(2025.10.14—2025.10.17, Korea Institute for Advanced Study, Seoul)
The 7th East Asia Joint Seminars on Statistical Physics

Lateral Predictive Coding

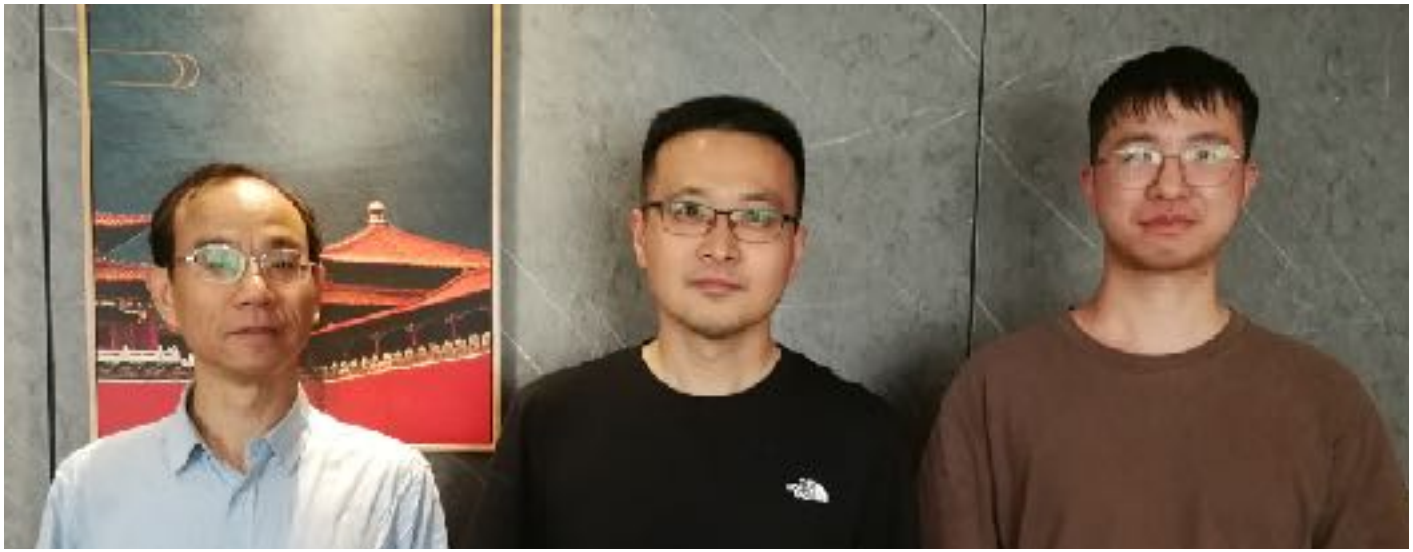
Hai-Jun Zhou （周海军)

Institute of Theoretical Physics, Chinese Academy of Sciences, Beijing
(中国科学院理论物理研究所，北京)





Collaborators



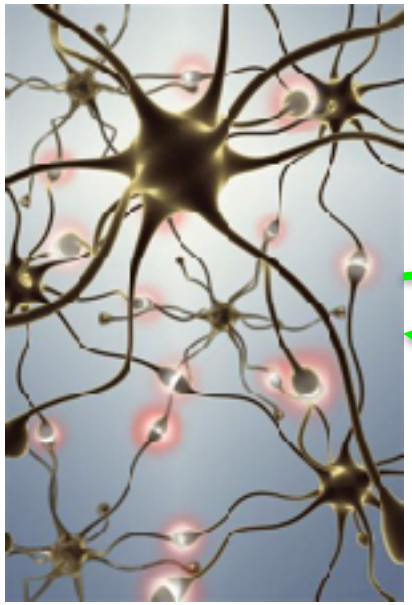
Weikang Wang
(王维康)

Zhenye Huang
（黄振业）

Outline

1. **Motivation**
2. Problem settings
3. Theoretical framework
4. Detecting one non-Gaussian feature
5. Detecting and separating two non-Gaussian features
6. Conclusion

