清華大學
Tsinghua University

Machine Learning Designation of the Chiral Magnetic Effect Observable

Reporter: Ziyi Liu

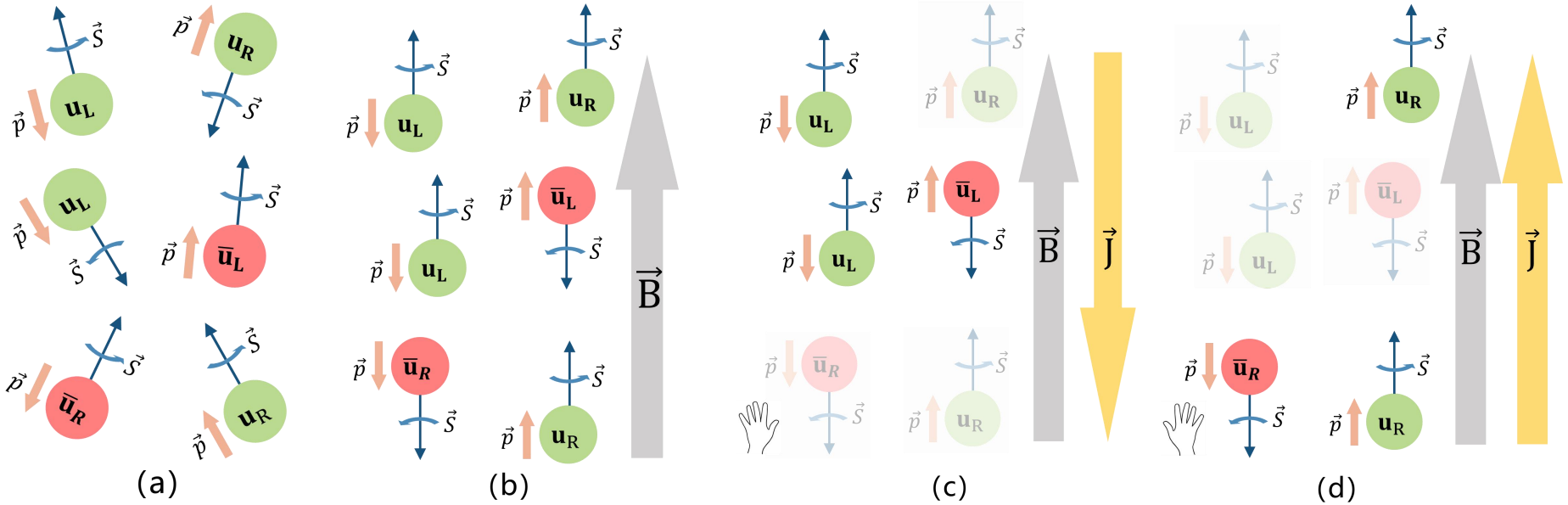
Author: Yuji Hirono, Kazuki Ikeda, Dmitri E. Kharzeev, **Ziyi Liu,** Shuzhe Shi

Contents

- Background
- Method
  - **Linear combination:** construct potential observable
  - **Neural network:** identify the signal
- Result
- Summary & Outlook

The Chiral Magnetic Effect

$$\vec{J} = \frac{e^2}{2\pi^2}\mu_5\vec{B} \qquad \mu_5 \sim \mu_R - \mu_L$$

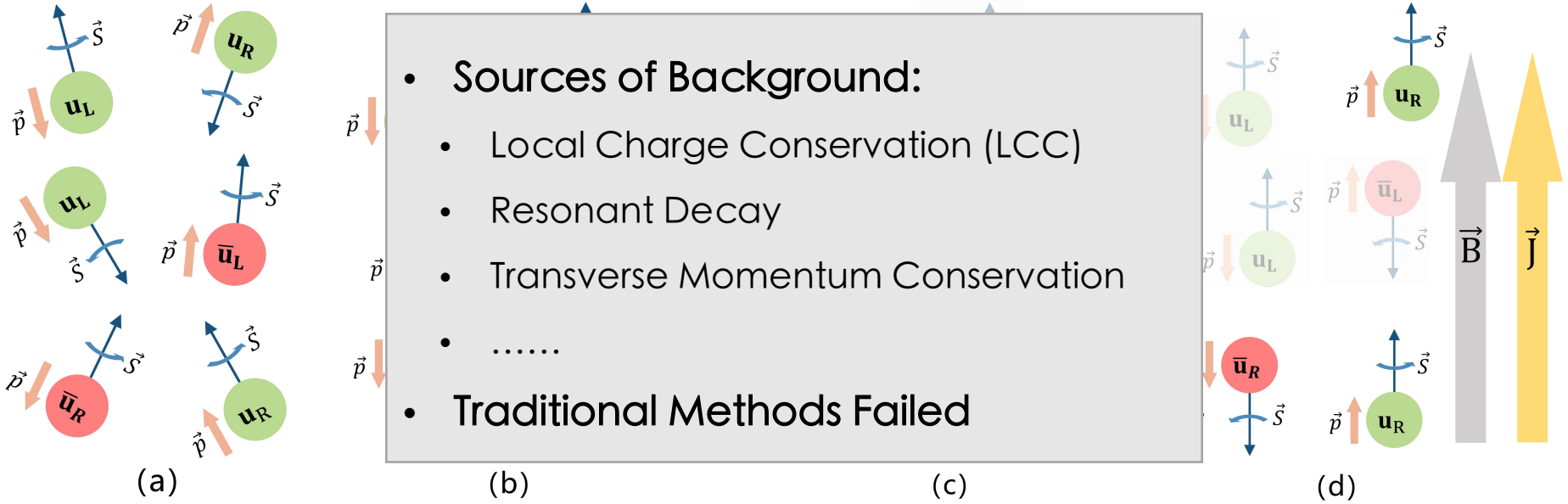


[1] D. E. Kharzeev, L. D. McLerran, and H. J. Warringa, Nuclear Physics A 803, 227 (2008).
[2] D. E. Kharzeev, Progress in Particle and Nuclear Physics 75, 133 (2014)

The Chiral Magnetic Effect

$$\vec{J} = \frac{e^2}{2\pi^2}\mu_5\vec{B} \qquad \mu_5 \sim \mu_R - \mu_L$$

- Sources of Background:
  - Local Charge Conservation (LCC)
  - Resonant Decay
  - Transverse Momentum Conservation
  - ……
- Traditional Methods Failed



[4] J. Zhao and F. Wang, Prog. Part. Nucl. Phys. 107, 200 (2019)
[5] D. E. Kharzeev, *et al*, Int. J. Mod. Phys. E 33, 2430007 (2024)
[6] S. Choudhury, *et al*. Chin. Phys. C 46, 014101 (2022)

