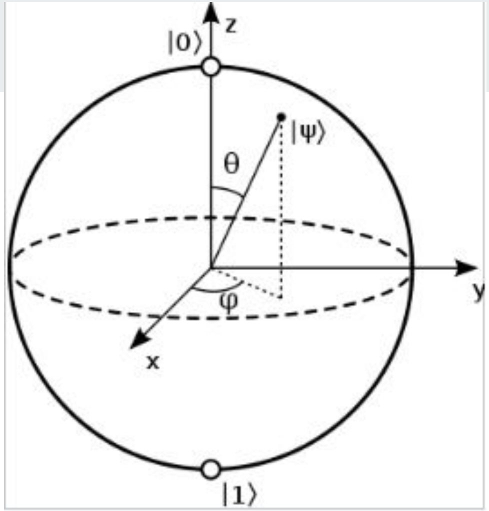
Reinforcement Learning with Neural Networks for Quantum Feedback

Zehra Abdrahim

Quantum information intro

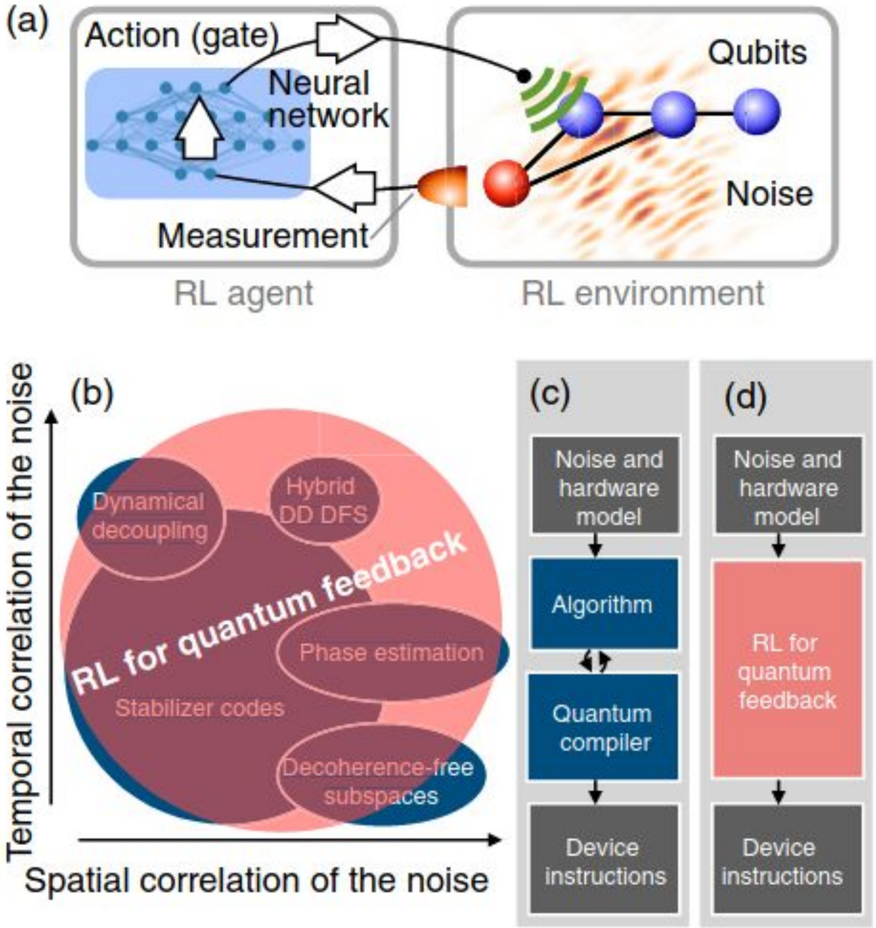


- Bit vs. qubit
- Qubit states - Coherence $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ $|\alpha|^2 + |\beta|^2 = 1$
- Bloch sphere representation
- Quantum logic gates
    - Quantum entanglement
    - Controlled NOT gate

Neural-Network-based RL in a few-qubit quantum system

- Finds quantum-error-correction strategies
  - Finds complex sequences like projective measurements and entangling gates
- Protects quantum memory from noise and decoherence
- Unified, fully autonomous, human-guidance free
- Finds hard-ware adapted solutions

- Key ideas in the method
    - 2-stage learning - state-aware network and recurrent network
    - A measure of the student quantum info (initial state to be fully restored)