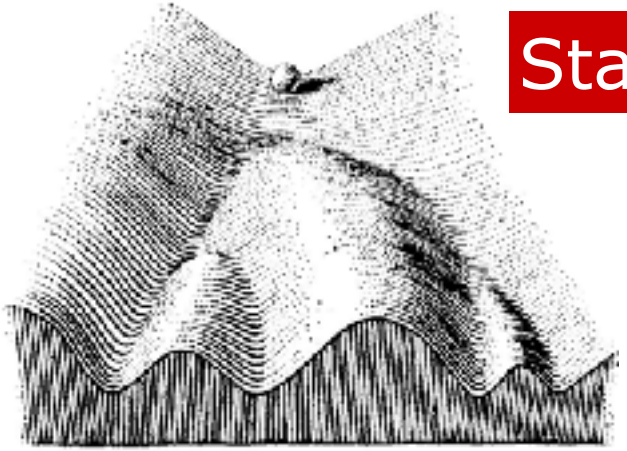
Intelligence Complexity



neuro-electric ——> action

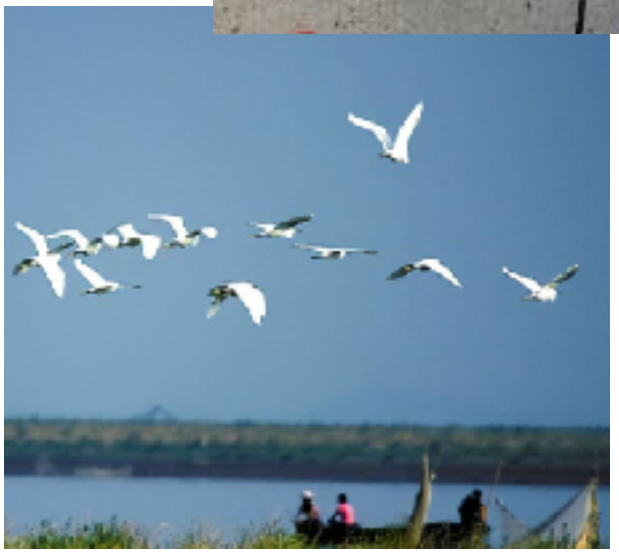
learning & feedback



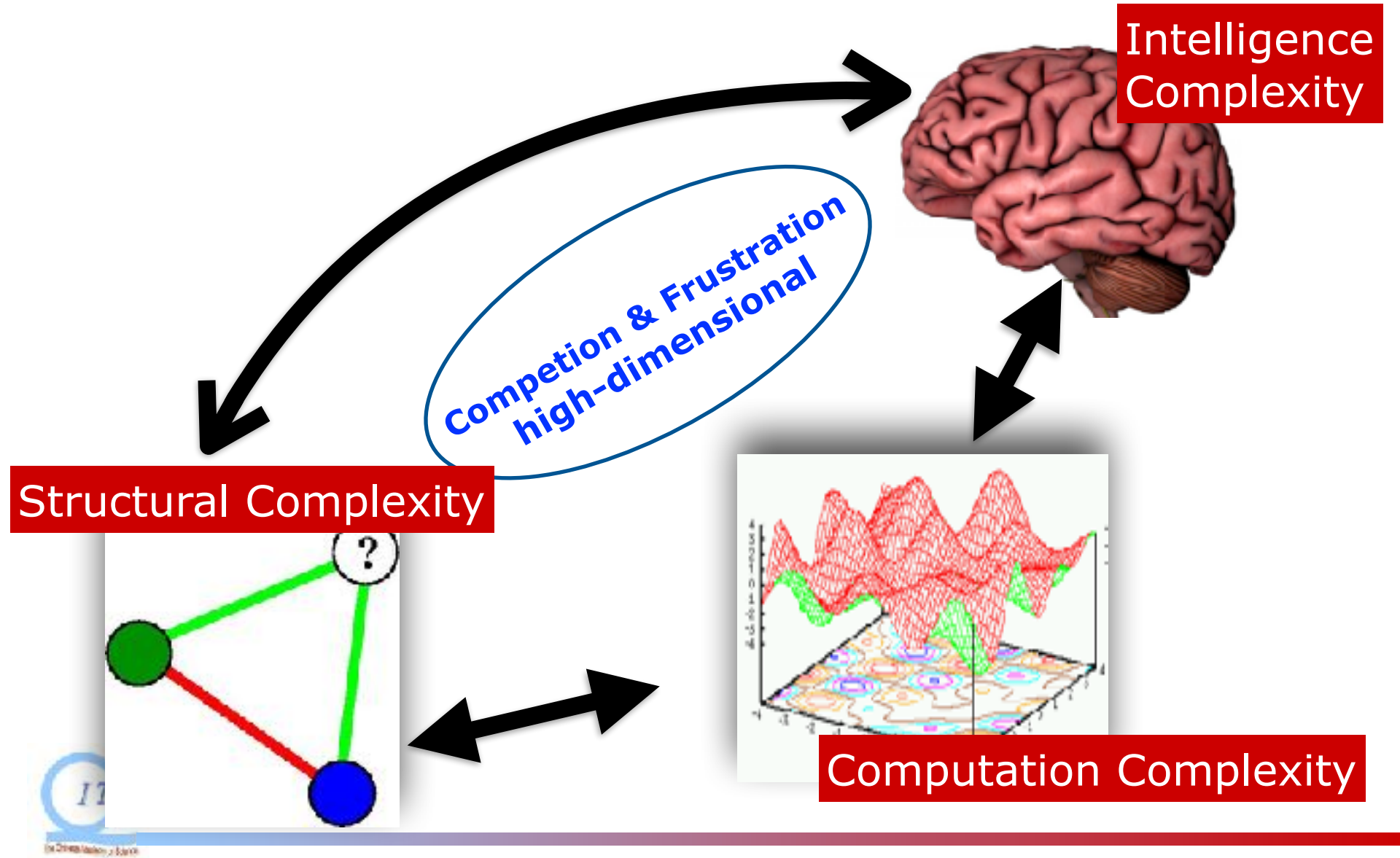


Stat.Mech. of Brain

Co-evolution of structure and dynamics



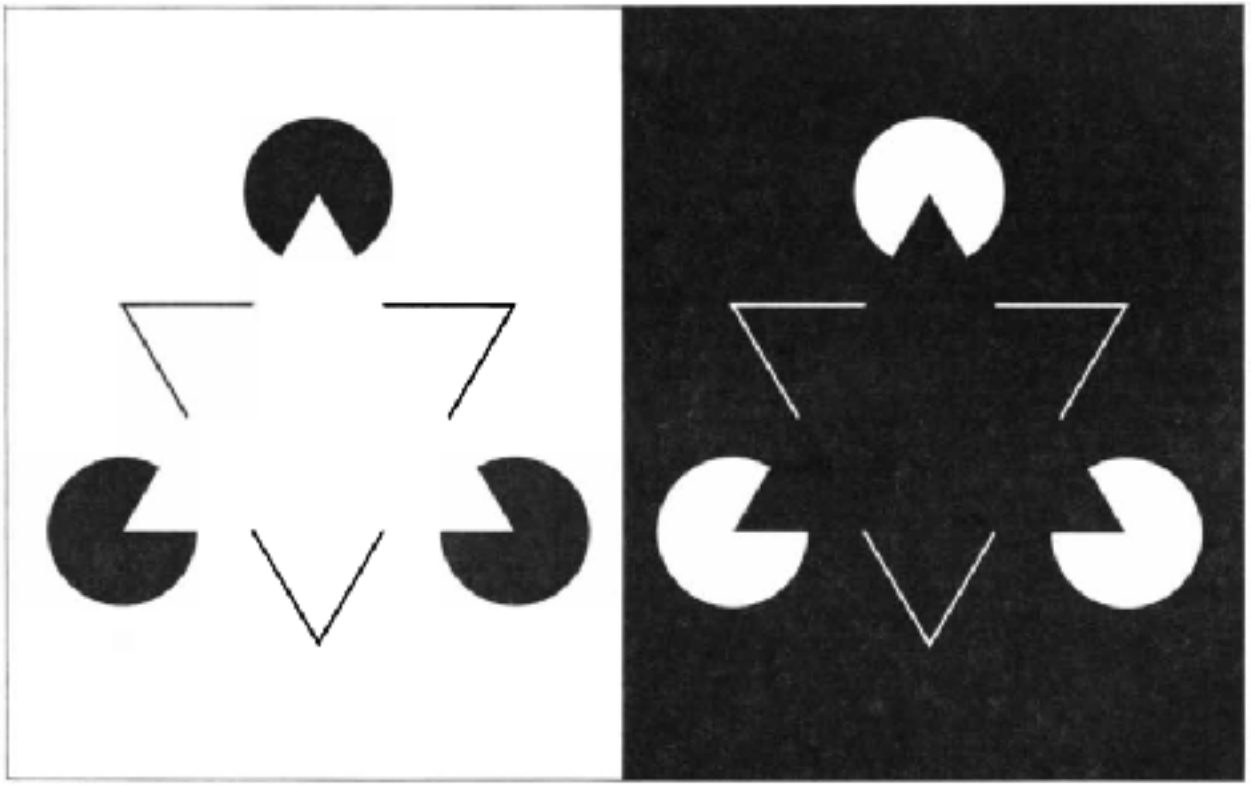
Complexity Frontier



Brain's powerful feature detection capability

human's visual system

G. Kanizsa, "Subjective Contours" (1976)



Vision is guided by internal model

… and feature separation capability

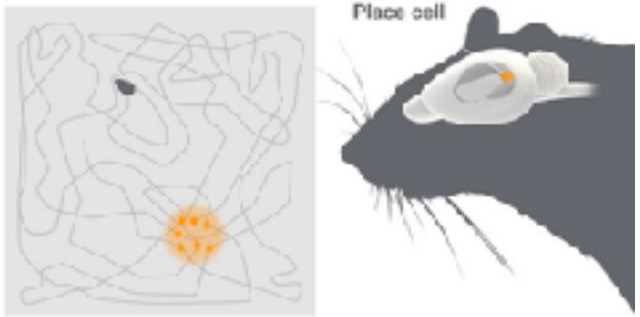
dog's olfactory system



human's auditorial system



place cells, concept cells

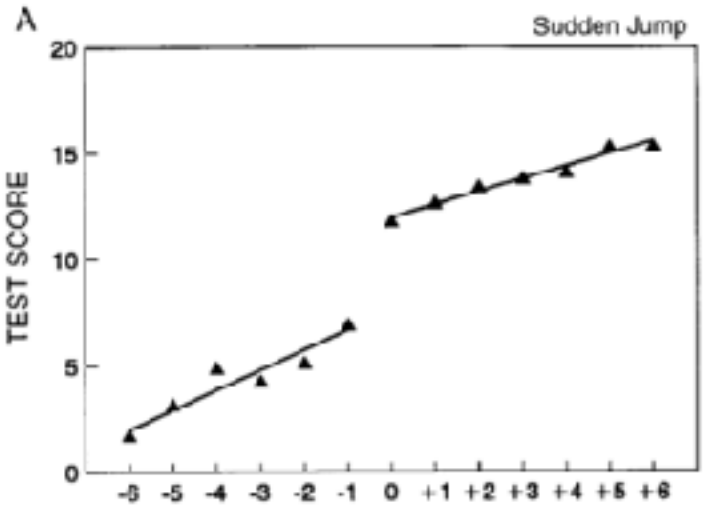


Decomposition into independent components

Discontinuity in building such capabilities

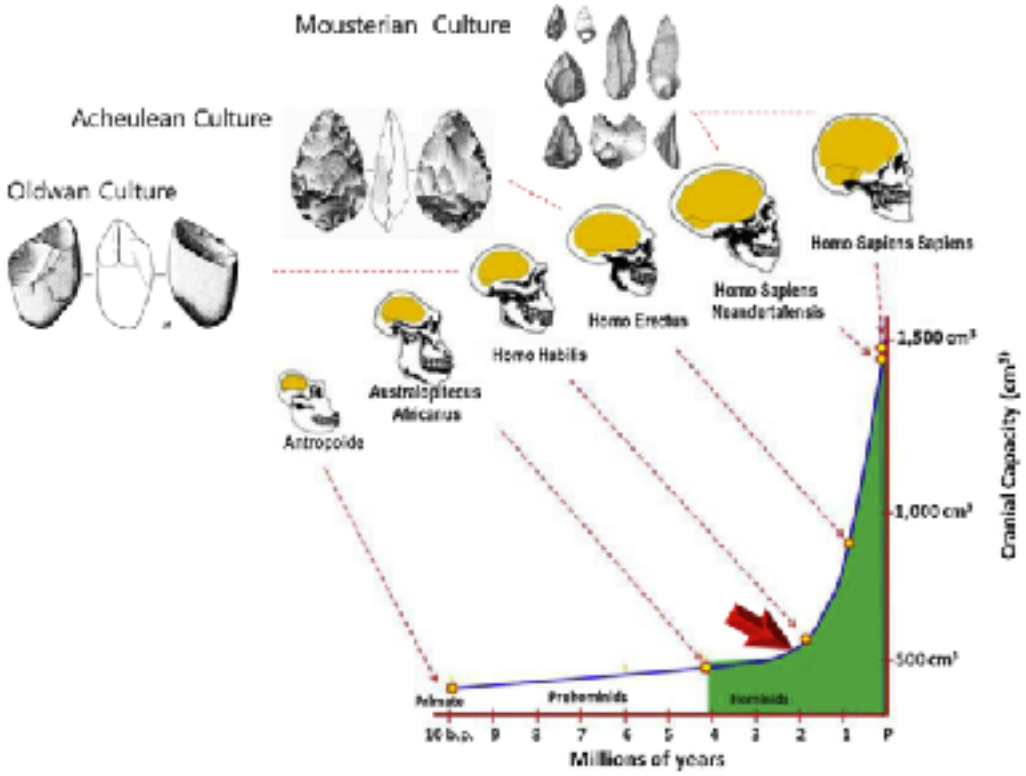
Short time scales

B. Hosenfeld et al., J. Experimental Child Psychology 64, 367-395 (1997)



