A Separability-Entanglement Classifier via Machine Learning

arXiv:1705.01523 [quant-ph] (2017)

Jun Li

July 4, 2017

Workshop on Machine Learning and Many-body Physics

Kavli Institute for Theoretical Sciences at UCAS

About Me

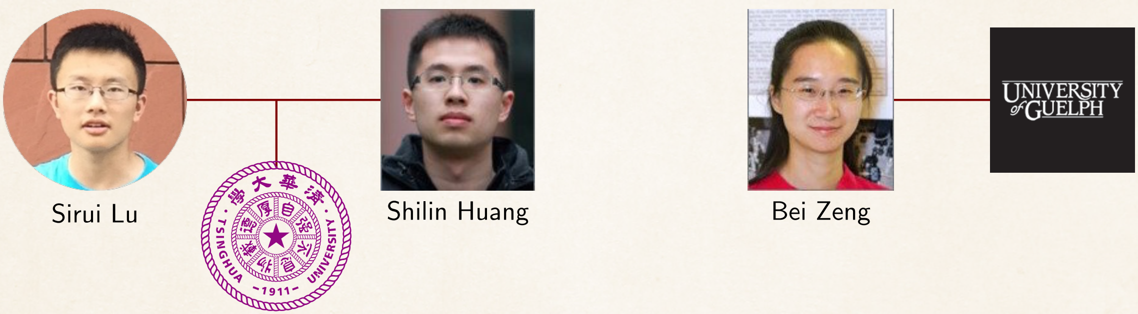
Joint postdoctoral fellow from

BEJING COMPUTATIONAL SCIENCE RESEARCH CENTER

UNIVERSITY OF WATERLOO | IQC Institute for Quantum Computing

Coauthors

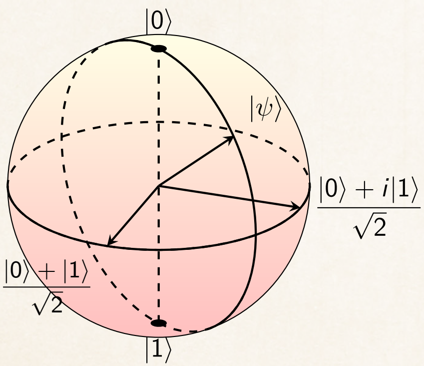
Sirui Lu, Shilin Huang, Keren Li, Jun Li, Jianxin Chen, Dawei Lu, Zhengfeng Ji, Yi Shen, Duanlu Zhou and Bei Zeng.