- Probabilistic policy

$$\pi_\theta(a_t|s_t)$$

- Policy gradient rule

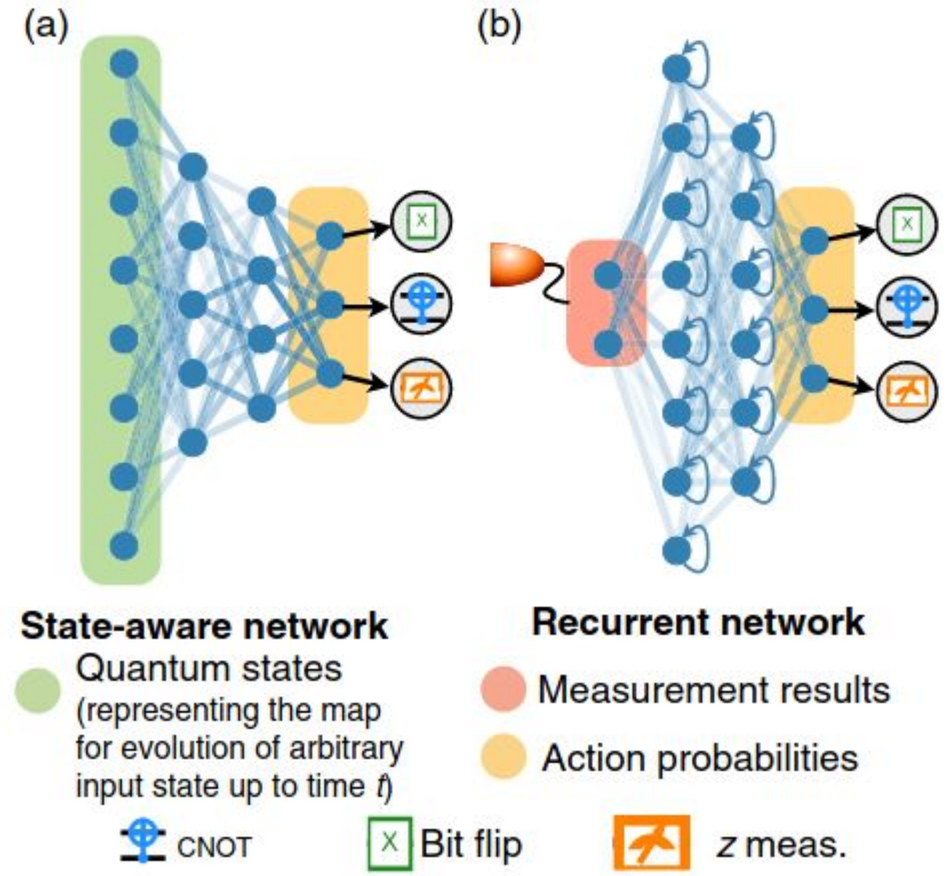
$$\delta\theta_j = \eta \frac{\partial \mathbb{E}[R]}{\partial \theta_j} = \eta \mathbb{E}\left[R \sum_t \frac{\partial}{\partial \theta_j} \ln \pi_\theta(a_t|s_t)\right]$$

RL Approach to Quantum Memory

- Steps for the stabilizer-code
  - the encoding, in which the logical state initially stored in one qubit is distributed over several physical qubits,
  - the detection of errors via measurement of suitable multiqubit operators,
  - The subsequent correction,
  - the decoding procedure that transfers the encoded state back into one physical qubit,
- $\alpha|0\rangle + \beta|1\rangle$