Long time scales

R.V. Bretas et al., Neuroscience Research 161, 1-7 (2020)



Vol 440|20 April 2006|doi:10.1038/nature04701 nature

ARTICLES

Nature 2006

Weak pairwise correlations imply strongly correlated network states in a neural population

Elad Schneidman[1,2,3], Michael J. Berry II[2], Ronen Segev[2] & William Bialek[1,3]

Biological networks have so many possible states that exhaustive sampling is impossible. Successful analysis thus depends on simplifying hypotheses, but experiments on many systems hint that complicated, higher-order interactions among large groups of elements have an important role. Here we show, in the vertebrate retina, that weak correlations between pairs of neurons coexist with strongly collective behaviour in the responses of ten or more neurons. We find that this collective behaviour is described quantitatively by models that capture the observed pairwise correlations but assume no higher-order interactions. These maximum entropy models are equivalent to Ising models, and predict that larger networks are completely dominated by correlation effects. This suggests that the neural code has associative or error-correcting properties, and we provide preliminary evidence for such behaviour. As a first test for the generality of these ideas, we show that similar results are obtained from networks of cultured cortical neurons.

Chaos in Neuronal Networks with Balanced Excitatory and Inhibitory Activity

C. van Vreeswijk and H. Sompolinsky

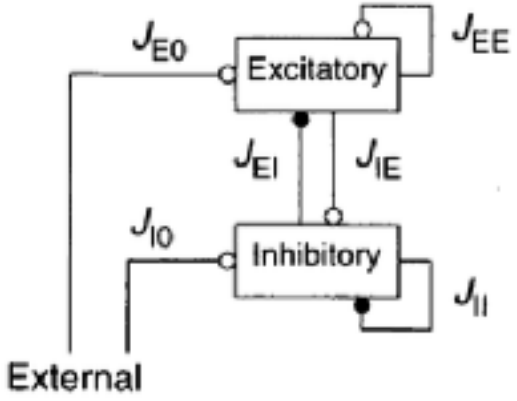
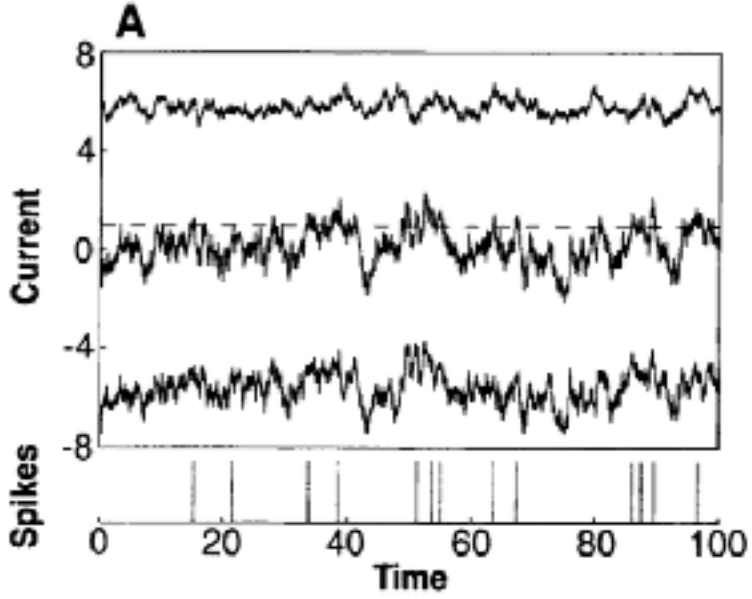


Fig. 1. Schematic representation of the network architecture. Excitatory connections between populations are shown as open circles; inhibitory ones, as filled circles.