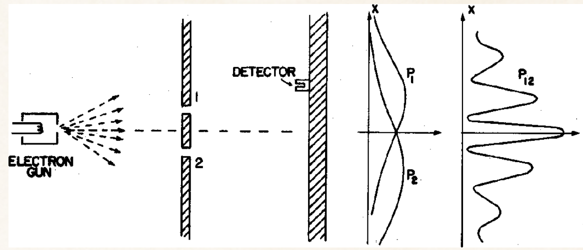
Outline

Quantum Features

Quantum Superposition



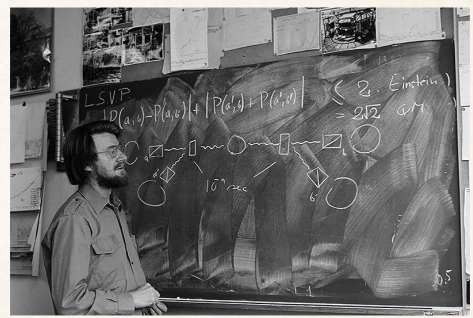
Wave-Particle Duality



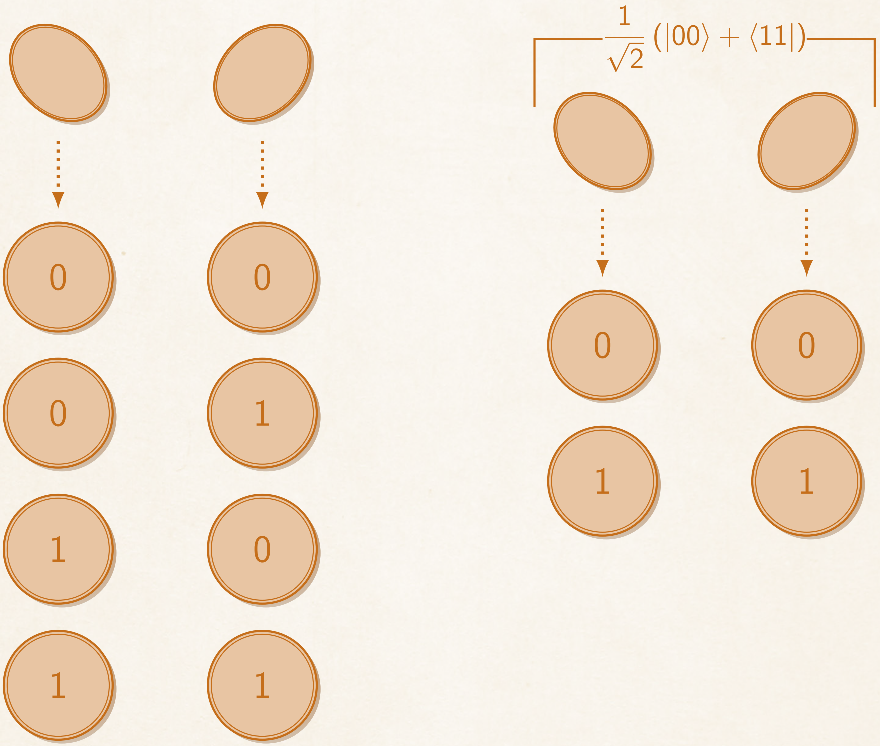
Uncertainty Principle

$$\Delta x \cdot \Delta p \geq \frac{\hbar}{2}$$

Nonlocality: Quantum Entanglement



EXAMPLE. Uncorrelated vs Entangled. Measurement of one particle affects the state of the other; no classical model of this behavior.



Entanglement Lies in the Heart of Quantum Information

Applications:

- ➢ Quantum cryptography
- ➢ Quantum teleportation
- ➢ Quantum computing
- ➢ Quantum metrology
- ➢ ...

Understanding and characterization of entanglement is one of the fundamental open problems in quantum mechanics.

China's satellite-based distribution of entangled photon pairs over 1,200 kilometers

Quantum State

A quantum state is described by a density matrix ρ:

(1) self-adjoint;

(2) semipositive definiteness $\rho \geq 0$;

(3) normalization: $\mathrm{Tr}\,\rho = 1$.

> The set of quantum states is convex.

Separable State

Consider a bipartite system AB with the Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$, where $\mathcal{H}_A$ has dimension d_A and $\mathcal{H}_B$ has dimension d_B, respectively.

For example, a qubit-qutrit system is a 2×3 system.

DEFINITION. A pure state $|\psi\rangle$ is called separable iff it can be written as

$$|\psi\rangle = |\psi_A\rangle \otimes |\psi_B\rangle,$$

otherwise it is entangled.

EXAMPLE. Bell states are entangled pure states.

Definition. A mixed state ρ_{AB} is separable if it can be written as a convex combination

$$\rho_{AB} = \sum_i \lambda_i \rho_{A,i} \otimes \rho_{B,i}$$

with a probability distribution $\lambda_i \geq 0$ and $\sum_i \lambda_i = 1$. Here $\rho_{A,i}$ and $\rho_{B,i}$ act on $\mathcal{H}_A$, $\mathcal{H}_B$ respectively. Otherwise, ρ_{AB} is entangled.

