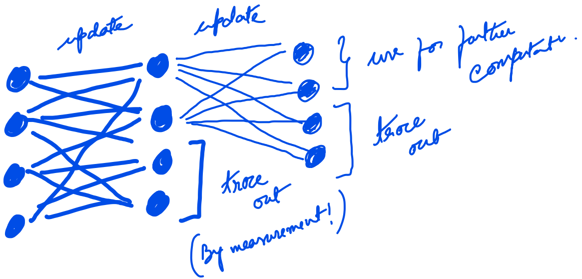
But how to make neurons Quantum ? [3], [2]

Non-linearity is an issue !

- Non-linear activation functions are necessary for neural networks.
- Quantum evolution is Unitary $\implies$ inherently **linear.**!

Possible ways !

- **Measurement based approach**
  Measurements are not unitary : Trace out part of the system on each iteration.
- **Basis Encoding**
  This involves transformations of the form, $|s\rangle|0\rangle \rightarrow |s\rangle|\phi(s)\rangle$, where $\phi(s)$: non-linear function. Using *Quine McClusky* method
- **Rotation Encoding**
  What we will be using.


update
update
use for further computation.
trace out
trace out
(By measurement!)

Miller's Model
[QHAM]

Encoding $\{\pm 1\}$ states into $|kets\rangle$

$|\psi(x_i)\rangle = cos(\frac{\pi}{4}x_i + \frac{\pi}{4})\,|0\rangle + sin(\frac{\pi}{4}x_i + \frac{\pi}{4})\,|1\rangle$
Thus, $\vec{x} = \{x_1, x_2, x_3..\} \implies |\psi(x_1, x_2..)\rangle = |\psi(x_1)\rangle\,|\psi(x_2)...\rangle$

Encoding W_{ij} into $R_y(\phi_{i,j})$

Idea: encode elements of W_{ij} as arguments to CR_y

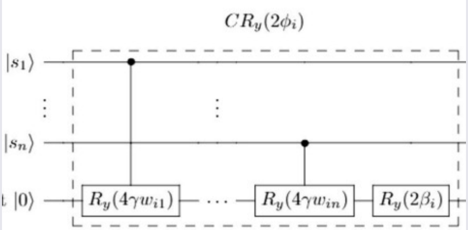
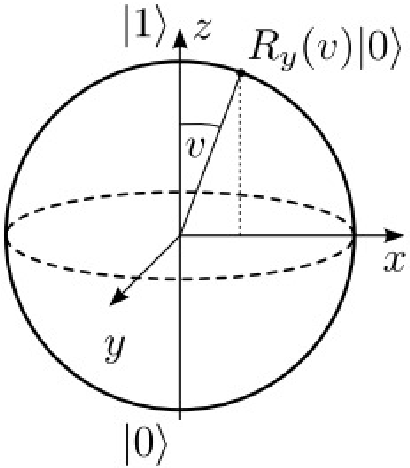


Figure: Rotation Encoding


$|1\rangle$
z
$R_y(v)|0\rangle$
v
x
y
$|0\rangle$

Miller's Model

Updating the system

Idea:

- Use CR_y to update the $|ancilla\rangle$ initialised at $|0\rangle$.
- **SWAP** $|ancilla\rangle$ with the qubit to be updated.

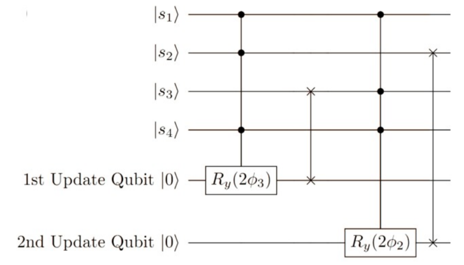
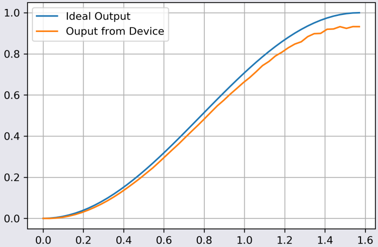


Figure: Updating qubit $|s3\rangle$ and $|s2\rangle$

Miller's Model

Effective Non-linearity



But how to get back the classical states ?

Idea: Use Majority voting

- Basically, run the circuit a number of times and estimate $P(|1\rangle)$ and $P(|0\rangle)$ my measuring it.
- Interpret the state as $(+1)$ if $P(|1\rangle) > 0.5$ else as (-1)

Illustrations

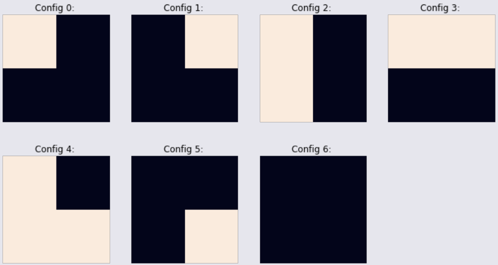
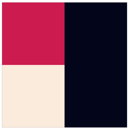
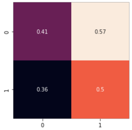
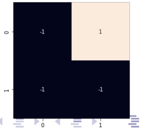


Figure: Stored-Patterns







‘

Best way to update states ?