Contents

Method 1

Linear combination: construct potential observable

Method 1: Gradient descent

- Define the harmonic coefficients

$$v_n^{(m)} \equiv \frac{\langle p_T^m \cos(n\varphi_n)\rangle}{\langle p_T^m \rangle}, \quad a_n^{(m)} \equiv \frac{\langle p_T^m \sin(n\varphi_n)\rangle}{\langle p_T^m \rangle}$$

- Separate them into P-odd and P-even observables

$$\{o\} \equiv \left\{ \left(a_{2j-1,c}^{(m)}, v_{2j-1,c}^{(m)} \right) \right\}_{\substack{j\in\mathbb{N},\, 0<j\leq N/2, \\ m\in\{0,\ldots,M\}, \\ c\in\{\pi^+,\pi^-\}}}$$

$$\{e\} \equiv \left\{ \frac{1}{N_{\mathrm{ch}}} \right\} \bigcup \left\{ \left(a_{2j,c}^{(m)}, v_{2j,c}^{(m)} \right) \right\}_{\substack{j\in\mathbb{N},\, 0<j\leq N/2, \\ m\in\{0,\ldots,M\}, \\ c\in\{\pi^+,\pi^-\}}}$$

Construct & Optimize Observables

- Construct a potential CME observable

$$O = \sum_j X_j^{(\mathrm{L})}\, e_j + \sum_{j,j'} X_{j,j'}^{(\mathrm{E})}\, e_j\, e_{j'} + \sum_{j,j'} X_{j,j'}^{(\mathrm{O})}\, o_j\, o_{j'}$$

- Using **Anomalous Viscous Fluid Dynamics**(AVFD), generate data:

$$O_{\mathrm{C,L}}\, O_{\mathrm{C,}\cancel{\mathrm{L}}}, O_{\cancel{\mathrm{C}},\mathrm{L}}, O_{\cancel{\mathrm{C}},\cancel{\mathrm{L}}}$$

- Using the **gradient descent** to iterate the parameters $\left\{X_j^{(\mathrm{L})}, X_{j,j'}^{(\mathrm{E})}, X_{j,j'}^{(\mathrm{O})}\right\}$, and minimize loss function.

- Construct loss function

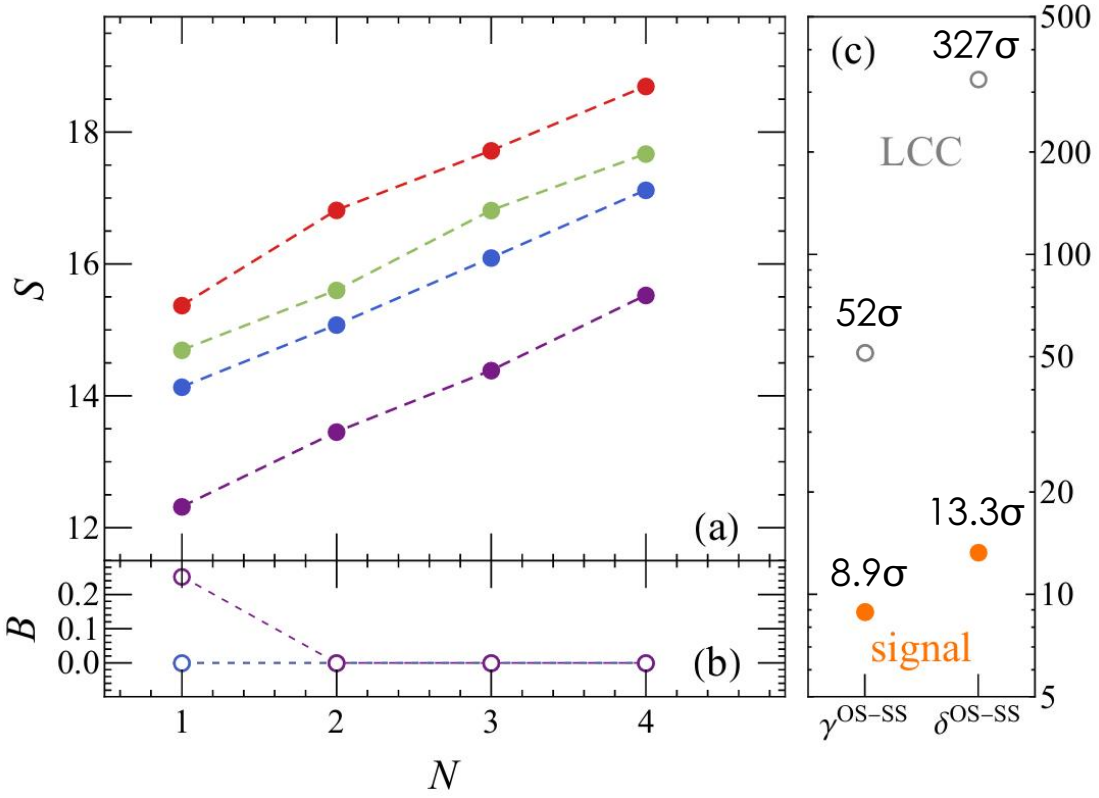
$$\mathcal{L} = \lambda \left(\frac{(O_{\mathrm{C,L}} - O_{\mathrm{C,}\cancel{\mathrm{L}}})^2}{\Delta_{\mathrm{C,L}}^2 + \Delta_{\mathrm{C,}\cancel{\mathrm{L}}}^2} + \frac{(O_{\cancel{\mathrm{C}},\mathrm{L}} - O_{\cancel{\mathrm{C}},\cancel{\mathrm{L}}})^2}{\Delta_{\cancel{\mathrm{C}},\mathrm{L}}^2 + \Delta_{\cancel{\mathrm{C}},\cancel{\mathrm{L}}}^2} \right) - \left(\frac{(O_{\mathrm{C,L}} - O_{\cancel{\mathrm{C}},\mathrm{L}})^2}{\Delta_{\mathrm{C,L}}^2 + \Delta_{\cancel{\mathrm{C}},\mathrm{L}}^2} + \frac{(O_{\mathrm{C,}\cancel{\mathrm{L}}} - O_{\cancel{\mathrm{C}},\cancel{\mathrm{L}}})^2}{\Delta_{\mathrm{C,}\cancel{\mathrm{L}}}^2 + \Delta_{\cancel{\mathrm{C}},\cancel{\mathrm{L}}}^2} \right)$$

Minimize the background

Maximize the CME signal.