Example. 2×2 Werner states are defined as[1]

$$\rho = p|\psi\rangle\langle\psi| + (1-p) I \otimes I/4,$$

where

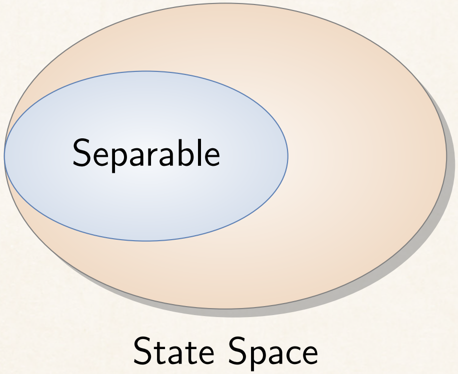
$$|\psi\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle).$$

When $p > 1/3$, ρ is entangled.

[1] R. F. Werner, Phys. Rev. A 40, 4277 (1989).

The Set of Separable States

The set of separable states is convex.



Separability Problem

To determine whether a given state is separable or entangled.

It is known that, generally, this problem is NP-hard[2].

There exist many methods and many many paper studying this problem.

- positive partial transpose criterion
- linear (or nonlinear) entanglement witness
- symmetric extension
- ...

[2] L. Gurvits, in Proceedings of the thirty-fifth annual ACM symposium on Theory of computing (ACM, 2003) pp. 1019.

Positive Partial Transpose Criterion (Peres-Horodecki)

If we write a $d_A \times d_B$ mixed state ρ as block matrix

$$\rho = \begin{bmatrix} \rho_{11} & \cdots & \rho_{1d_A} \\ \vdots & \ddots & \vdots \\ \rho_{d_A 1} & \cdots & \rho_{d_A d_A} \end{bmatrix},$$

where each ρ_{ij} $(i, j = 1, \ldots, d_A)$ is $d_B \times d_B$. The partial transpose operation:

$$\rho^{T_B} = \begin{bmatrix} \rho_{11}^T & \cdots & \rho_{1d_A}^T \\ \vdots & \ddots & \vdots \\ \rho_{d_A 1}^T & \cdots & \rho_{d_A d_A}^T \end{bmatrix}.$$

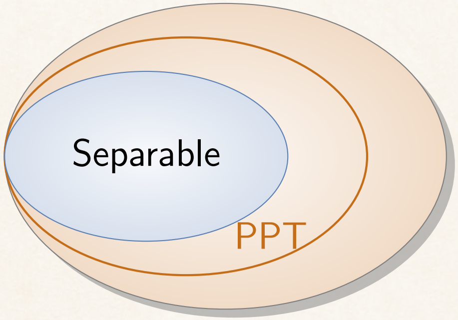
Definition. If the partial transpose ρ^{T_B} has non-negative eigenvalues, we say the state ρ fulfills the positive partial transpose (PPT) condition.

The following provides a separability criterion[3,4].

PROPOSITION.

(1) If ρ is separable, then it is PPT.

(2) The converse is true only for 2×2 or 2×3 systems.



State Space

[3] A. Peres, Separability criterion for density matrices. Phys. Rev. Lett., 77, 1413 (1996).
[4] M. Horodechi, P. Horodecki and R. Horodecki, Phys. Lett. A 223, 1 (1996).

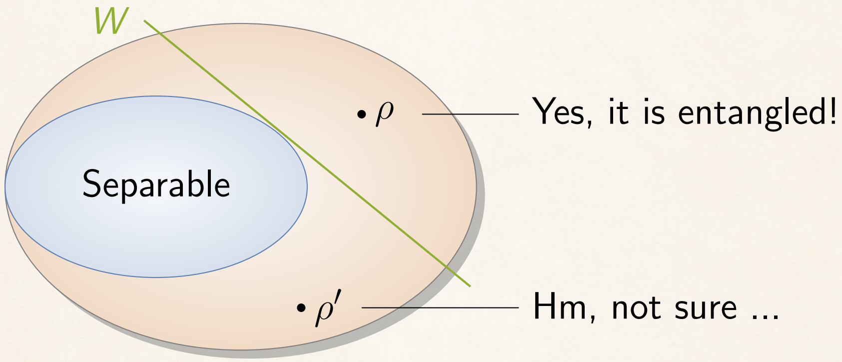
Entanglement Witness

For every entangled state ρ, there exists an observable W such that

(1) $\mathrm{Tr}(W\rho) < 0$;

(2) $\mathrm{Tr}(W\sigma) \geq 0$, if σ is separable.

