QML Journal Club @ CQT

Machine Learning Bell Nonlocality

Kishor Bharti
(CQT)

Motivation

Machine Learning

$$P(T) \propto E$$

"A Computer program is said to learn from experience E with respect to some class of tasks T and performance P, if its performance at tasks in T, as measured by P improves with experience E"

Tom Mitchell, 1997

Motivation

Foundational:
Understanding the principles behind intelligence

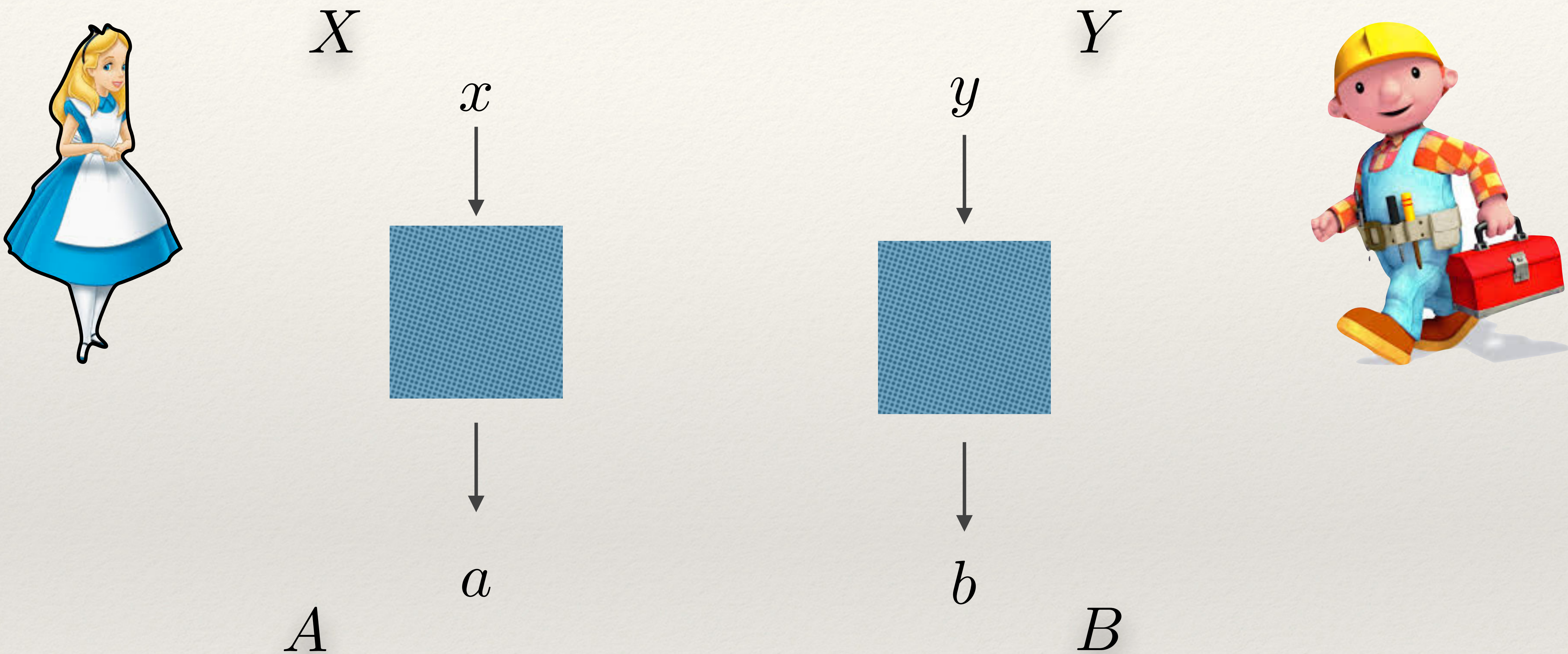
Human vs a machine

Motivation

- Practical:
  - "AI is the new electricity "
    1. Healthcare
    2. Transportation
    3. Communication
  - Applications in Physics

https://physicsml.github.io/pages/papers.html

Bell Nonlocality



Local realism

$$P_c\left(a,b|x,y\right)=\sum_{\lambda}p\left(\lambda\right)p\left(a|x,\lambda\right)p\left(b|y,\lambda\right)$$

Quantum theory

$$P_Q\left(a,b|x,y\right)=\mathrm{Tr}\left[\left(M_a^x\otimes M_b^y\right)\rho\right]$$

Bell Nonlocality

$\boxed{\text{Local realism}}$

$$P_c\left(a,b|x,y\right)=\sum_{\lambda}p\left(\lambda\right)p\left(a|x,\lambda\right)p\left(b|y,\lambda\right)$$

A statistics which can't be expressed as above, is said to exhibit Bell nonlocality.

$\mathcal{L}$

- The set of Bell local statistics
- A convex polytope
- Boundaries are Bell inequalities

$(2,2,2)$

$\vec{P}\in\mathbb{R}^{16}$

Bell test: A game whose winning strategy is described by Bell nonlocal statistics.

$\boxed{\text{Quantum theory is Bell nonlocal.}}$

Applications of Bell Nonlocality

Randomness certification

Quantum cryptography

Distributed computing

Self testing

Over the Table

Transforming Bell's inequalities into state classifiers with machine learning
[(Ma and Yung) npj Quantum Information 2018]

- Uses machine learning and Bell inequalities to develop state classifiers
- Tools from machine learning:
    - Feed-forward neural networks

Over the Table

Machine learning non-local correlations
[(Canabarro, Brito, Chaves) arXiv May 2018]

- Uses machine learning for detection and quantification of non locality
- Tools from machine learning:
  - Feed forward neural network
  - TPOT

Machine Learning Based State Classifier

Machine Learning Based State Classifier

Entangled state: a quantum state is entangled if it can't be expressed as a convex combination of product states.

$$\rho_{\text{sep}} = \sum_i p_i \rho_i^1 \otimes \rho_i^2 \otimes \cdots \otimes \rho_i^n$$

$$0 \leq p_i \leq 1 \qquad\qquad \sum_i p_i = 1$$

Goal: To classify states as entangled or separable

Challenges:

- Obtaining full description of a quantum state becomes resource consuming as the number of qubits increases.
- Detecting entanglement is NP hard even if all the information about the states is given. [Gurvits, STOC '03 10 (2003)]

Bell inequalities predictor

Is it possible to use Bell inequalities to classify states?

Bell inequalities can't be used as reliable tool for entanglement detection.

$$|\psi_-\rangle = \frac{|00\rangle - |11\rangle}{\sqrt{2}}$$

$$\rho = p|\psi_-\rangle\langle\psi_-| + (1-p)\,\frac{I}{4}$$

$p > \frac{1}{3}$ Is entangled

$p > \frac{1}{\sqrt{2}} \approx 0.707$ Violates CHSH inequality

Angle dependence

Another reason: measurement angles depends on the quantum state.

$$\Pi_{\mathrm{CHSH}} \equiv A_0B_0 - A_0B_1 + A_1B_0 + A_1B_1$$

$$A_0 = \sigma_z \qquad B_0 = \frac{(\sigma_x - \sigma_z)}{\sqrt{2}}$$

$$A_1 = \sigma_x \qquad B_1 = \frac{(\sigma_x + \sigma_z)}{\sqrt{2}}$$

$$|\psi_{\theta,\phi}\rangle = \cos\left(\frac{\theta}{2}\right)|00\rangle + \exp\left(i\phi\right)\sin\left(\frac{\theta}{2}\right)|11\rangle$$

$$\langle\psi_{\theta,\phi}|\Pi_{\mathrm{CHSH}}|\psi_{\theta,\phi}\rangle = \sqrt{2}\left(\sin\theta\cos\phi - 1\right)$$