[7] S. Shi, Y. Jiang, E. Lilleskov, and J. Liao, Annals Phys. 394, 50 (2018)
[8] Y. Jiang, S. Shi, Y. Yin, and J. Liao, Chin. Phys. C 42, 011001 (2018)

Background Suppression & Sensitivity



- sig./bkg. to statistics ratio

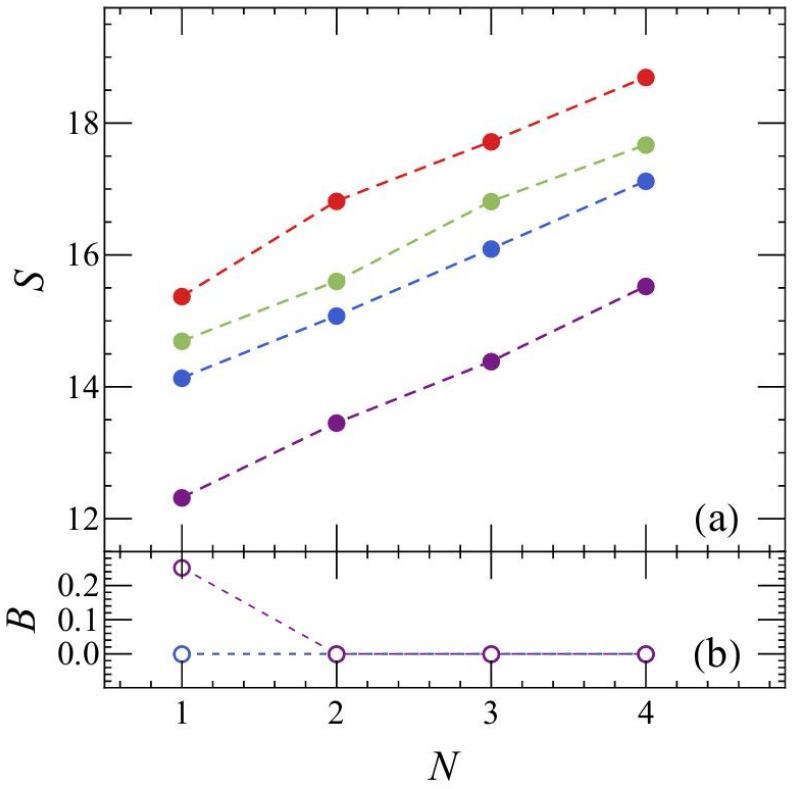
$$S = \frac{1}{2}\left(\frac{O_{C,L} - O_{\bar{C},L}}{\sqrt{\frac{\Delta^2_{C,L}+\Delta^2_{\bar{C},L}}{N_{ev}}}} + \frac{O_{C,\bar{L}} - O_{\bar{C},\bar{L}}}{\sqrt{\frac{\Delta^2_{C,\bar{L}}+\Delta^2_{\bar{C},\bar{L}}}{N_{ev}}}}\right)$$

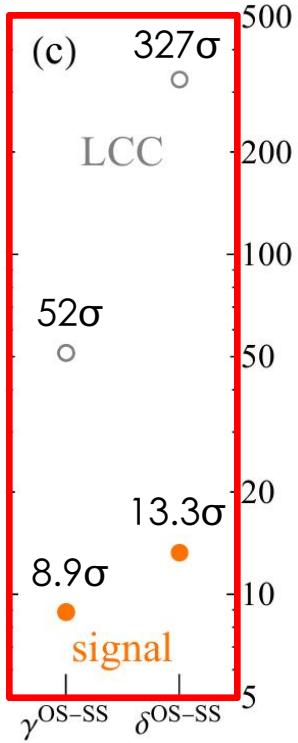
$$B = \frac{1}{2}\left(\frac{O_{C,L} - O_{C,\bar{L}}}{\sqrt{\frac{(\Delta^2_{C,L}+\Delta^2_{C,\bar{L}})}{N_{ev}}}} + \frac{O_{\bar{C},L} - O_{\bar{C},\bar{L}}}{\sqrt{\frac{(\Delta^2_{\bar{C},L}+\Delta^2_{\bar{C},\bar{L}})}{N_{ev}}}}\right)$$

harmonic order N=1-4,

p_T weight M=0-3

Background Suppression & Sensitivity





- sig./bkg. to statistics ratio

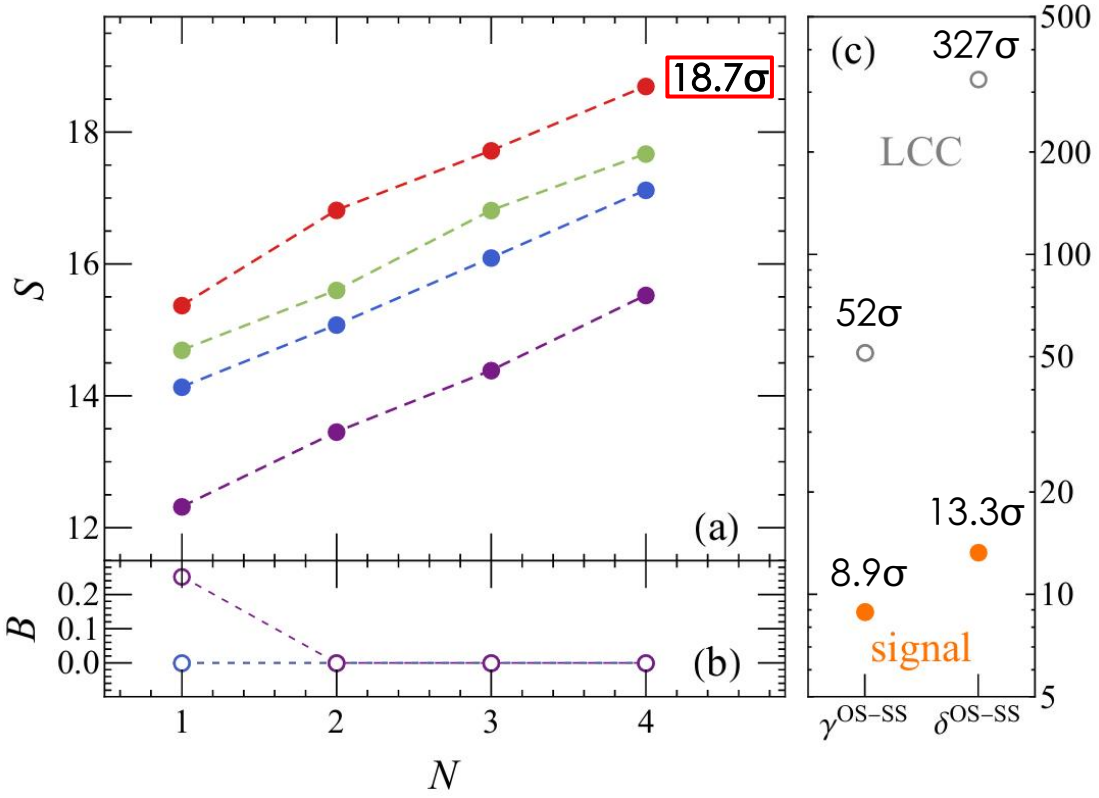
$$S = \frac{1}{2}\left(\frac{O_{C,L} - O_{\bar{C},L}}{\sqrt{\frac{\Delta^2_{C,L}+\Delta^2_{\bar{C},L}}{N}}} + \frac{O_{C,\bar{L}} - O_{\bar{C},\bar{L}}}{\sqrt{\frac{\Delta^2_{C,\bar{L}}+\Delta^2_{\bar{C},\bar{L}}}{N}}} \right)$$