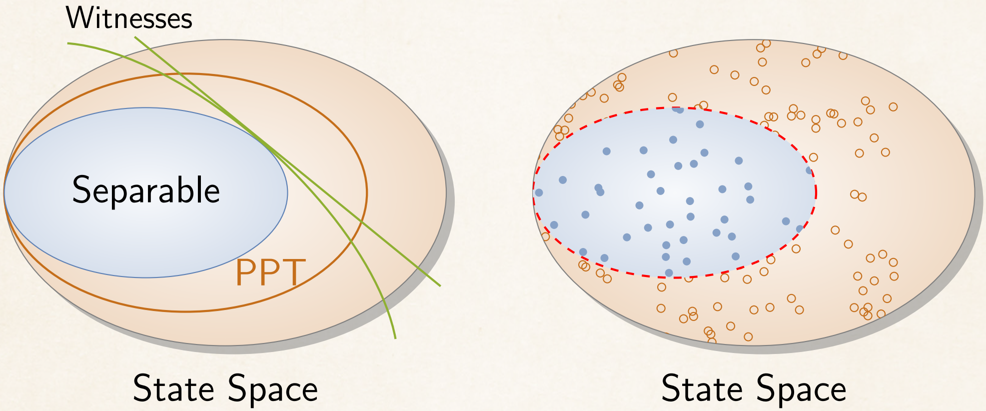
W: entanglement witness.



M. Horodechi, P. Horodecki and R. Horodecki, Phys. Lett. A 223, 1 (1996).

Machine Learning Approach

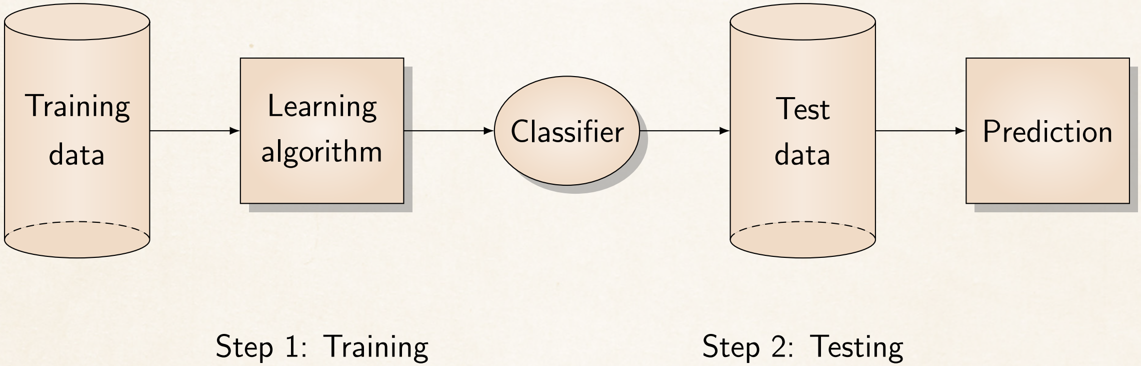
Many criteria only detect a limited set of entangled states. In contrast, if we take machine learning approach to build a classifier. This classifier, after sufficient traning, can handle a variety of input states.



The bipartite separability problem can be formulated as a supervised binary classification task.

Classification

- Learning (training): learn a model using the training data;
- Testing: test the model using unseen test data to assess the model accuracy.