Outline

1. **Motivation**
2. **Problem settings**
3. Theoretical framework
4. Detecting one non-Gaussian feature
5. Detecting and separating two non-Gaussian features
6. Conclusion

Basic viewpoint

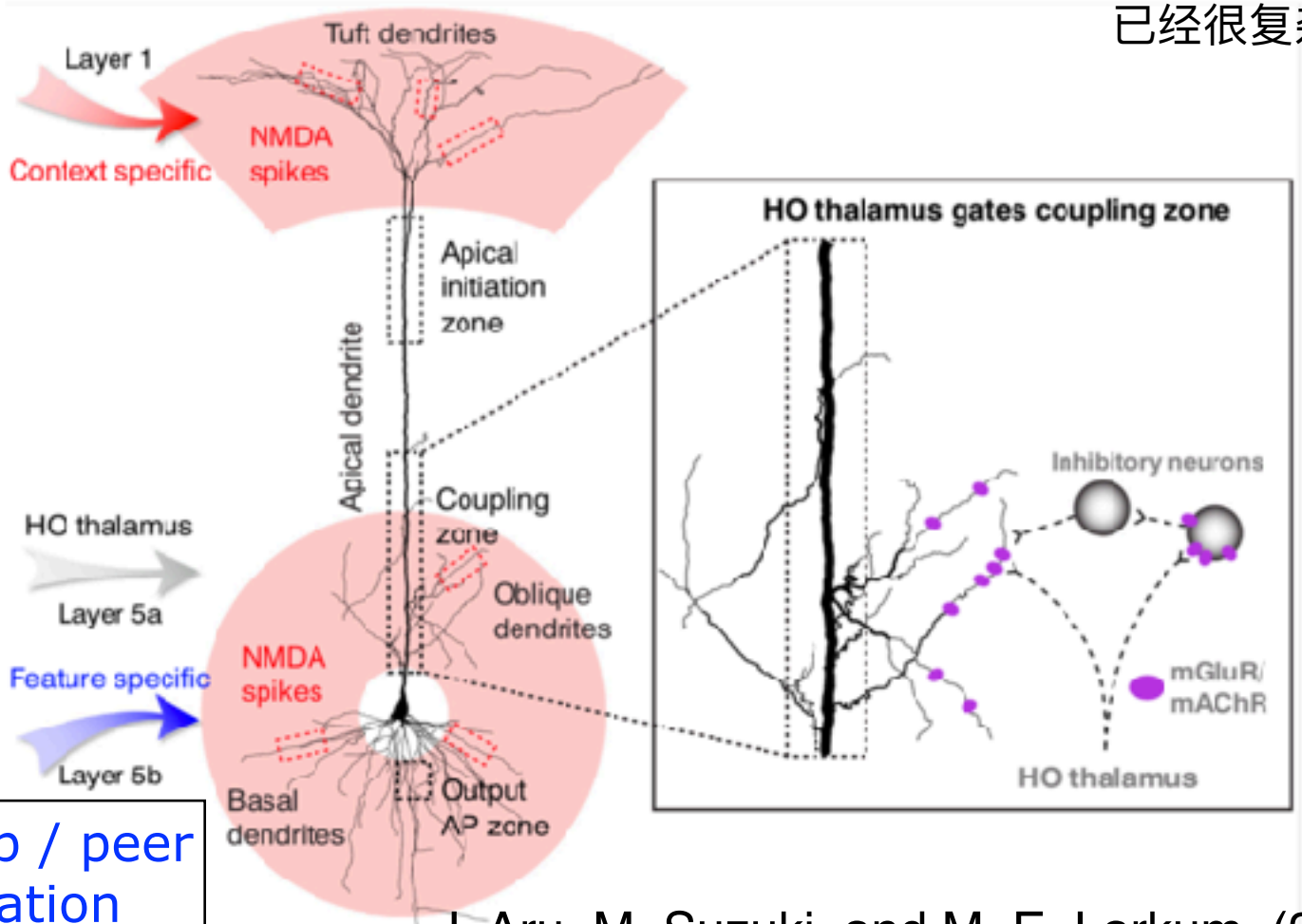
- Brain and Cell: **information processing** system
- Input signals: temporal and spatial **regularity**
- Internal **predictive model**
- **Energy — Information trade-off**

Single neuron

Information processing within a single neuron

已经很复杂了!

top down information



bottom up / peer information

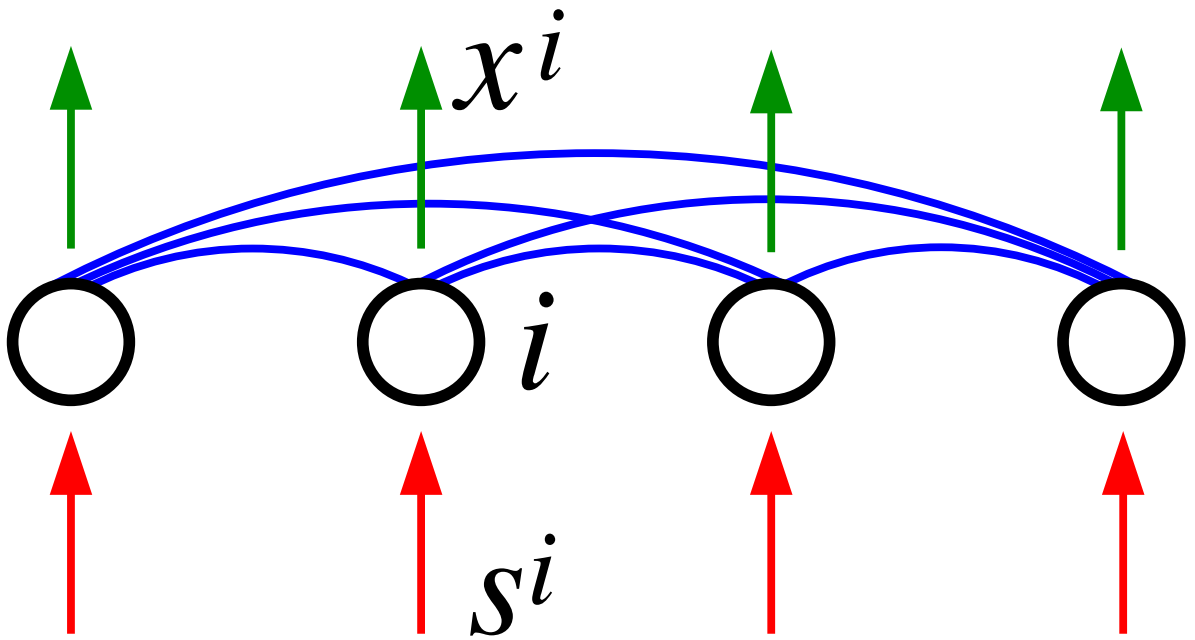
J. Aru, M. Suzuki, and M. E. Larkum (2020).