γ^{SS-OS} and δ^{SS-OS} :
sensitive to CME signal, but
more sensitive to LCC background.

harmonic order N=1-4,

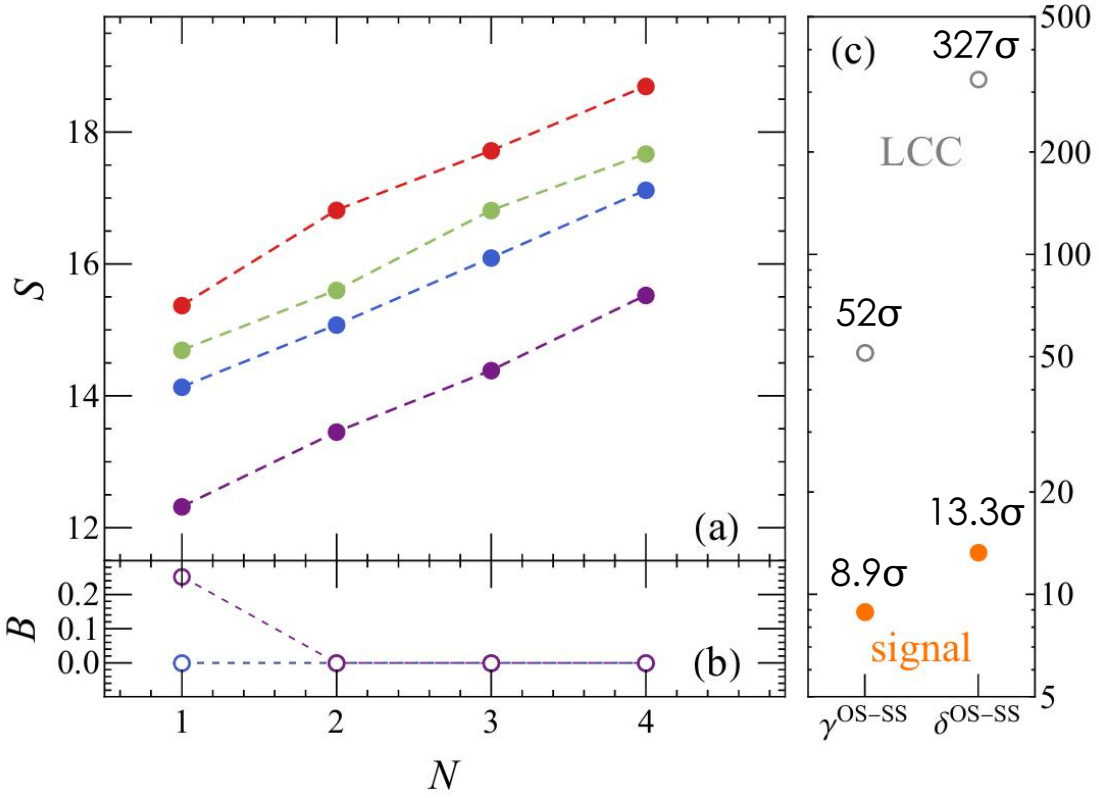
p_T weight M=0-3

Background Suppression & Sensitivity



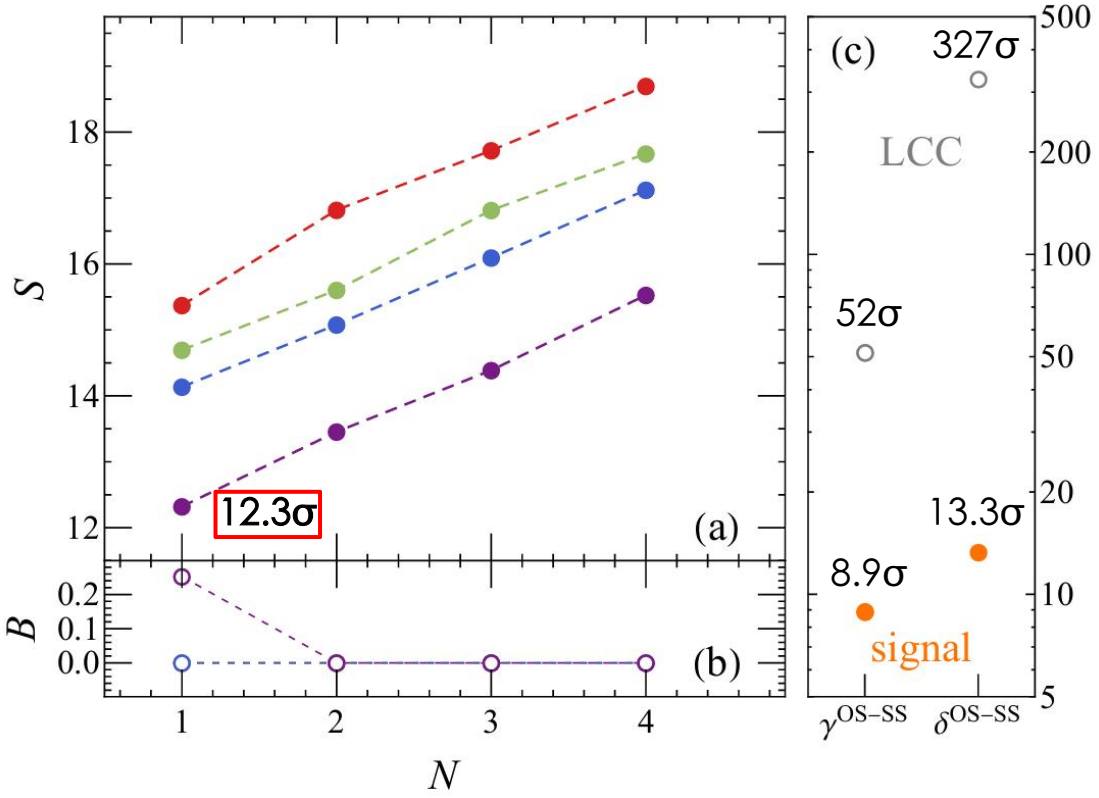
- Dramatically Improved Sensitivity.
- Complete Background Suppression.
- Effective Even at Low Complexity.

Background Suppression & Sensitivity



- Dramatically Improved Sensitivity.
- Complete Background Suppression.
- Effective Even at Low Complexity.

Background Suppression & Sensitivity



- Dramatically Improved Sensitivity.
- Complete Background Suppression.
- Effective Even at Low Complexity.