Feature Vector Representation

Definition. Any quantum state ρ acting on $\mathcal{H}_A \otimes \mathcal{H}_B$ can be represented as a real vector $x \in X := \mathbb{R}^{d_A^2 d_B^2 - 1}$. We refer x as the feature vector of ρ and X the feature space.

Training Dataset

$\mathcal{D}_{\text{train}} = \{(x_1, y_1), (x_2, y_2), \ldots\}$, $x_i \in X$ is the i-th sample, and y_i is its label:

$$\begin{cases} y_i = 1, & \text{if } x_i \text{ is entangled,} \\ y_i = -1, & \text{if } x_i \text{ is separable.} \end{cases}$$

Inferring a Classifer

The aim of supervised learning is to infer a classifier $h : X \to \{-1, 1\}$, where h is expected to be close to the true decision function.

To evaluate how well h fits the training data $\mathcal{D}_{\text{train}}$, a loss function is defined as

$$\mathcal{L}(h, \mathcal{D}_{\text{train}}) = \frac{1}{|\mathcal{D}_{\text{train}}|} \sum_{(x_i, y_i) \in \mathcal{D}_{\text{train}}} \mathbf{1}(y_i \neq h(x_i)), \tag{1}$$

Generalization

For a generic new input test dataset $\mathcal{D}_{\text{test}}$ that contains previously unseen data, function $\mathcal{L}(h, \mathcal{D}_{\text{test}})$ gives a quantification of the generalization error from $\mathcal{D}_{\text{train}}$ to $\mathcal{D}_{\text{test}}$.

Learning Algorithm

Premilinary Trials

Error rate of classifiers trained by different algorithms.

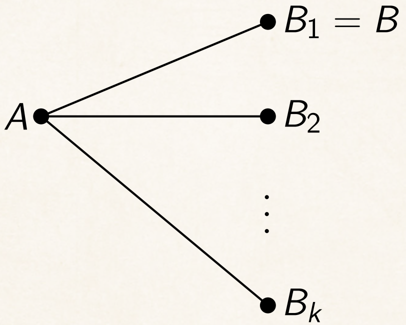
Method	Bagging	Boosting	SVM	Decision Tree
Error (%)	12.03	14.8	8.4	23.3

> The error rate is difficult to be reduced. This suggests that the boundary of separable states is of very complicated shape.

Maybe we can try to add some prior information to the learning process.

Approximate From Outside—k-Extendible States

Definition. A bipartite quantum state ρ_{AB} is k-symmetric extendible if there is a quantum state $\rho_{AB_1 \cdots B_k}$ whose martials on A, B_i all equal to ρ_{AB}.



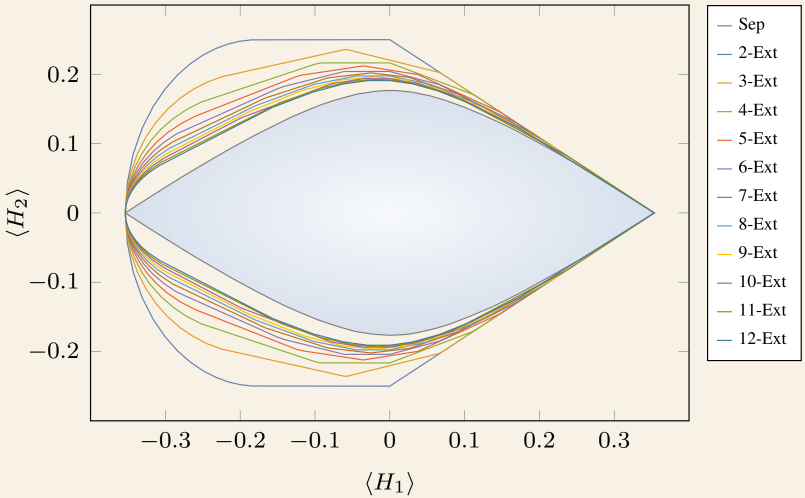
Let E_k denote the set of k-symmetric extendible set. Then[5,6]

$$E_1 \supset E_2 \supset \cdots E_\infty = \mathcal{S}.$$

[5] A. C. Doherty, Pablo A. Parrilo, and Federico M. Spedalieri. Phys. Rev. Lett. 88, 187904 (2002).