Phase estimation

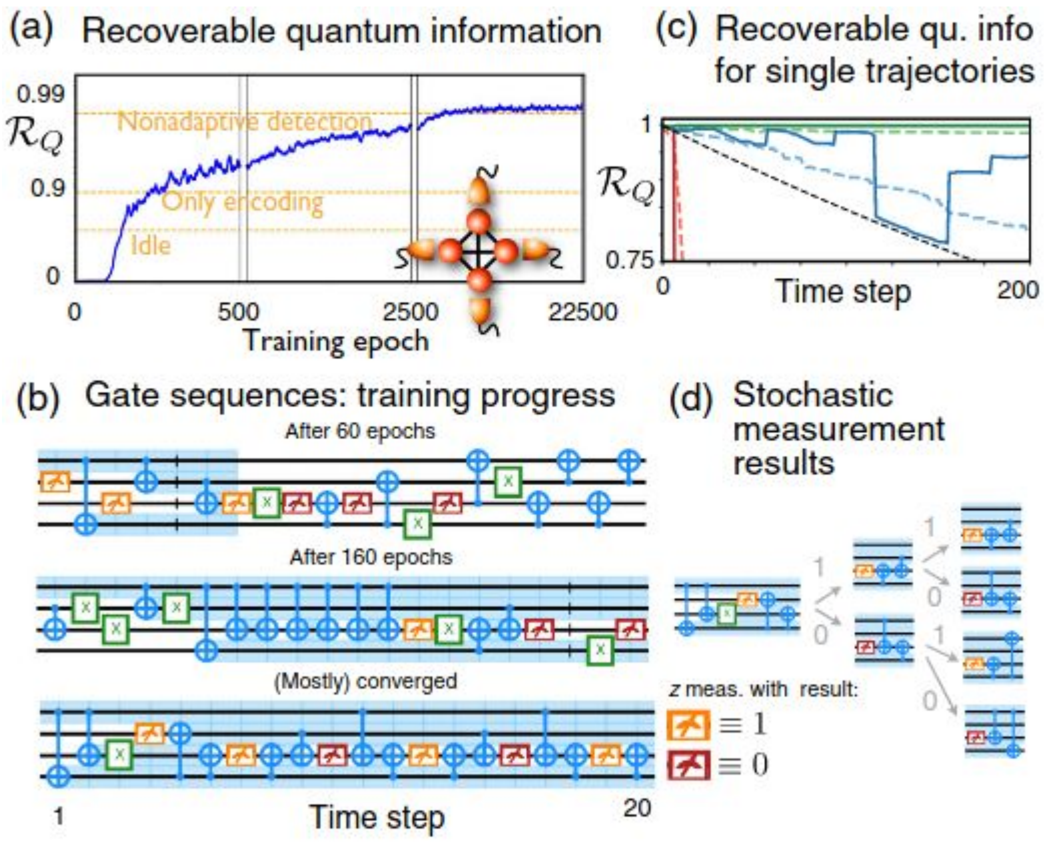
- 4 initial quantum states $\hat{\rho}(0)$ for one logical qubit
- Completely positive map Φ of the multiqubit system that maps $\hat{\rho}(0)$ o $\hat{\rho}(t)$
- Partially observed Markov proccess