Contents

Method 2

Neural network: identify the signal

Method 2: Neural network

$$v_n^{(m)} \equiv \frac{\langle p_T^m \cos(n\varphi_n)\rangle}{\langle p_T^m\rangle}, \quad a_n^{(m)} \equiv \frac{\langle p_T^m \sin(n\varphi_n)\rangle}{\langle p_T^m\rangle}$$

p_T weight $\mathrm{m} = 0$

harmonic order $\mathrm{n} = 1\sim4$

C−parity conservation

$$e = (e_{\pi^+} + e_{\pi^-})/2$$
$$o = (o_{\pi^+} - o_{\pi^-})/2$$

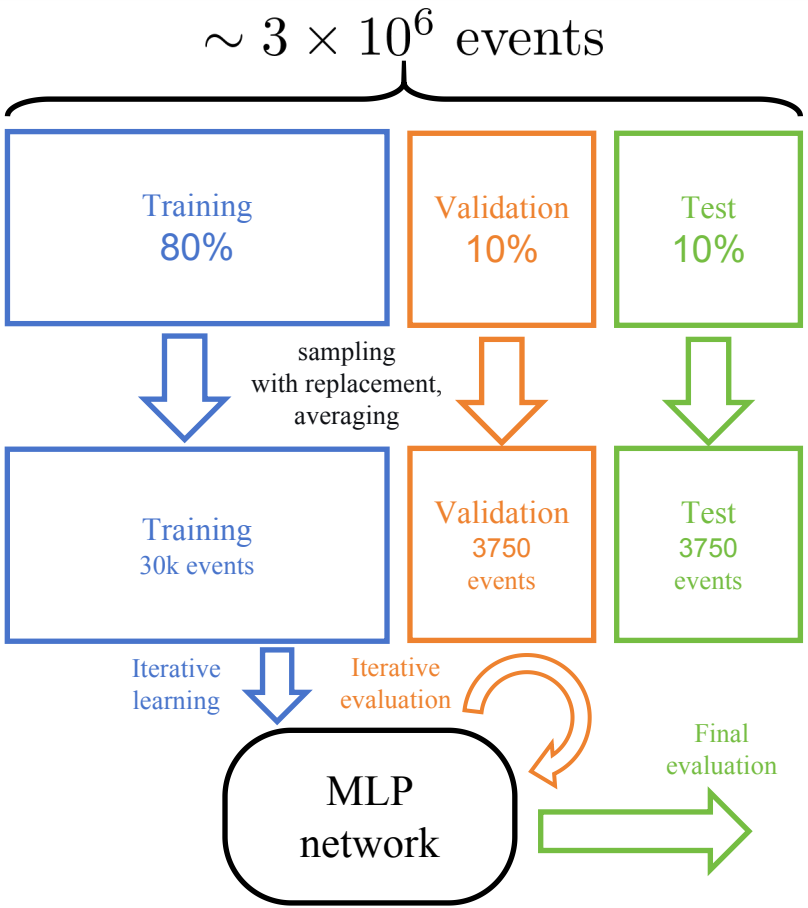
$$\frac{1}{N}, a_1, v_1, a_2, v_2, a_3, v_3, a_4, v_4$$

$$\frac{1}{N}, a_2, v_2, a_4, v_4$$
$$(\frac{1}{N})^2, (a_2)^2, (v_2)^2, (a_4)^2, (v_4)^2$$
$$\frac{1}{N}a_2, \frac{1}{N}v_2, \frac{1}{N}a_4, \frac{1}{N}v_4, a_2v_2, a_2a_4, a_2v_4, v_2a_4, v_2v_4, a_4v_4$$
$$(a_1)^2, (v_1)^2, (a_3)^2, (v_3)^2$$
$$a_1v_1, a_1a_3, a_1v_3, v_1a_3, v_1v_3, a_3v_3$$

P-even observables as inputs to the network

Dataset Construction

AVFD simulation framework
↓
Four datasets:
no CME & **no** LCC,
has CME & **no** LCC,
no CME & **has** LCC,
has CME & **has** LCC