$\theta = \frac{\pi}{2}, \phi = \pi$ ✓ $\qquad$ $\theta = \frac{\pi}{2}, \phi = \frac{\pi}{2}$ ✗

Refined Goal

Refined Goal: To transform Bell's inequalities into a reliable entanglement-separable states classifier

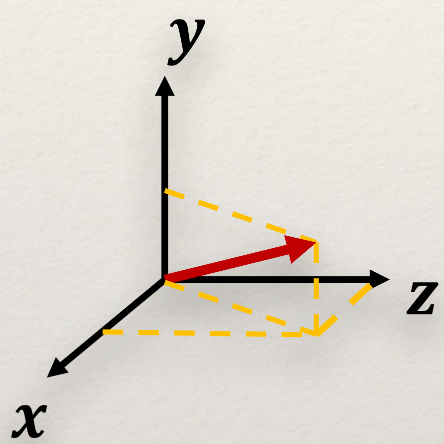
- Given the same measurement settings, is it possible to optimise the coefficients of CHSH inequality for a better performance, compared with the values (1,-1,1,1,2)?
- Is it possible to use hidden layers to improve the optimisation?
- Is it possible to construct a universal state classifier for detecting quantum entanglement?

Main challenge: labelling data

- PPT criteria: works up to d = 6
- Semidefinite programming

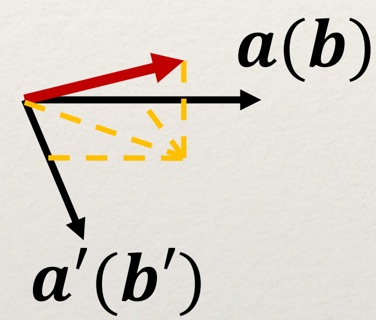
Predictors

Tomographic Predictor



- Uses all information of a given quantum state
- Used to benchmark the performance of Bell-like predictor

Bell-like Predictor



- Uses a subset of non-orthogonal measurements

Two qubits case: 15 versus 4

Types of predictors

CHSH

- The weights are fixed
- The measurement angles are fixed
- Linear in structure

$$\mathrm{CHSH}_{\mathrm{ml}} \equiv \mathbf{w_1}\langle \mathbf{a_0 b_0}\rangle + \mathbf{w_2}\langle \mathbf{a_0 b_0'}\rangle + \mathbf{w_3}\langle \mathbf{a_0' b_0}\rangle + \mathbf{w_4}\langle \mathbf{a_0' b_0'}\rangle + \mathbf{w_0}$$

- The weights are random
- The measurement angles are fixed
- Linear in structure

$\mathrm{Bell}_{\mathrm{ml}}\,(n, n_f, n_e)$ (number of qubits, number of features, number of neurons in the hidden layer)

- The weights are random
- The measurement angles are random
- Non-linear with hidden layer

Labelling quantum states

$$\rho_{\theta,\phi} = p|\psi_{\theta,\phi}\rangle\langle\psi_{\theta,\phi}| + (1-p)\,\frac{I}{4}$$

$0 \leq p \leq 1$

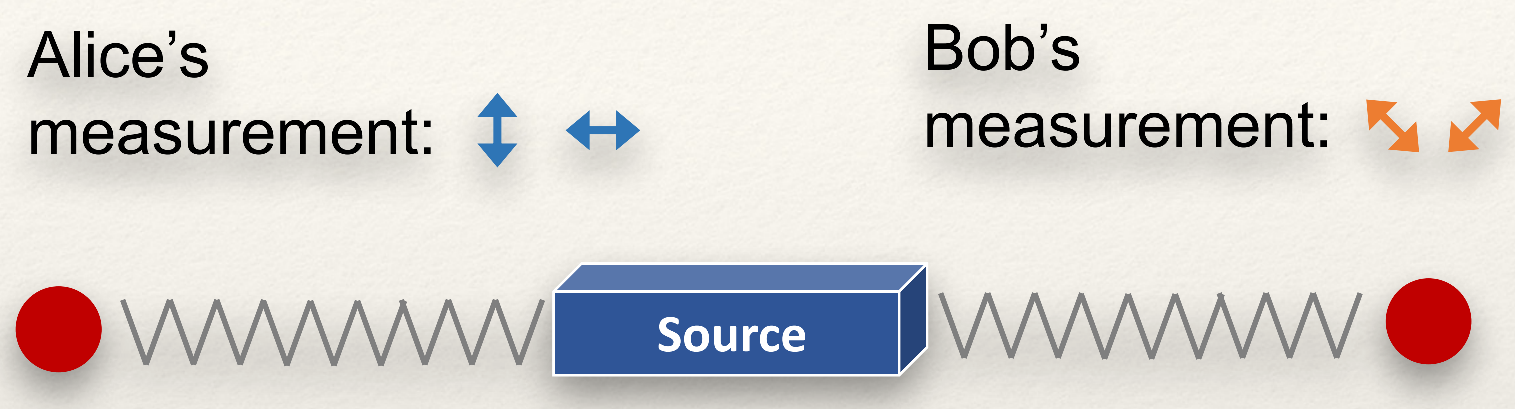
$$\lambda_{\min}\left(\rho_{\theta,\phi}^{T_B}\right) = \frac{1-p}{4} - p\cos\frac{\theta}{2}\sin\frac{\theta}{2}$$

$\lambda_{\min}\left(\rho_{\theta,\phi}^{T_B}\right) < 0 \implies$ **State is entangled**