Repeated vs. Final Measurement

- Measure all qubits after every updating step.
- Results are more accurate
- But involves mid-circuit measurements

- Measure all qubits only at the end.
- Results are often ambiguous.
- Avoids mid-circuit measurements. Evolution is more quantum in nature.

Best way to update states ?

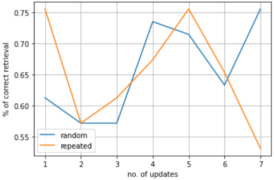


Figure: n= 4

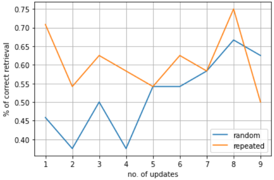


Figure: n= 5

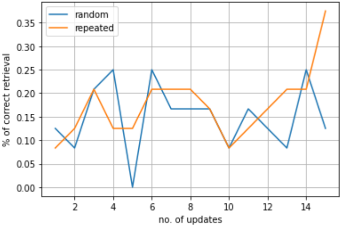
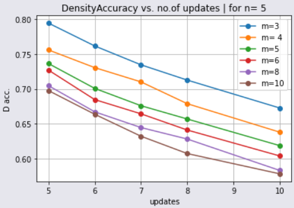


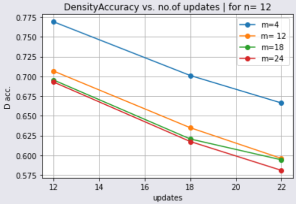
Figure: n= 8

How well can it retrieve the patterns ?

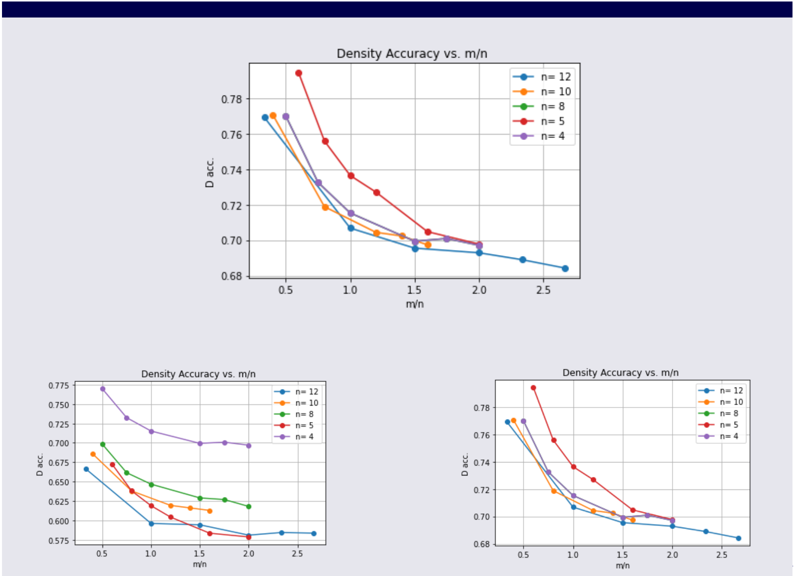
Scaling of Eff. Capacity

- Check variation of **Density Accuracy** of the network against **(m/n)**, where **m** is the no. of patterns, **n** is the no. of qubits, tuning **no.of updates** to n and 2n respectively.
- Check variation of **Density Accuracy** of the network against **no.of updates**, for different cases of **m** and **n**.





How well can it retrieve the patterns ?



Where does it go wrong ?

Majority Voting

- Unjustly collapses the state to the computational basis $\implies$ looses essential quantum information.
- Looses all information on the phase of the state.
- Density accuracy falls drastically with no. of (m/n) ratio.

Repeated Updates

- Sequence of updates affect trajectory as update operators are non-commuting.
- Postponing the measurements to the end lead to states where $P(|0\rangle) = P(|1\rangle) = 0.5$ for all qubits.

Things to do better..

Rajarsi Pal

- This model is not truly Quantum as it does not incorporate effects like superposition and entanglement into its design.
- There is no way to probe into the dynamics of the system i.e to find metrics like current **Energy**, **Hamming distance** to stored configurations.
- Capacity of the network indicates that it barely manages to satisfy the classical limit. No real quantum advantage !

References

[1] J J Hopfield. "Neural networks and physical systems with emergent collective computational abilities.". In: *Proceedings of the National Academy of Sciences* 79.8 (1982), pp. 2554–2558. DOI: 10.1073/pnas.79.8.2554. eprint: https://www.pnas.org/doi/pdf/10.1073/pnas.79.8.2554. URL: https://www.pnas.org/doi/abs/10.1073/pnas.79.8.2554.

[2] Florian Marquardt. "Machine Learning and Quantum Devices". In: *SciPost Phys. Lect. Notes* (2021), p. 29. DOI: 10.21468/SciPostPhysLectNotes.29. URL: https://scipost.org/10.21468/SciPostPhysLectNotes.29.

[3] Maria Schuld and Francesco Petruccione. *Supervised Learning with Quantum Computers*. 1st. Springer Publishing Company, Incorporated, 2018. ISBN: 3319964232.