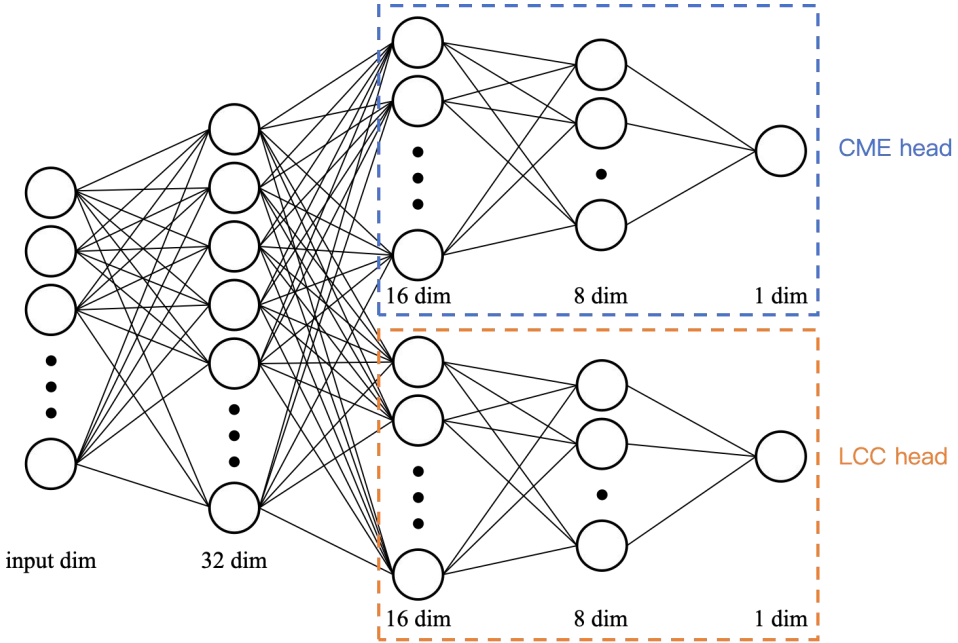


the number of averaged events

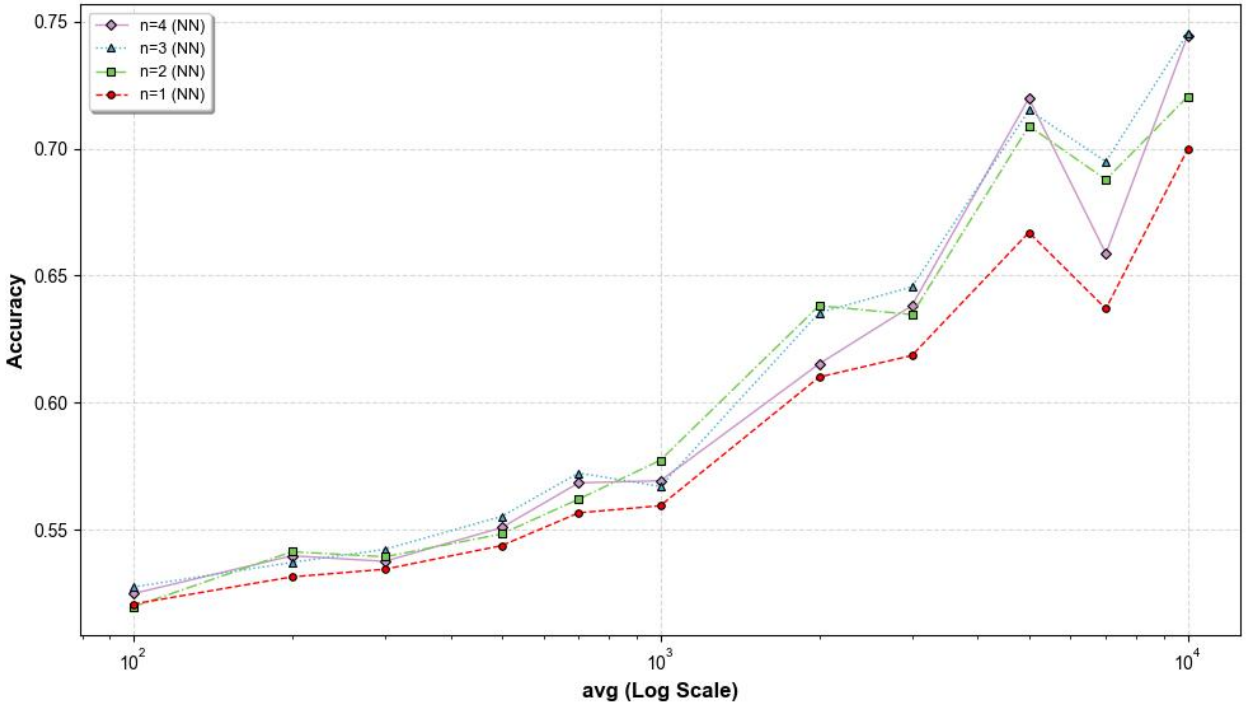
The average of ***X rows*** of data to generate a single row of data, and repeat this operation.

Network Architecture

- Multi-Layer Perceptron (MLP)
  - Typical representative of fully connected neural networks.
  - A shared backbone with two output heads, classifying CME and LCC.



Results at n=1~4 (CME)



n = 4

$$e(0-4): \frac{1}{N}, a_2, v_2, a_4, v_4$$
$$e^2(5-9): (\frac{1}{N})^2, (a_2)^2, (v_2)^2, (a_4)^2, (v_4)^2$$
$$e*e(10-19): \frac{1}{N}a_2, \frac{1}{N}v_2, \frac{1}{N}a_4, \frac{1}{N}v_4, a_2v_2, a_2a_4, a_2v_4, v_2a_4, v_2v_4, a_4v_4$$
$$o^2(20-23): (a_1)^2, (v_1)^2, (a_3)^2, (v_3)^2$$
$$o*o(24-29): a_1v_1, a_1a_3, a_1v_3, v_1a_3, v_1v_3, a_3v_3$$

n = 3

$$e(0-2): \frac{1}{N}, a_2, v_2$$
$$e^2(3-5): (\frac{1}{N})^2, (a_2)^2, (v_2)^2$$
$$e*e(6-8): \frac{1}{N}a_2, \frac{1}{N}v_2, a_2v_2$$
$$o^2(9-12): (a_1)^2, (v_1)^2, (a_3)^2, (v_3)^2$$
$$o*o(13-18): a_1v_1, a_1a_3, a_1v_3, v_1a_3, v_1v_3, a_3v_3$$

n = 2

$$e(0-2): \frac{1}{N}, a_2, v_2$$
$$e^2(3-5): (\frac{1}{N})^2, (a_2)^2, (v_2)^2$$
$$e*e(6-8): \frac{1}{N}a_2, \frac{1}{N}v_2, a_2v_2$$
$$o^2(9-10): (a_1)^2, (v_1)^2$$
$$o*o(11): a_1v_1$$

n = 1

$$e(0): \frac{1}{N}$$
$$e^2(1): (\frac{1}{N})^2$$
$$o^2(2-3): (a_1)^2, (v_1)^2$$
$$o*o(4): a_1v_1$$

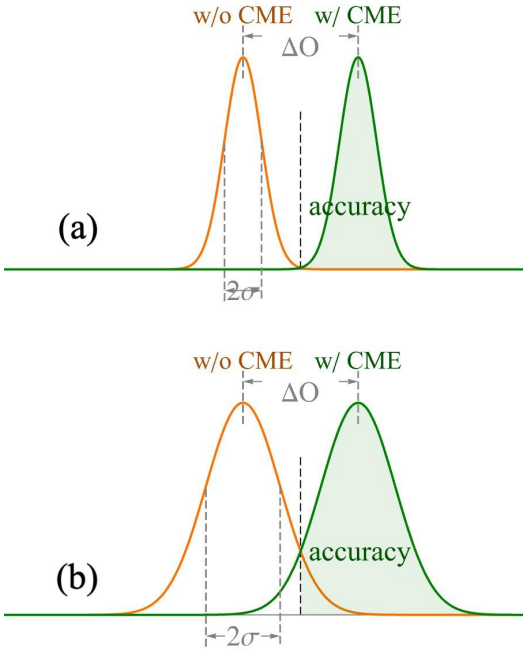
Theoretical of Avg - Acc Curves

definition of sensitivity $$S = \frac{f_c}{2}\left(\frac{O_{\mathrm{C,L}} - O_{\mathrm{C,\bar{L}}}}{\sqrt{\frac{\Delta^2_{\mathrm{C,L}}+\Delta^2_{\mathrm{C,\bar{L}}}}{N_{\mathrm{ev}}}}} + \frac{O_{\mathrm{C,\bar{L}}} - O_{\mathrm{C,L}}}{\sqrt{\frac{\Delta^2_{\mathrm{C,\bar{L}}}+\Delta^2_{\mathrm{C,L}}}{N_{\mathrm{ev}}}}}\right)$$

harmonic orders (N) = 1~4

Sensitivities of optimized observables: 12.32, 13.46, 14.39, 15.52

sensitivity
↓
overlap
↓
shadow area
↓
accuracy



Theoretical of Avg - Acc Curves

definition of sensitivity $S = \frac{f_c}{2}\left(\frac{O_{\mathrm{C,L}} - O_{\mathrm{G,L}}}{\sqrt{\frac{\Delta^2_{\mathrm{C,L}}+\Delta^2_{\mathrm{G,L}}}{N_{\mathrm{ev}}}}} + \frac{O_{\mathrm{C,R}} - O_{\mathrm{G,R}}}{\sqrt{\frac{\Delta^2_{\mathrm{C,R}}+\Delta^2_{\mathrm{G,R}}}{N_{\mathrm{ev}}}}}\right)$