Model system

Lateral predictive coding, linear internal model, prediction error

Feature detection by lateral predictive coding

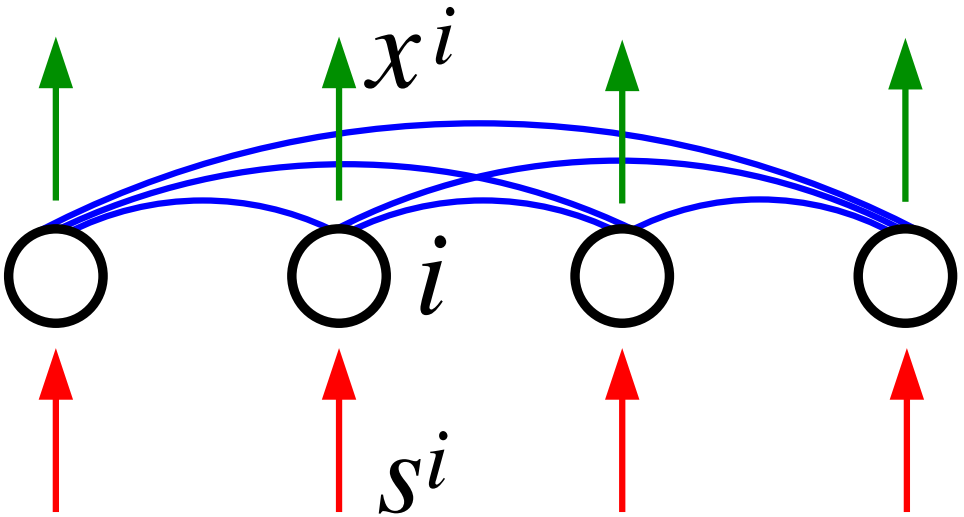
Output: $\boldsymbol{x} = (x^1, x^2, \ldots, x^N)^\top$



Lateral Interactions

Input: $\boldsymbol{s} = (s^1, s^2, \ldots, s^N)^\top$

lateral predictive coding



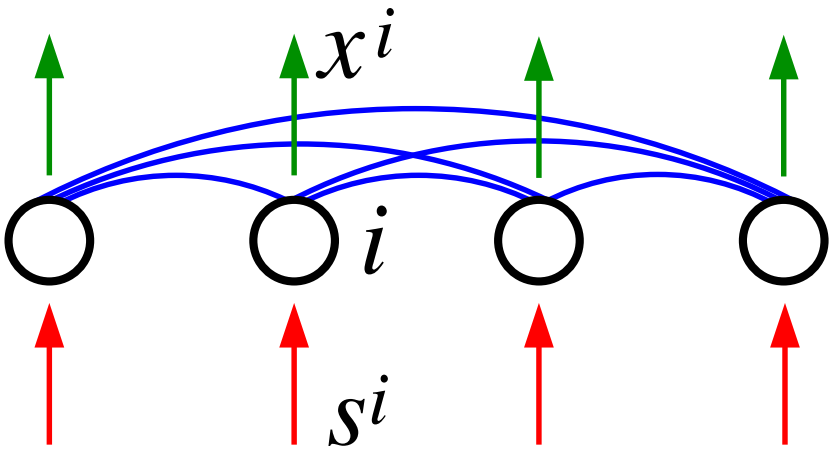
Short-time Dynamics:

$$\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} = \boldsymbol{s} - \boldsymbol{x} - \boldsymbol{f}(\boldsymbol{x})$$

Srinivasan, Laughlin, Dubs (1982);
Pineda, PRL (1987); ……

linear internal model

$$\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} = \boldsymbol{s} - \boldsymbol{x} - \boldsymbol{f}(\boldsymbol{x})$$



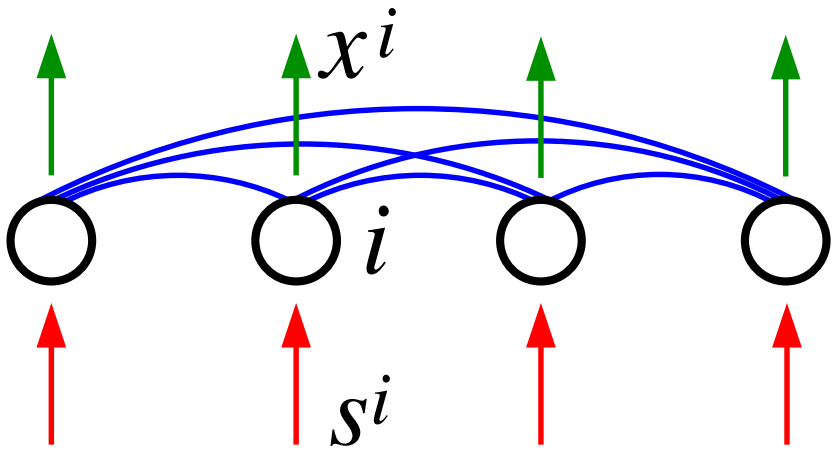
External signal

internal state

peer prediction

linear internal model

$$\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} = \boldsymbol{s} - \boldsymbol{x} - \boldsymbol{f}(\boldsymbol{x})$$