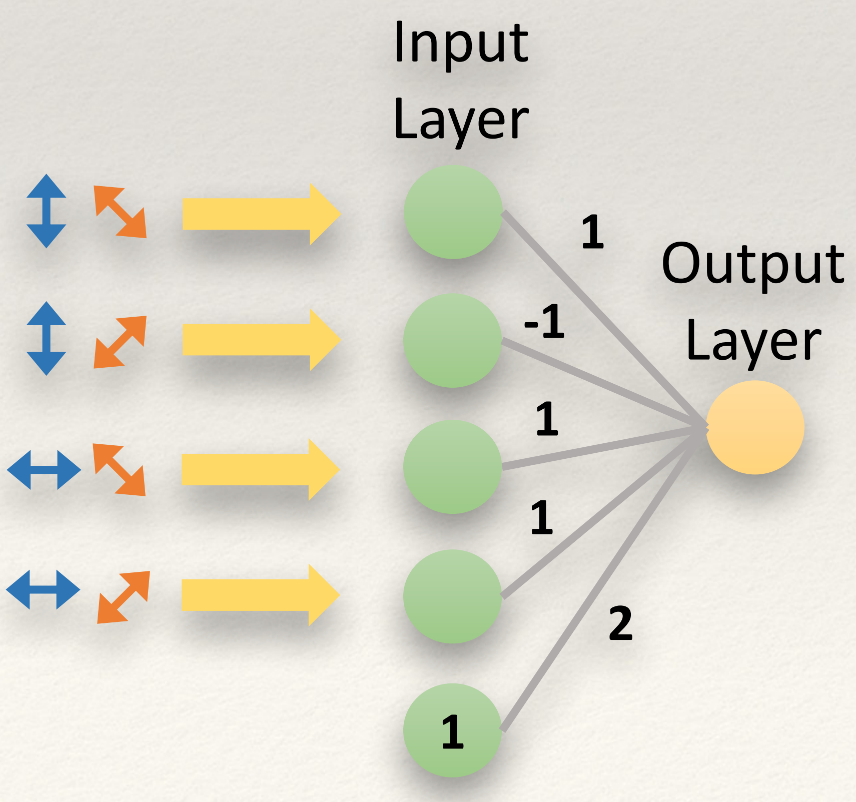
Entanglement does not depend on ϕ

For a general state, one requires complete tomography. $4^n - 1$

Linear predictor



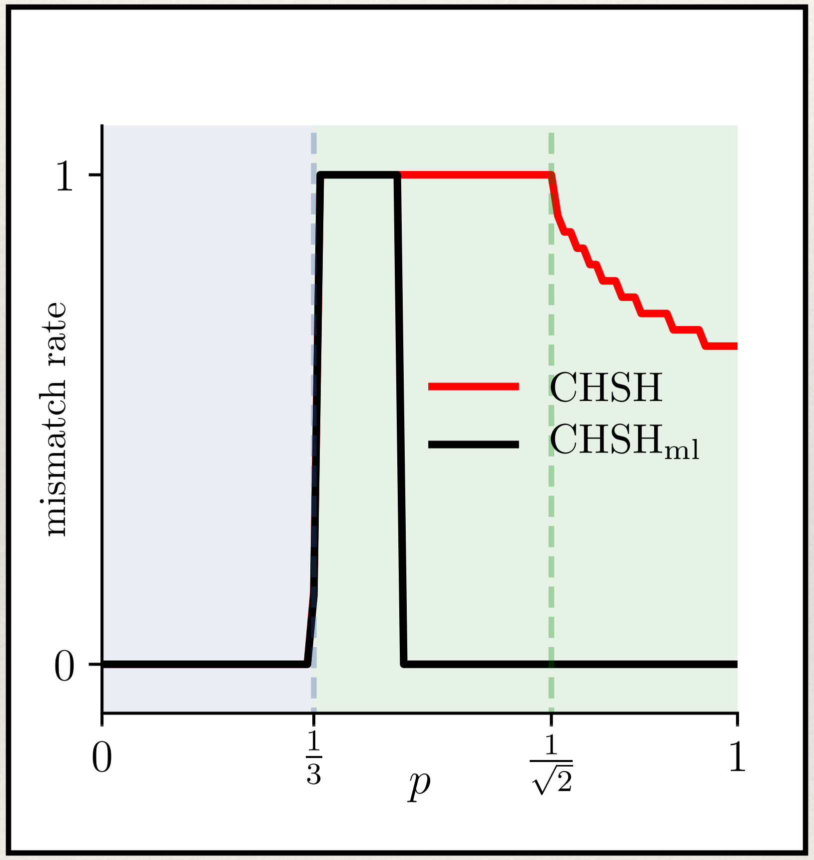
$$R_{\mathrm{mm}}\left(p,\theta\right) \equiv \Pr\left(1_{\mathrm{ML}}|0_{\mathrm{PPT}}\right) + \Pr\left(0_{\mathrm{ML}}|1_{\mathrm{PPT}}\right) \qquad \text{averaged over } \phi$$

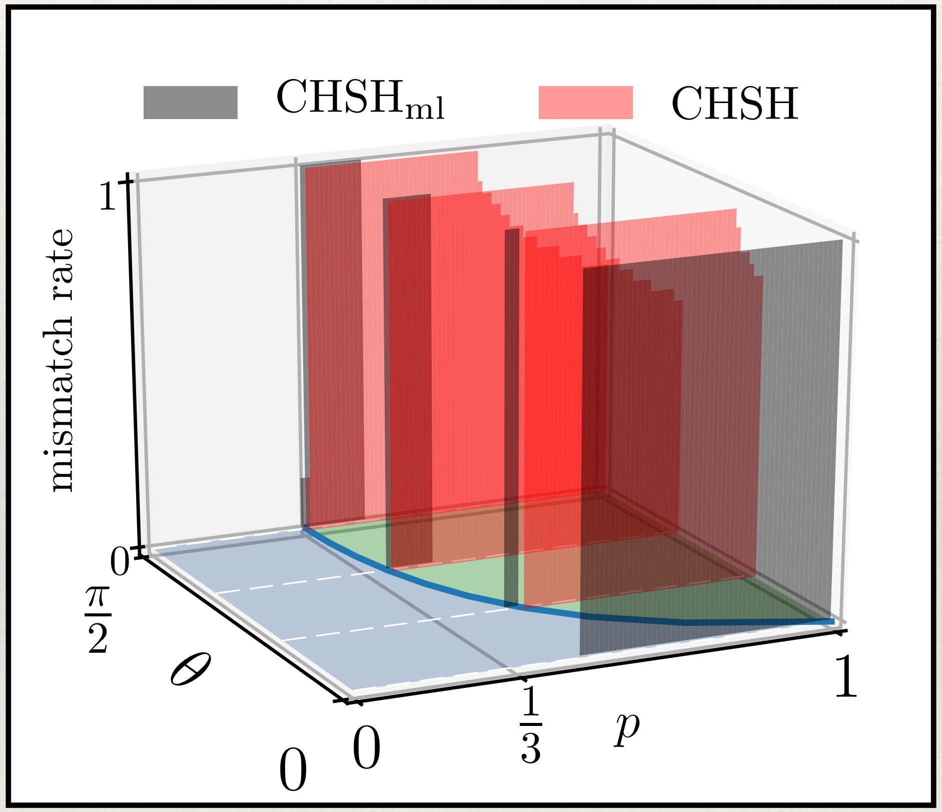


$$\mathrm{CHSH}_{\mathrm{ml}} \equiv \mathbf{w_1}\langle \mathbf{a_0 b_0}\rangle + \mathbf{w_2}\langle \mathbf{a_0 b_0'}\rangle + \mathbf{w_3}\langle \mathbf{a_0' b_0}\rangle + \mathbf{w_4}\langle \mathbf{a_0' b_0'}\rangle + \mathbf{w_0}$$

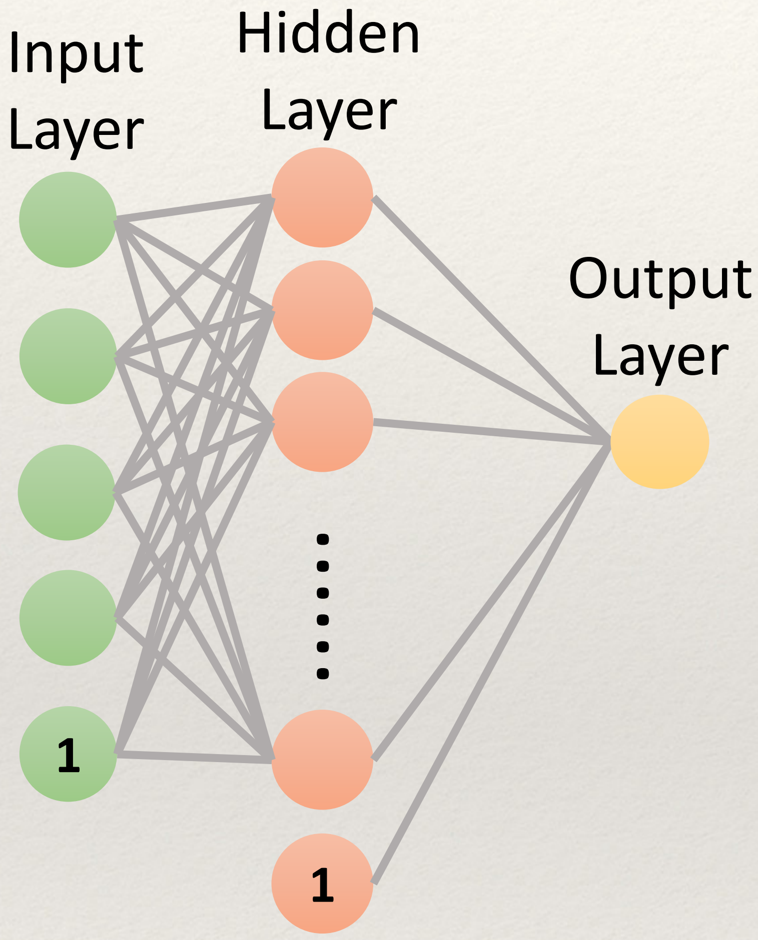
$$w_1 = 0.521, w_2 = -0.603, w_3 = -0.025, w_4 = 0.016, w_0 = 0.373$$