harmonic orders (N) = 1~4

Sensitivities of optimized observables: 12.32, 13.46, 14.39, 15.52

averaging 10^6 events

Reduced to 1/100

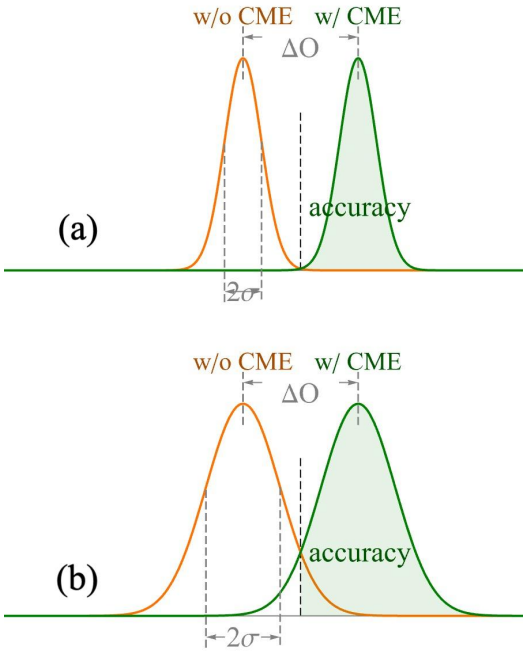
averaging 10^4 events

$S \propto \sqrt{N}$

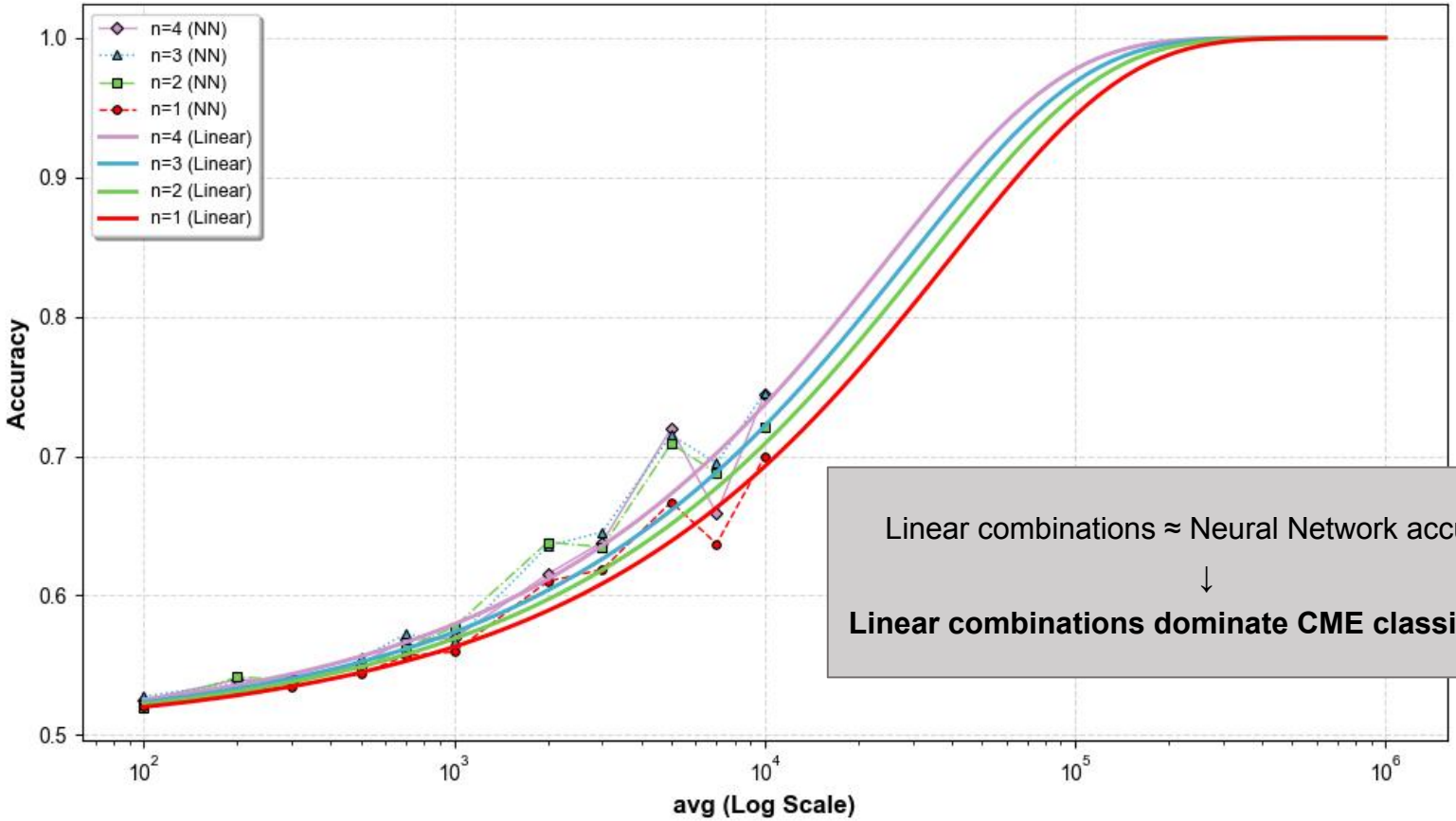
Reduced to 1/10

~13σ

~1.3σ



Theoretical and Experimental Comparison

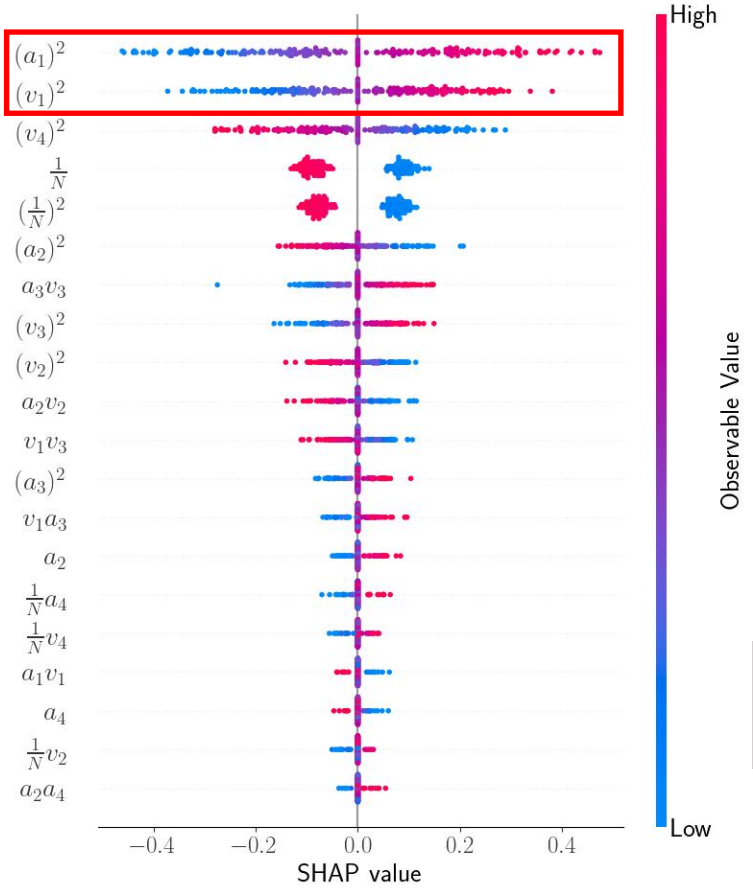


Linear combinations ≈ Neural Network accuracy

↓

Linear combinations dominate CME classification

Network Interpretability Analysis - SHAP

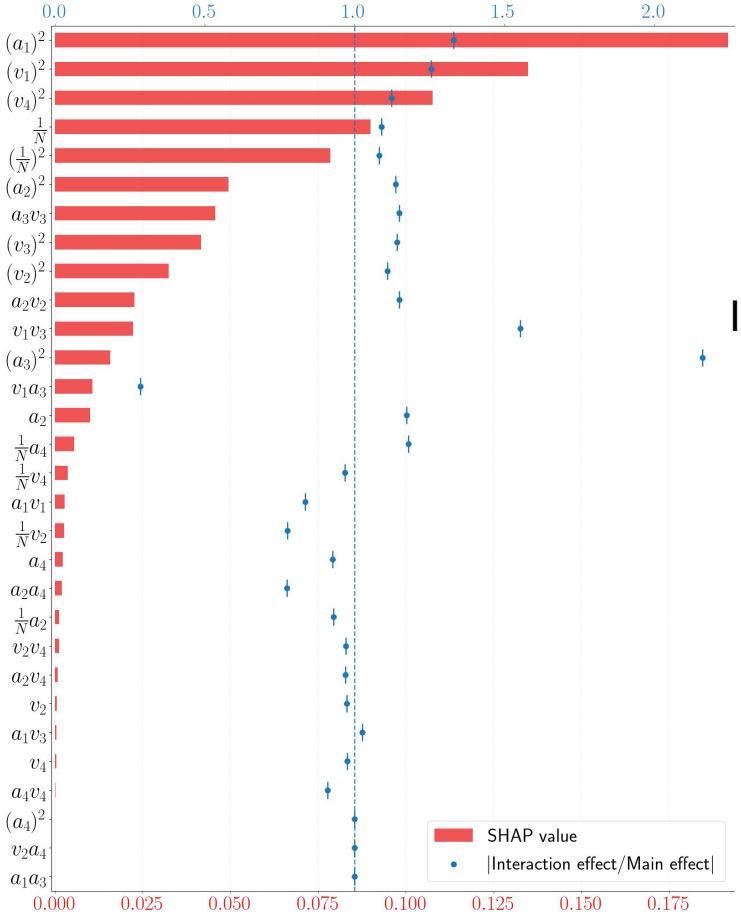


Features	SHAP value
$(a_1)^2$	0.195698
$(v_1)^2$	0.143833
$(v_4)^2$	0.115854
$\frac{1}{N}$	0.085744
$(\frac{1}{N})^2$	0.074707

≈ Contribution of Observable

- $(a_1)^2$ and $(v_1)^2$ are **key features** for CME classification, consistent with physical intuition.

Main Effect and Interaction Effect

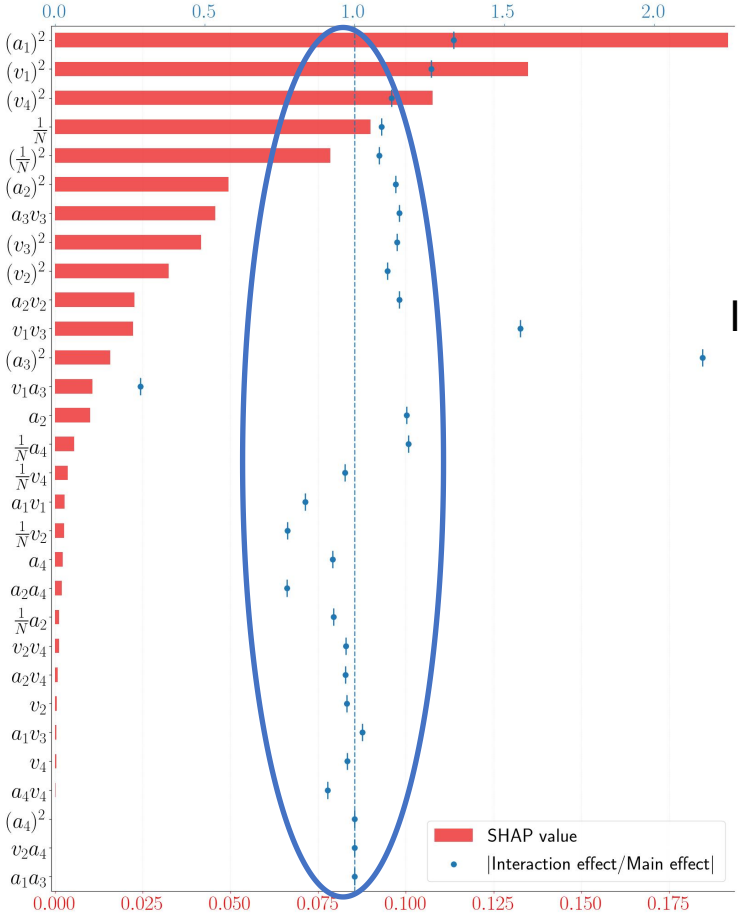


Contribution of Observables

Main Effect → Independent contribution

Interaction Effect → Cross synergy

Main Effect and Interaction Effect



Contribution of Observables

Main Effect → Independent contribution

Interaction Effect → Cross synergy

$$\left|\frac{\text{Interaction Effect}}{\text{Main Effect}}\right| \sim 1$$

Observables:

- **Independent predictive power**
- **Non-negligible interactions**

↓

Introduce more observables

Summary & Outlook

- **Method 1- Linear combination**
  - Optimised observable has higher sensitivity and near-zero background
- **Method 2- Neural network**
  - Accuracy achieves ~74%
  - Linear combinations dominate CME classification
  - Interpretability
- **Outlook**
  - Apply the method to measured Au + Au, isobar (Ru + Ru, Zr + Zr) and Pb + Pb data
  - Add particle-species and centrality dependence
  - Extend to higher-order harmonics



Thank you !

清華大學 Tsinghua University