[6] Andrew C. Doherty, Pablo A. Parrilo, and Federico M. Spedalieri. Phys. Rev. A 69, 022308 (2004).

Example. Two-qubit case:



Here

$$H_1 = |0\rangle\langle 0| \otimes \sigma_z/\sqrt{2},$$

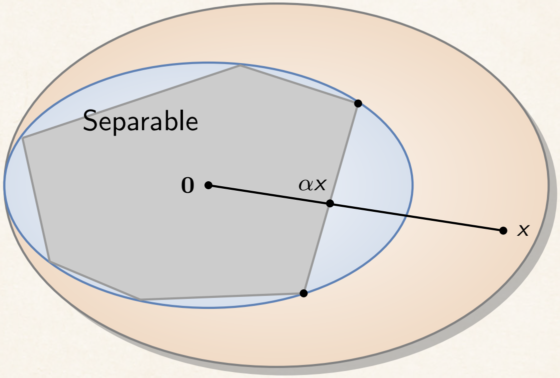
$$H_2 = (\sigma_y \otimes \sigma_x - \sigma_x \otimes \sigma_y)/2.$$

Approximate From Inside—Convex Hull

We randomly sample m separable pure states $c_1, \ldots, c_m \in X$ to form a convex hull $\mathcal{C} = \text{conv}(\{c_1, \ldots, c_m\})$. With $\mathcal{C}$ we can approximately tell whether a state ρ is separable or not by testing if it is in $\mathcal{C}$.

$$\max \quad \alpha \quad \text{s.t} \quad \alpha x \in \mathcal{C},$$

$$\text{i.e.} \quad \alpha x = \sum_{i=1}^{m} \lambda_i c_i, \quad \lambda_i \geq 0, \quad \sum\nolimits_i \lambda_i = 1.$$



A Two-qutrit Example

Consider a 2-qutrit state[7]:

$$\rho_p = p\mathbb{I}/(d_A d_B) + (1-p)\rho_{\text{tiles}}, \quad 0 \le p \le 1,$$

where

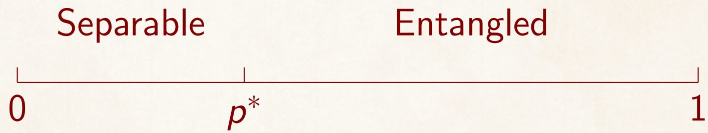
$$\rho_{\text{tiles}} = (\mathbb{I} - \sum_{i=1}^{5} |v_i\rangle\langle v_i|)/4.$$

where

$$|v_1\rangle = (|00\rangle - |01\rangle)/\sqrt{2},$$
$$|v_2\rangle = (|21\rangle - |22\rangle)/\sqrt{2},$$
$$|v_3\rangle = (|02\rangle - |12\rangle)/\sqrt{2},$$
$$|v_4\rangle = (|10\rangle - |20\rangle)/\sqrt{2},$$
$$|v_5\rangle = (|0\rangle + |1\rangle + |2\rangle)^{\otimes 2}/3.$$

[7] I. Bengtsson and K. Zyczkowski, Geometry of quantum states: an introduction to quantum entanglement (Cambridge University Press, 2007).

It is known that ρ_{tiles} is entangled. There must exist a critical point $p^* \in [0, 1)$



Result

In a lecture note by Nathaniel Johnston[8]:

> ... Thus there is a rather large gap of values $p \in (0.1351, 0.4357)$ where we do not know whether or not ρ_p is entangled.