External signal

internal state

peer prediction

$$f_i(\boldsymbol{x}) = \sum_{j \neq i} w_{ij} x^j$$

linear predictive dynamics

weight matrix *W*

linear internal model

$$f_i(\boldsymbol{x}) = \sum_{j \neq i} w_{ij} x^j$$

internal state as prediction error

$$\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} = \mathrm{e}^{-(\boldsymbol{I}+\boldsymbol{W})t}\boldsymbol{s}$$

$$\boldsymbol{x}_\alpha = (\boldsymbol{I}+\boldsymbol{W})^{-1}\boldsymbol{s}_\alpha$$

$$\boldsymbol{W} = \begin{pmatrix} 0 & w_{12} & \cdots & w_{1N} \\ w_{21} & 0 & \cdots & w_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ w_{N1} & w_{N2} & \cdots & 0 \end{pmatrix}$$

$$\boldsymbol{s}_\alpha = \boldsymbol{W}\boldsymbol{x}_\alpha + \boldsymbol{x}_\alpha = \boldsymbol{p}_\alpha + \boldsymbol{x}_\alpha$$

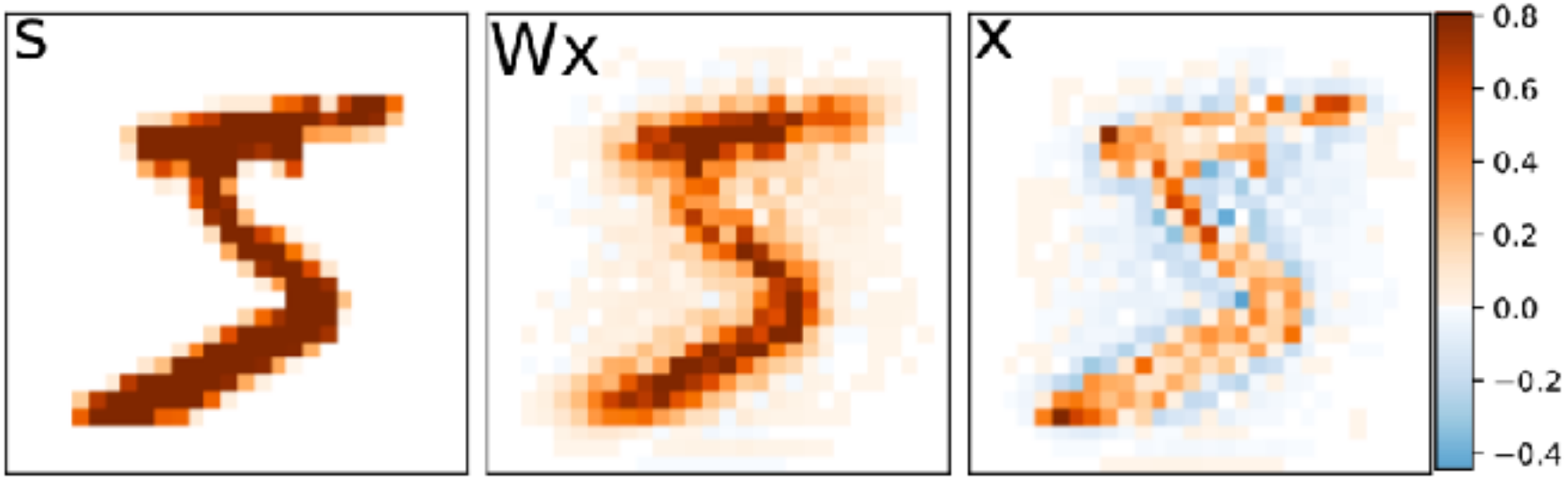
Input alpha

prediction for alpha

prediction-error for alpha

ITP

Huang, Fan, Zhou, Zhou, Commun. Theor. Phys. 74, 095601 (2022)



$$\boldsymbol{s}_\alpha = \boldsymbol{W}\boldsymbol{x}_\alpha + \boldsymbol{x}_\alpha = \boldsymbol{p}_\alpha + \boldsymbol{x}_\alpha$$

$(\boldsymbol{I} + \boldsymbol{W})$: eigenvalues have positive real part

Task 1:
Detecting a single non-Gaussian feature

$$\boldsymbol{s} = \underbrace{a_1}_{\text{non-Gaussian}} \boldsymbol{\phi}_1 + \sum_{j=2}^{N} \underbrace{b_k}_{\text{Gaussian}} \boldsymbol{\phi}_k$$

$$\langle a_i \rangle = 0\,, \quad \langle b_j \rangle = 0$$

$$\langle a_i^2 \rangle = 1\,, \quad \langle b_j^2 \rangle = 1$$

$\boldsymbol{\phi}_i,\ \boldsymbol{\phi}_j,\ \ldots$ orthogonal random bases

$\langle \boldsymbol{s}\boldsymbol{s}^\top \rangle = \boldsymbol{I}\,,$ Intractable by PCA