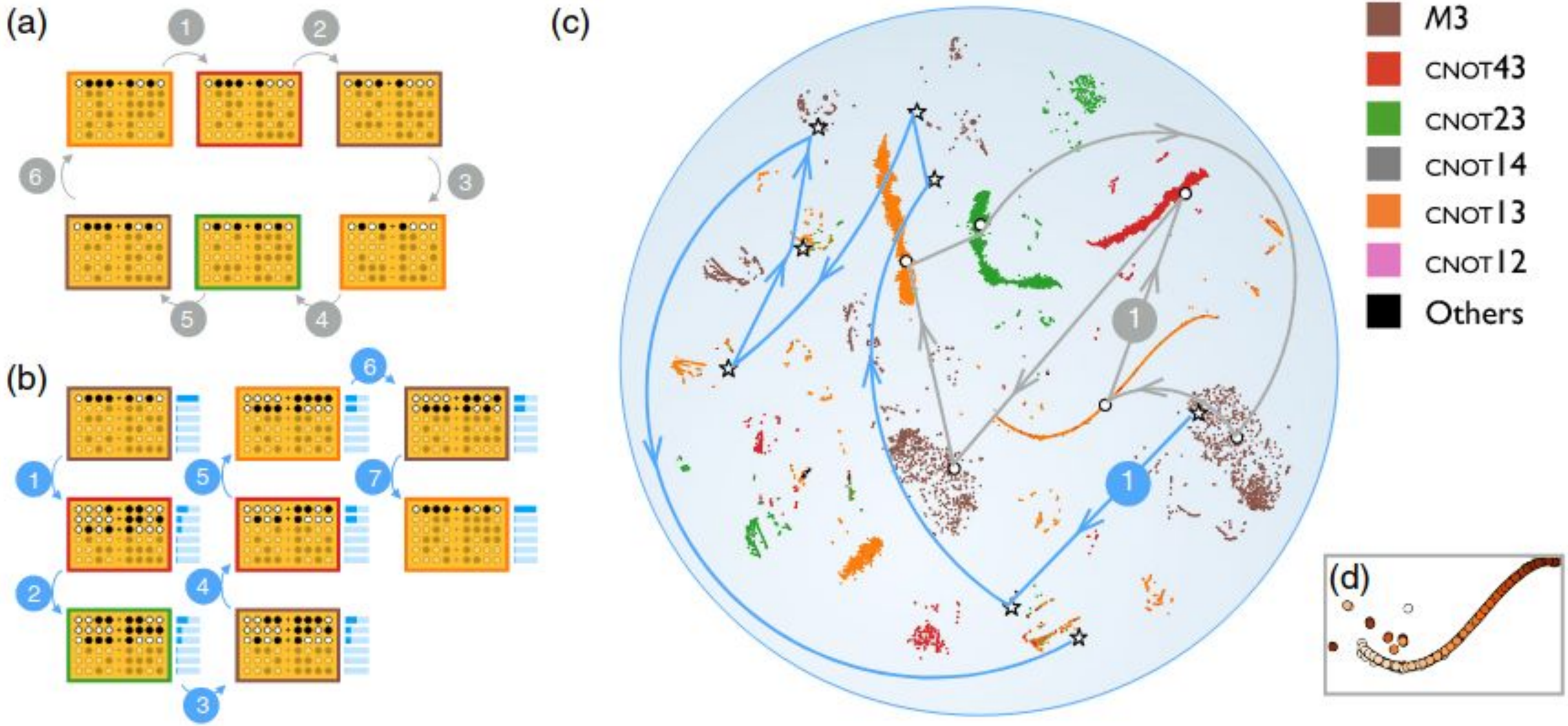


Reward

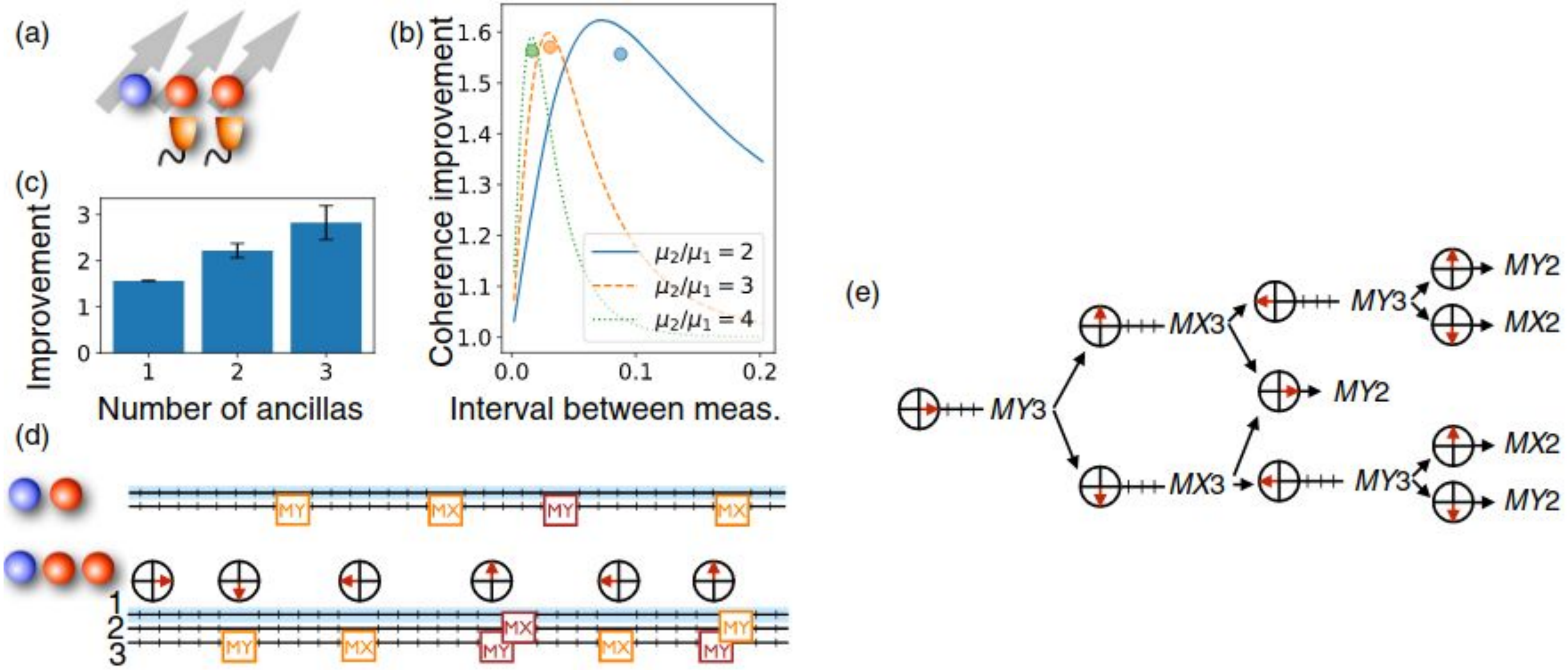
- The probability to distinguish two quantum states by after mapping:
  - Trace distance $\frac{1}{2}\|\hat{\rho}_{\vec{n}}(t) - \hat{\rho}_{-\vec{n}}(t)\|_1$
- Recoverable quantum info
  - $\mathcal{R}_Q(t) = \frac{1}{2}\min_{\vec{n}} \|\hat{\rho}_{\vec{n}}(t) - \hat{\rho}_{-\vec{n}}(t)\|_1$