Comparing the performance





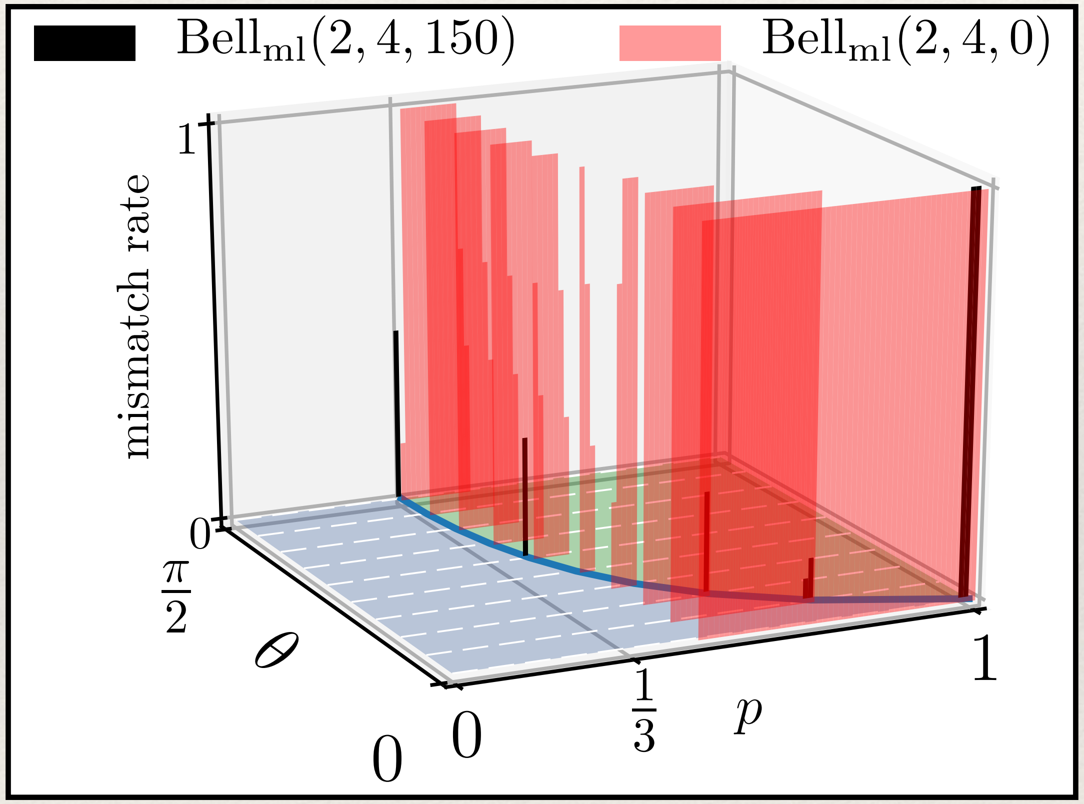
Nonlinear predictor

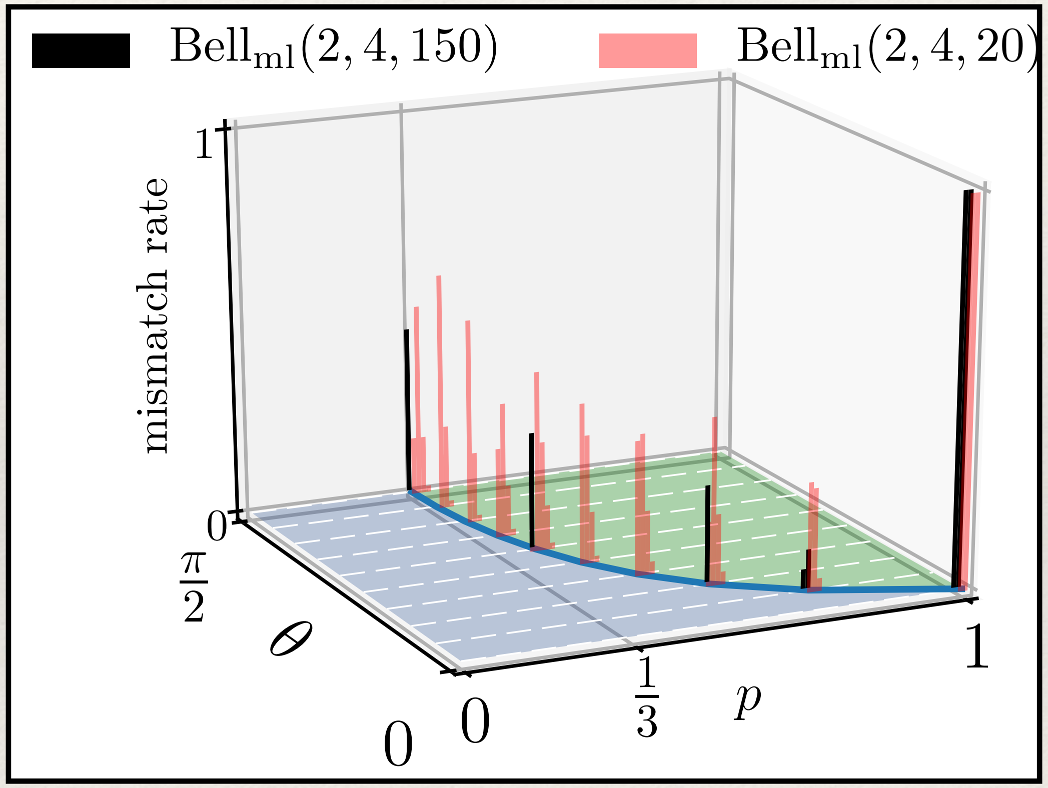


ReLu is used as activation function in the hidden layer.

Sigmoid is used in the final layer.