Size of $\mathcal{C}$	2000	5000	10000	20000	50000	...
deteced p^*	0.4763	0.4132	0.3613	0.3241	0.2850	...

Our result:

$p^* \in (0.1351, 0.1352)$

[8] N. Johnston, "Entanglement detection," (2014).

Combining Convex Hull Approximation and Supervised Learning

To incorparating the convex hull information into the dataset, We extend the feature vector

$$x \to (x, \alpha(\mathcal{C}, x)).$$

So the new feature vector contains the geometric information of the convex hull.

Correspondingly, the training dataset is

$$\mathcal{D}_{\text{train}} = \{(x_1, \alpha_1, y_1), ..., (x_n, \alpha_n, y_n)\}.$$

Numerical Tests on Two-qubit and Two-qutrit Systems

Our learning precedure:

(1) build a convex hull approximation

(2) adding information of the convex hull in the training dateset

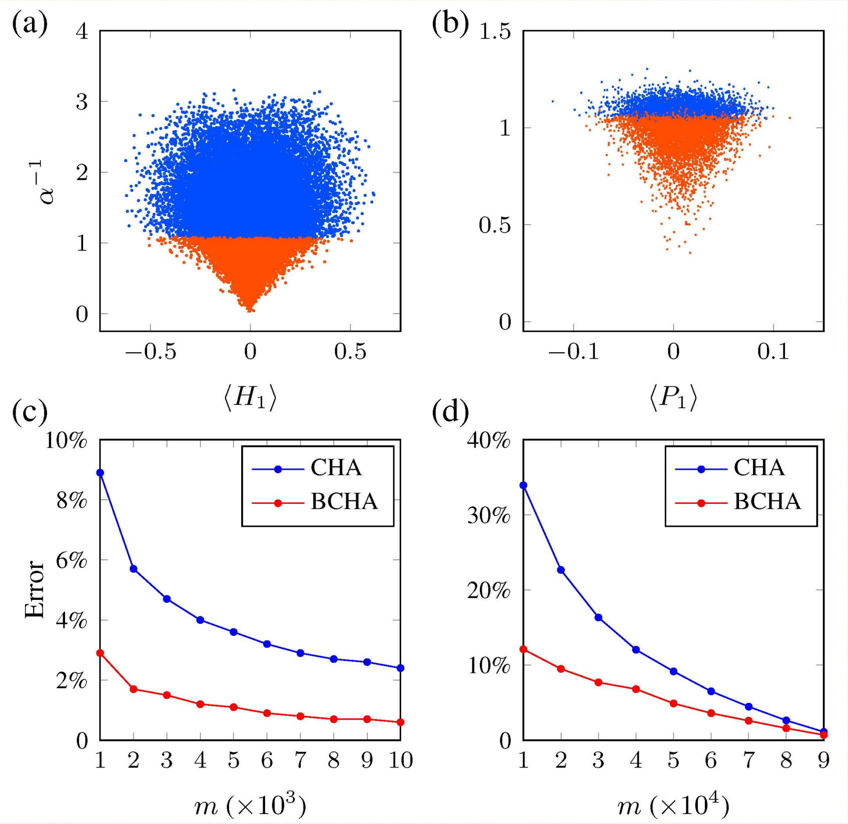
(3) randomly drawing a training subset from the whole dateset

(4) decision trees learning: build a sub-classifer for the training subset

(5) repeat step (3); (4) many times

(6) ensemble learning: combine the sub-classifiers to form the final classifier

Results


(a)
α^{-1}
$\langle H_1 \rangle$
(b)
$\langle P_1 \rangle$
(c)
Error
CHA
BCHA
$m\ (\times 10^3)$
(d)
CHA
BCHA
$m\ (\times 10^4)$

Conclusion

Outperform the existing popular entanglement detection methods in accuracy and speed in the case of 2×2 and 3×3 systems.

THE END THANKS!