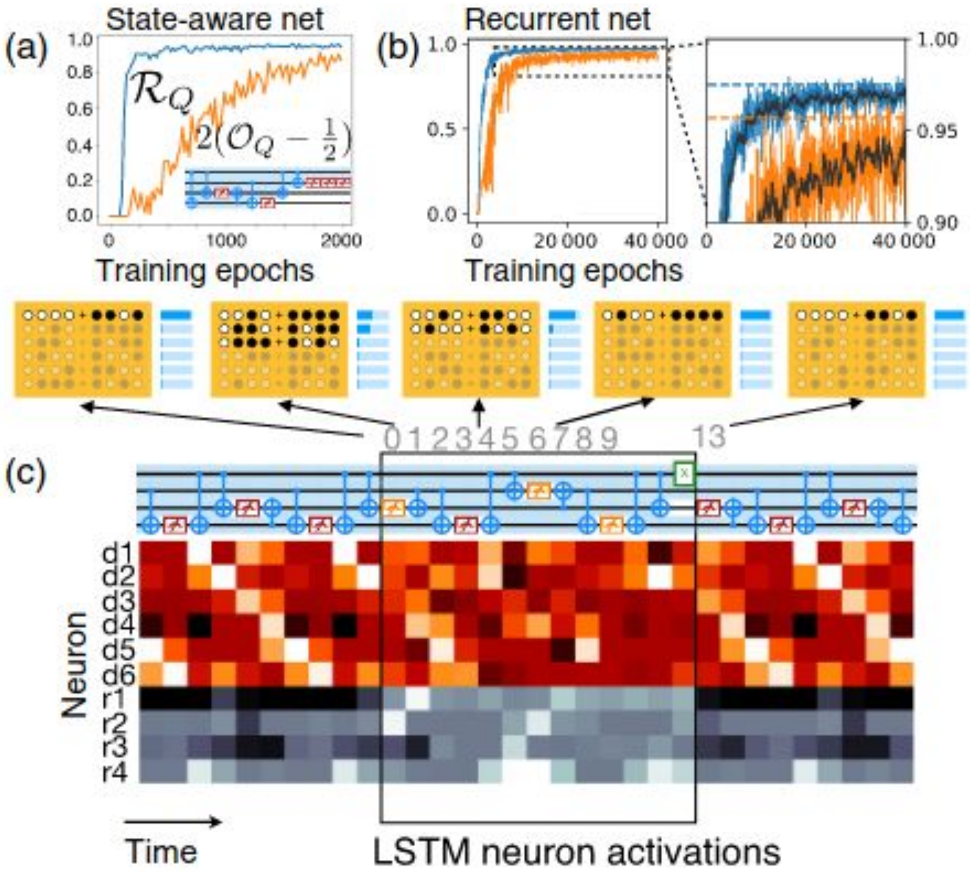
Results


(a) Recoverable quantum information
0.99
$\mathcal{R}_Q$
0.9
0
Nonadaptive detection
Only encoding
Idle
0
500
2500
22500
Training epoch
(c) Recoverable qu. info for single trajectories
1
$\mathcal{R}_Q$
0.75
0
Time step
200
(b) Gate sequences: training progress
After 60 epochs
After 160 epochs
(Mostly) converged
1
Time step
20
(d) Stochastic measurement results
z meas. with result:
≡ 1
≡ 0



(a)
(b)
(c)
(d)
M3
CNOT43
CNOT23
CNOT14
CNOT13
CNOT12
Others


(a)
(b)
(c)
(d)
(e)
Improvement
Number of ancillas
Coherence improvement
Interval between meas.
$\mu_2/\mu_1 = 2$
$\mu_2/\mu_1 = 3$
$\mu_2/\mu_1 = 4$
MY
MX
MY3
MX3
MY2
MX2


(a)
State-aware net
$\mathcal{R}_Q$
$2(\mathcal{O}_Q - \frac{1}{2})$
Training epochs
(b)
Recurrent net
Training epochs
0 1 2 3 4 5 6 7 8 9
13
(c)
Neuron
d1
d2
d3
d4
d5
d6
r1
r2
r3
r4
Time
LSTM neuron activations

Future Applications and Conclusion

- Cavities as long-lived quantum memory
- Qubit connectivity
- Cooptimize the hardware layout (taking into account physical constraints)
- non-Markovian noise,
- weak measurements,
- qubit-cavity systems,
- error-corrected transport of quantum information through networks

Thank You for Your Attention!