Comparing the performance

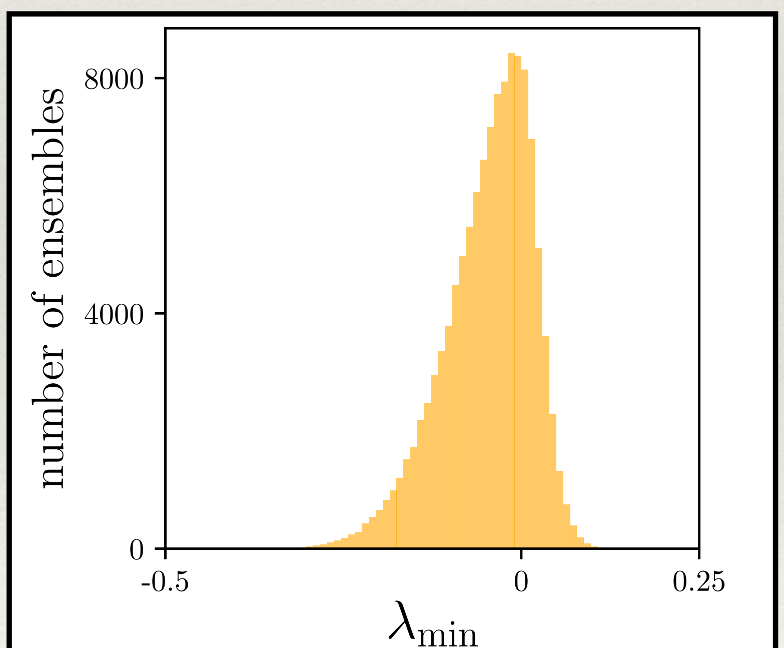




Classifying general two qubit states

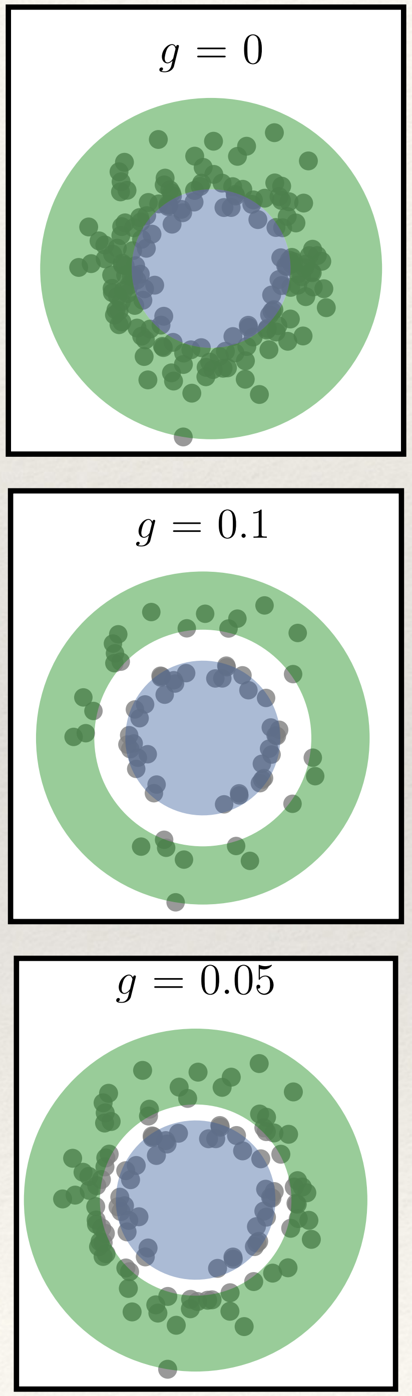
Is it possible to build universal state classifier by using partial information?

No. Look at Lu, D *et. al* PRL **116**, 230501 2016. However, you can use tomographic predictors.



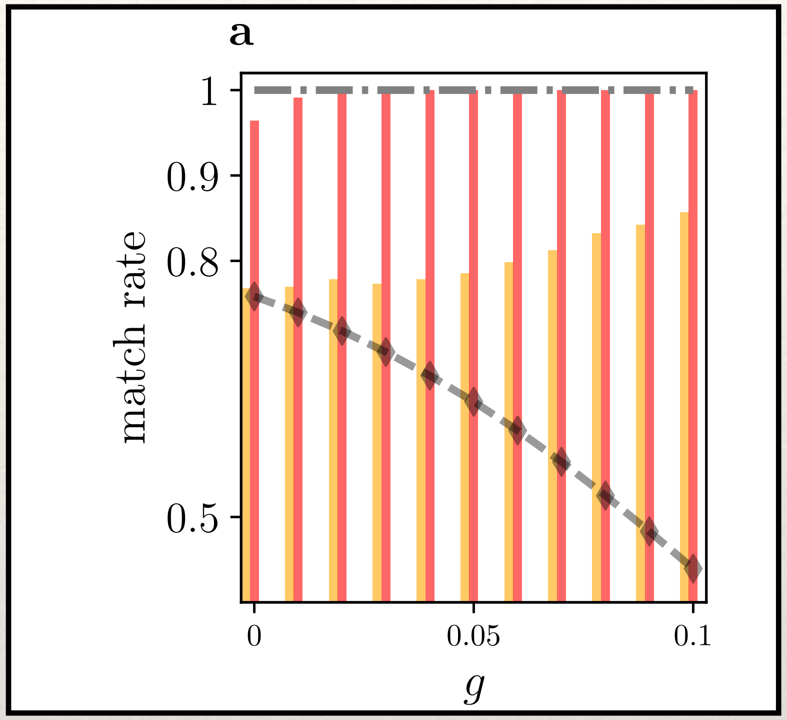
The performance of the predictor depends on training data as well as test data.

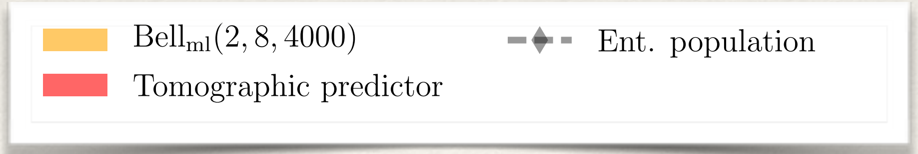
Classifying general two qubit states



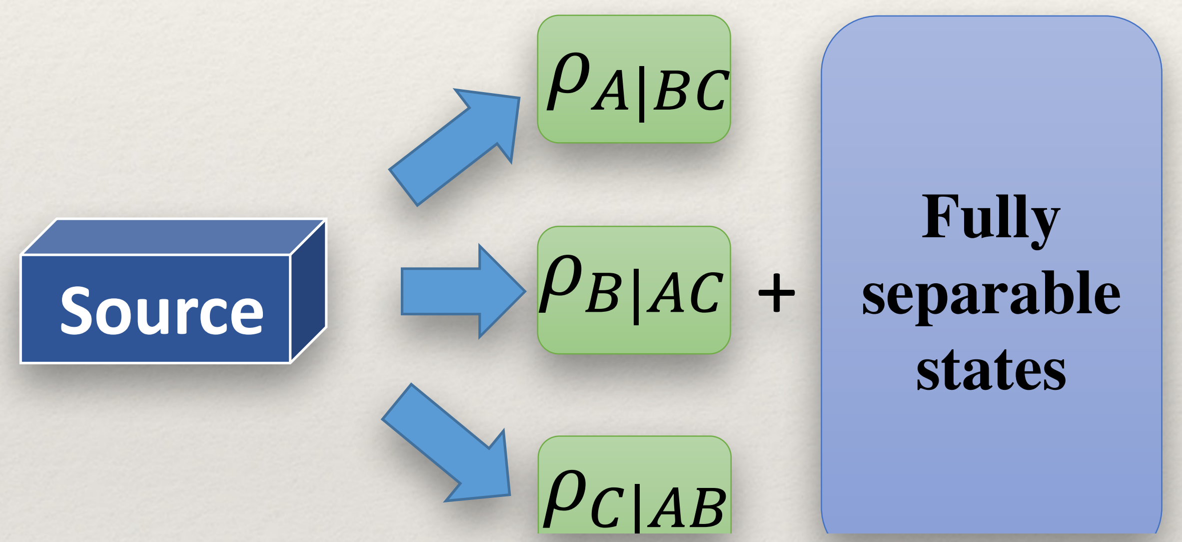
Green: entangled
Blue: separable

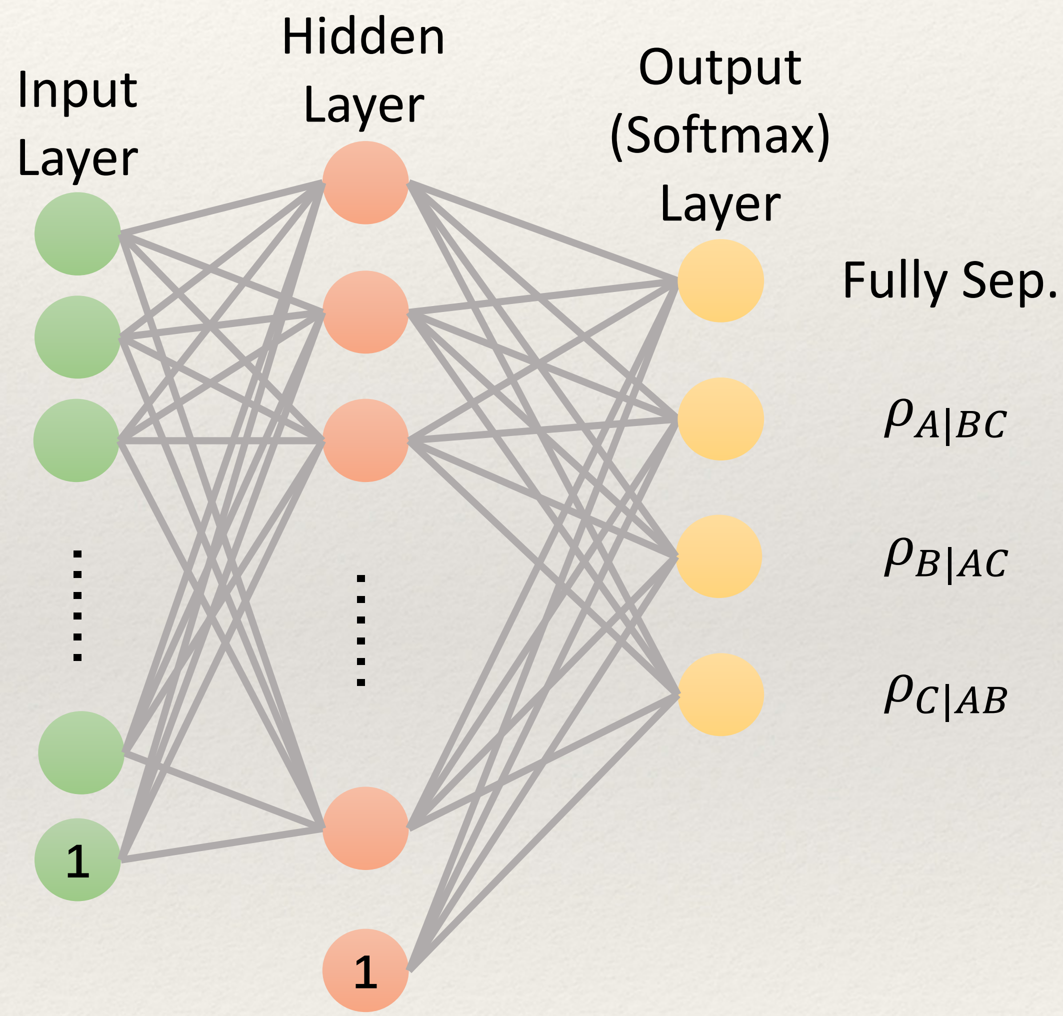
$$|\lambda_{\text{min}}| > g$$





Machine Learning Based State Classifier





N qubits

- Tomographic predictor for N qubit case
  - Challenge: labelling states

- Detecting bound-entangled state in a family of three qubit systems

- Identifying N qubit GHZ type from fully separable states states

$$4^N - 1 \text{ vs } 2N$$

Machine learning non-local correlations

Machine learning non-local correlations

"How far a given correlation is from the local set"

$\mathcal{L}$ Local set

$$q \equiv q\left(a, b|x, y\right)$$

$$p \equiv p\left(a, b|x, y\right)$$

Minimum trace distance

$$NL\left(q\right) = \frac{1}{2|x||y|}\min_{p\in\mathcal{L}} \sum_{a,b,x,y} |q - p|$$

$$|x| = |y| = m$$

Machine learning non-local correlations

- Training was done in supervised manner.

$$\{\vec{f}, t\}$$

features labels

$\langle A_x B_y \rangle$ Components of feature vector

$NL(q)$ Labels

- Uses TPOT

TPOT

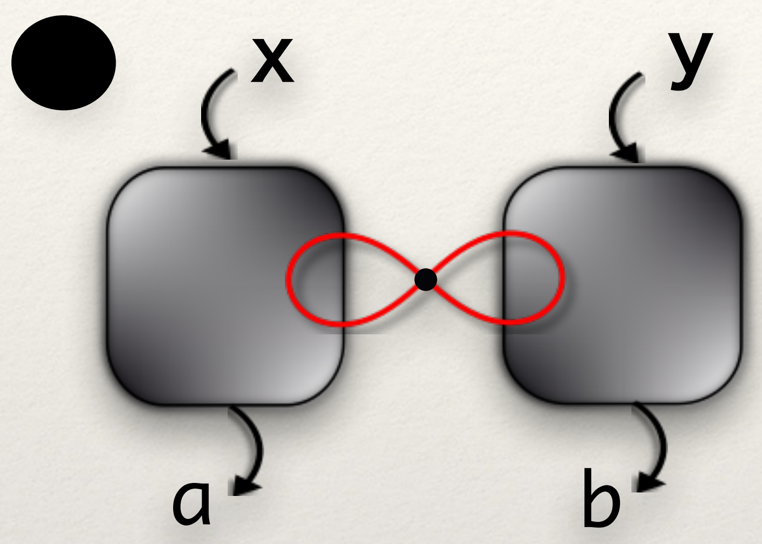
- No a priori machine learning algorithm more suitable for the task at hand
- Requires testing a variety of machine learning approaches with vast hyper-parameter space

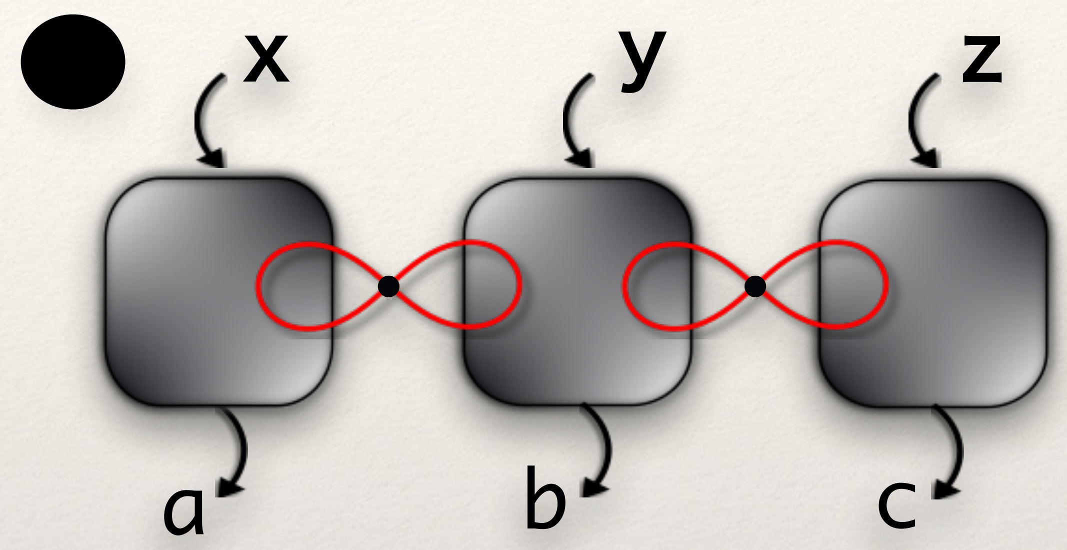
- Python based scikit-learn resource developed by Randal Olson while he was a postdoctoral researcher at University of Pennsylvania

- Free automated machine learning tool
- For a regression or classification task, it gives optimal pipeline suitable for the data

Randal Olson

Machine learning non-local correlations

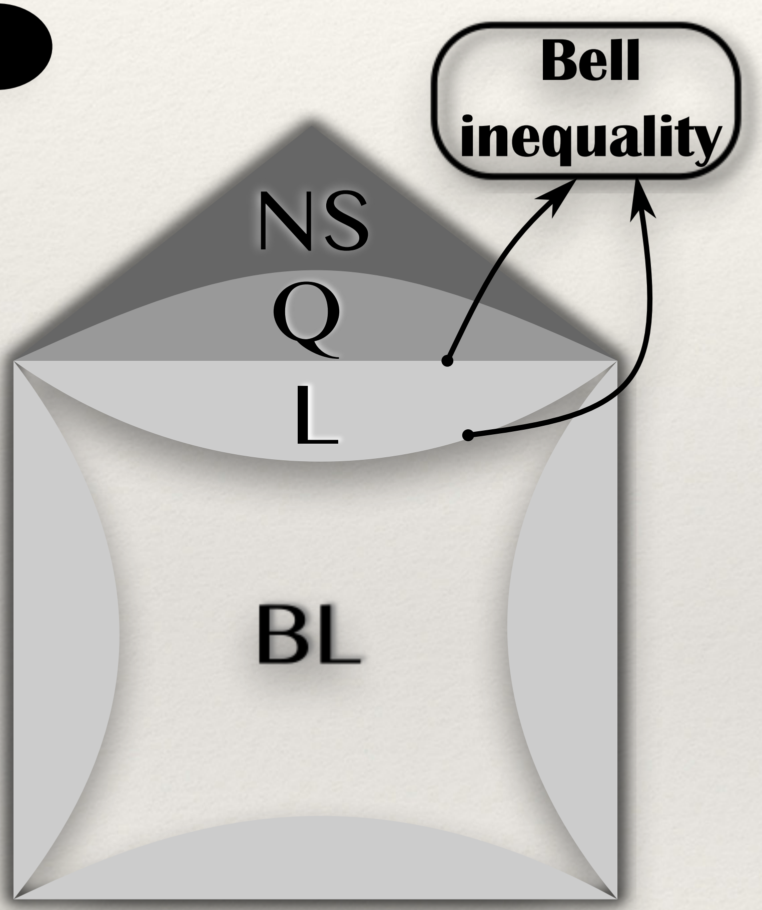


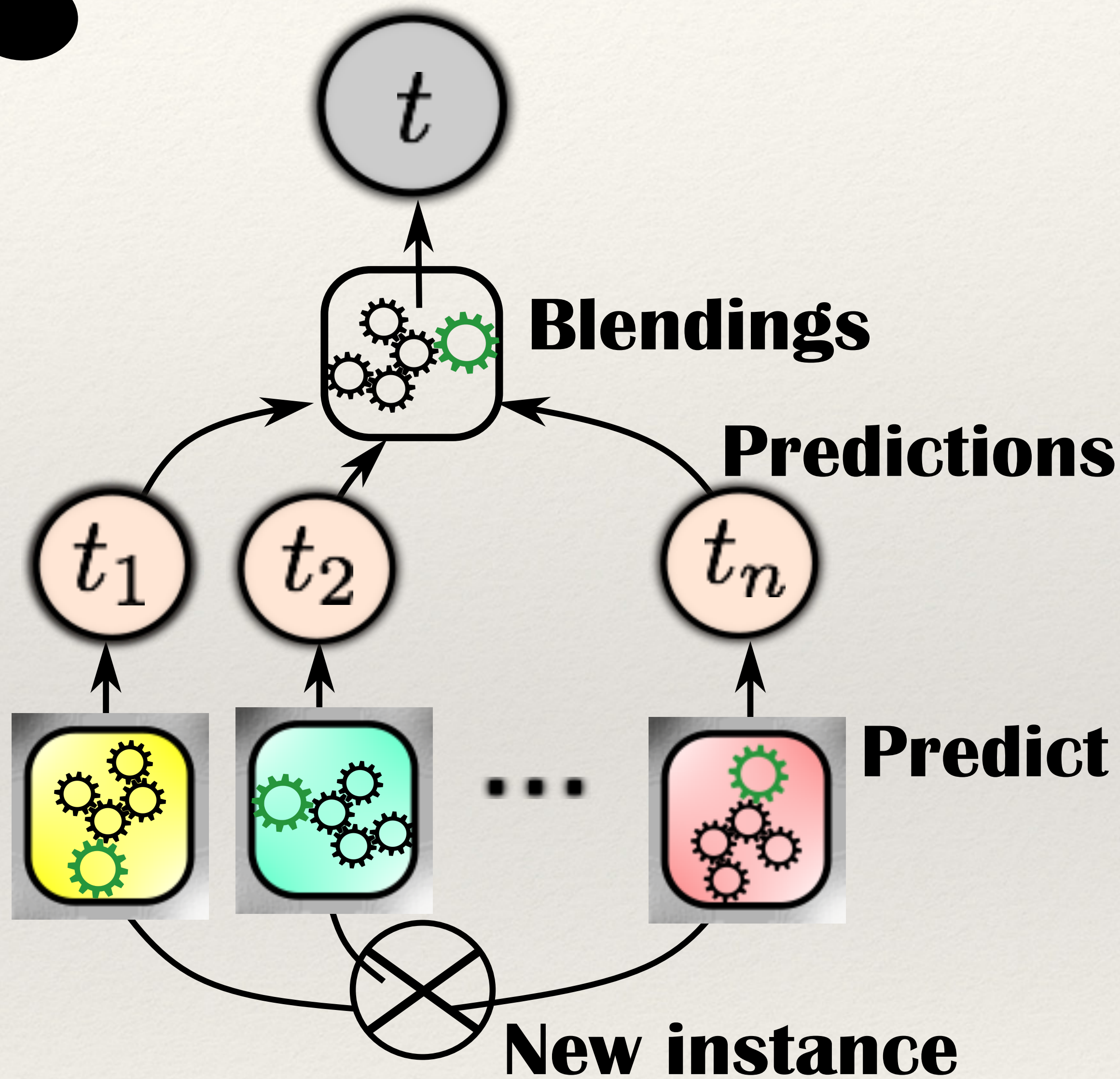


$$p\left(a,b,c|x,y,z\right)=\sum_{\lambda_1,\lambda_2}p\left(a|x,\lambda_1\right)p\left(b|y,\lambda_1,\lambda_2\right)p\left(c|z,\lambda_2\right)p\left(\lambda_1\right)p\left(\lambda_2\right)$$

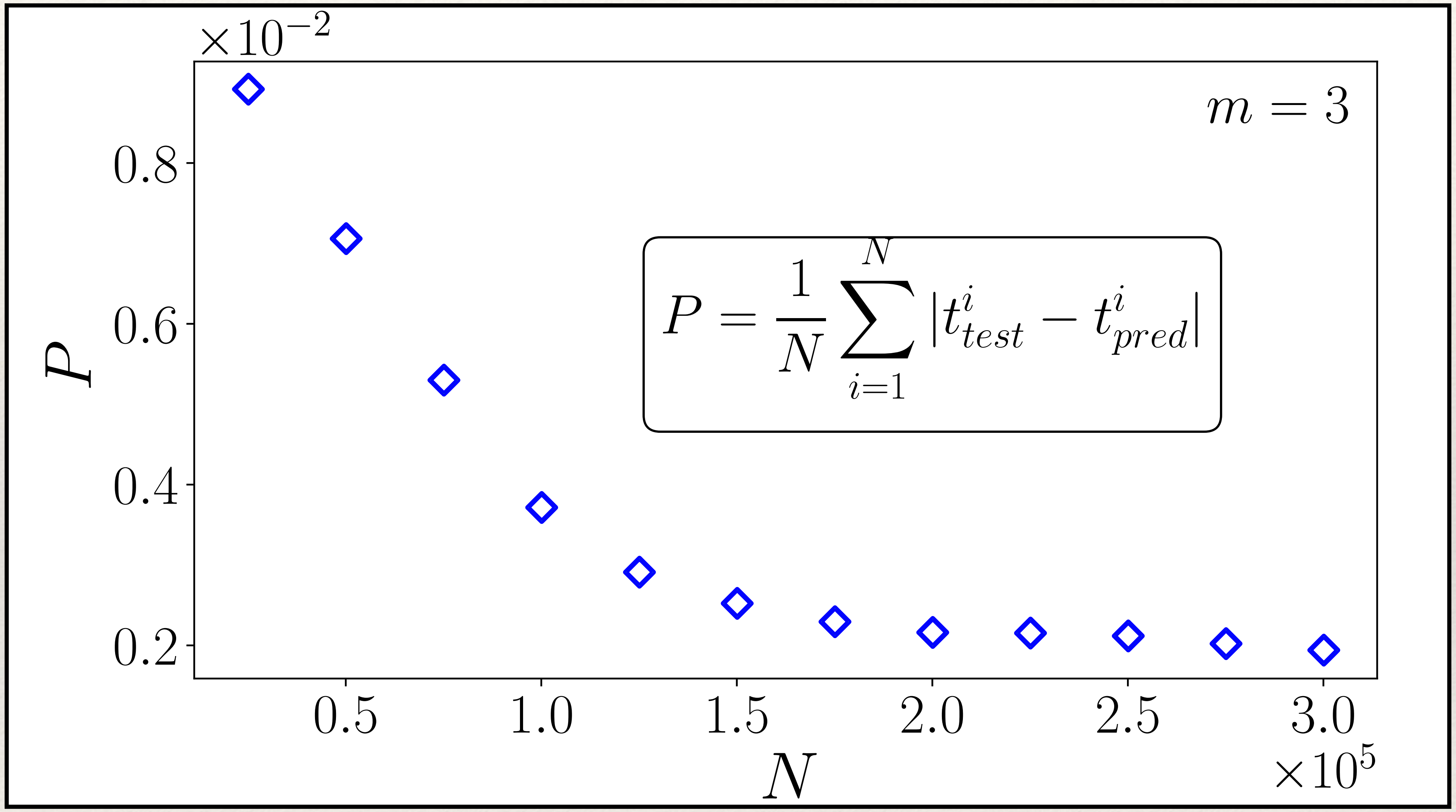
$$p\left(\lambda_1,\lambda_2\right)=p\left(\lambda_1\right)p\left(\lambda_2\right)$$

Machine learning non-local correlations

- 

- 

Machine learning non-local correlations

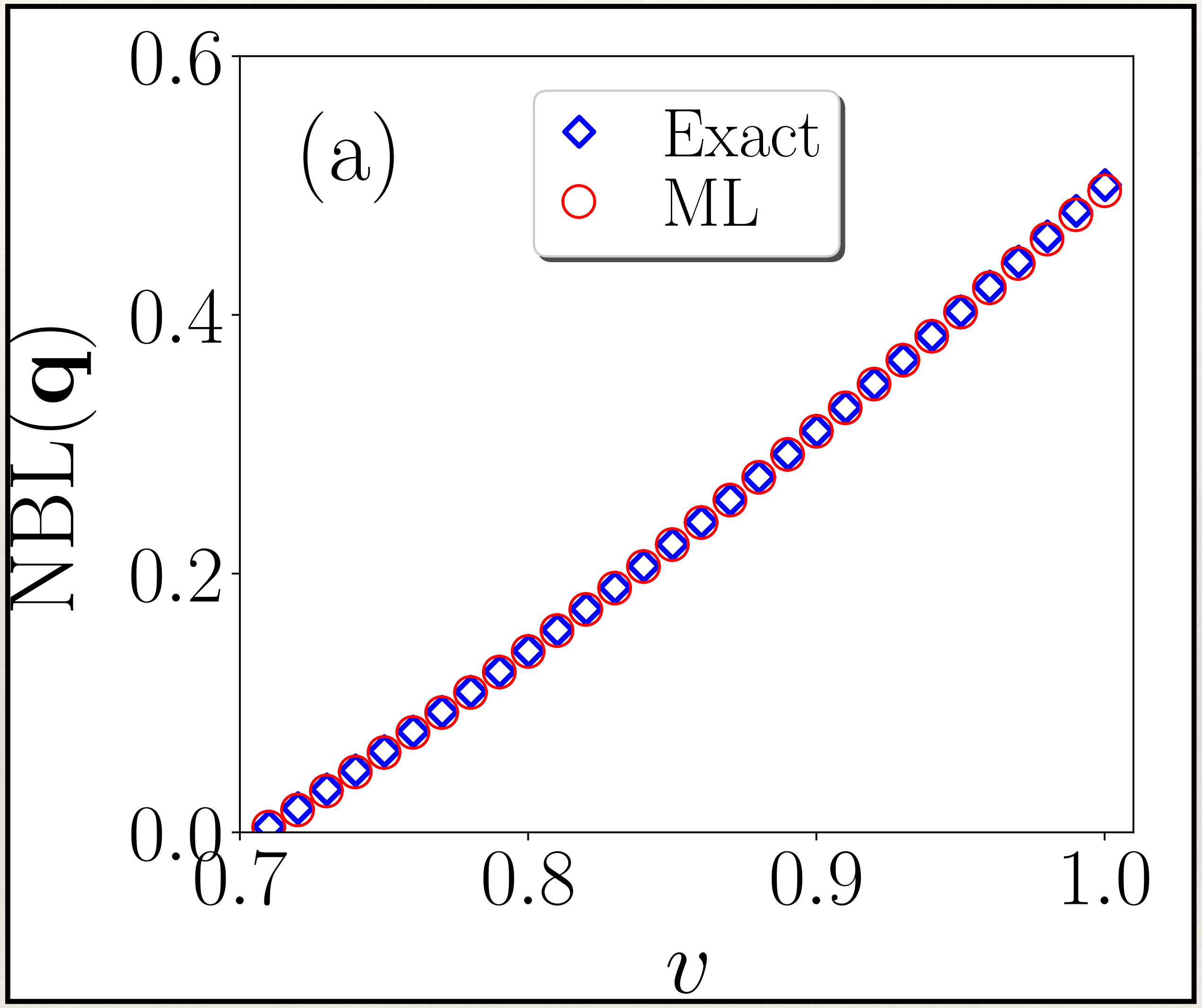


Average trace distance error for m =3 bipartite Bell scenario versus size N of the training set

Machine learning non-local correlations

$$\rho = v|\phi^+\rangle\langle\phi^+| + (1-v)\,\frac{I}{4}$$

$$\mathrm{NBL}\,(q) = v^2 - \frac{1}{2}$$



Summary and Future

- Transforming Bell inequalities into state-classifier with machine learning

$$\mathrm{CHSH_{ml}} \equiv w_1\langle \mathbf{a_0 b_0}\rangle + \mathbf{w_2}\langle \mathbf{a_0 b_0'}\rangle + \mathbf{w_3}\langle \mathbf{a_0' b_0}\rangle + \mathbf{w_4}\langle \mathbf{a_0' b_0'}\rangle + \mathbf{w_0}$$

- Using ANN to learn the set of non-local correlations

Future

- Machine learning contextual correlations

Questions

Steps

- Generate a two qubit density matrix from the required family
- Label it as entangled/separable
- Find out the feature vectors for
  - n = 4, 6, 8, 10, 12 , 14 and 16
- Train the single layer classifier and get weights for these values
- Compare the performance with regular n-cycle CHSH
- Add hidden layers and compare the performance with layer
- Finally compare the same for tomographic predictors
- Look for a different family of density matrices and do the same