Task 2:
Detecting and separating two non-Gaussian features

$$\boldsymbol{s} = \underbrace{a_1}_{\text{non-Gaussian}} \left[\cos\frac{\theta}{2}\,\boldsymbol{\phi}_1 + \sin\frac{\theta}{2}\,\boldsymbol{\phi}_2\right] +$$

$$\underbrace{a_2}_{\text{non-Gaussian}} \left[\cos\frac{\theta}{2}\,\boldsymbol{\phi}_1 - \sin\frac{\theta}{2}\,\boldsymbol{\phi}_2\right]$$

$$+ \sum_{j=3}^{N} \underbrace{b_k}_{\text{Gaussian}}\ \boldsymbol{\phi}_k$$

$$
\begin{aligned}
\boldsymbol{s} \;=\; & \underbrace{a_1}_{\text{non-Gaussian}} \left[\cos\frac{\theta}{2}\,\boldsymbol{\phi}_1 + \sin\frac{\theta}{2}\,\boldsymbol{\phi}_2\right] + \\
& \underbrace{a_2}_{\text{non-Gaussian}} \left[\cos\frac{\theta}{2}\,\boldsymbol{\phi}_1 - \sin\frac{\theta}{2}\,\boldsymbol{\phi}_2\right] \\
& + \sum_{j=3}^{N} \underbrace{b_k}_{\text{Gaussian}}\; \boldsymbol{\phi}_k
\end{aligned}
$$

$$
\begin{aligned}
\hat{\boldsymbol{\phi}}_1 &= \cos\frac{\theta}{2}\boldsymbol{\phi}_1 + \sin\frac{\theta}{2}\boldsymbol{\phi}_2 \,, \\
\hat{\boldsymbol{\phi}}_2 &= \cos\frac{\theta}{2}\boldsymbol{\phi}_1 - \sin\frac{\theta}{2}\boldsymbol{\phi}_2 \,.
\end{aligned}
$$

$$
\begin{aligned}
\theta = \frac{\pi}{2} &: \quad \text{orthogonal} \,, \\
\theta = \frac{\pi}{4} &: \quad \text{partially aligned} \,.
\end{aligned}
$$

Outline

1. **Motivation**
2. **Problem settings**
3. **Theoretical framework**
4. Detecting one non-Gaussian feature
5. Detecting and separating two non-Gaussian features
6. Conclusion

Mechanism

Energy—information trade-off

Theory L1-norm Energy

$$\boldsymbol{x} = \frac{\boldsymbol{I}}{\boldsymbol{I}+\boldsymbol{W}}\boldsymbol{s}\,, \qquad x_i = \left[\frac{\boldsymbol{I}}{\boldsymbol{I}+\boldsymbol{W}}\ \boldsymbol{s}\right]_i$$

$$E = \sum_{i=1}^{N}\langle|x_i|\rangle$$

$$= \sum_{i}^{N} P_{\text{input}}(\boldsymbol{s})\left|\left[\frac{\boldsymbol{I}}{\boldsymbol{I}+\boldsymbol{W}}\ \boldsymbol{s}\right]_i\right|$$

Minimize energy of information transmission

Theory Information robustness

Mutual information ***x*** and ***s***:

$$\begin{aligned} I[\boldsymbol{x};\boldsymbol{s}] &= H[\boldsymbol{s}] + H[\boldsymbol{x}] - H[\boldsymbol{s},\boldsymbol{x}] \\ &= H[\boldsymbol{x}] - H[\boldsymbol{x}|\boldsymbol{s}] \\ &= H[\boldsymbol{x}] + constant \end{aligned}$$

Marginal distribution **x**:

$$P(\boldsymbol{x}) = \int \mathrm{d}\boldsymbol{s} P(\boldsymbol{s}) \delta\left(\boldsymbol{x} - \frac{\boldsymbol{I}}{\boldsymbol{I}+\boldsymbol{W}}\boldsymbol{s}\right)$$

$$H[\boldsymbol{x}] = -\log\left[\mathrm{Det}(\boldsymbol{I}+\boldsymbol{W})\right] + constant$$

$$S = -\log\left[\mathrm{Det}\left(\boldsymbol{I}+\boldsymbol{W}\right)\right]$$ Entropy

Maximize output entropy
(output points more widely dispersed)

Theory Energy-information tradeoff

Free energy:

$$
\begin{aligned}
F &= E - T\,S \\
&= \sum_i \left\langle \left[\frac{\boldsymbol{I}}{\boldsymbol{I}+\boldsymbol{W}}\ \boldsymbol{s} \right]_i \right\rangle + T \log\left[\mathrm{Det}\left(\boldsymbol{I}+\boldsymbol{W}\right) \right]
\end{aligned}
$$

Minimum free energy at temperature ***T***.
Optimal internal model (weight matrix ***W***).

***T* is balancing energetic cost and functional robustness.**

Algorithm Stochastic Annealing at fixed S

1. Change one column/row of $\boldsymbol{W}$ under constraint $\mathrm{Det}(\boldsymbol{I}+\boldsymbol{W}) = e^{-S}$.
2. Check eigenvalues λ of $(\boldsymbol{I}+\boldsymbol{W})$.
3. If $\mathrm{Re}[\lambda] > 0$, then go back to (1).
   Else, compute energy change ΔE.
4. If $\Delta E \leq 0$ or if $\mathrm{Rand}() < e^{-\kappa \Delta E}$, accept updated $\boldsymbol{W}$.
5. After $\mathcal{N} = 10^5$ repeats of (1)-(4), increase $\kappa \leftarrow 1.01\kappa$.
6. Stop if $\kappa > 10^8$. Report $\boldsymbol{W}$ of the minimum energy.

SCIENCE CHINA
Physics, Mechanics & Astronomy

• Article •

June 2024 Vol. 67 No. 6: 260511
https://doi.org/10.1007/s11433-024-2341-2

Reference 1

Energy-information trade-off induces continuous and discontinuous phase transitions in lateral predictive coding

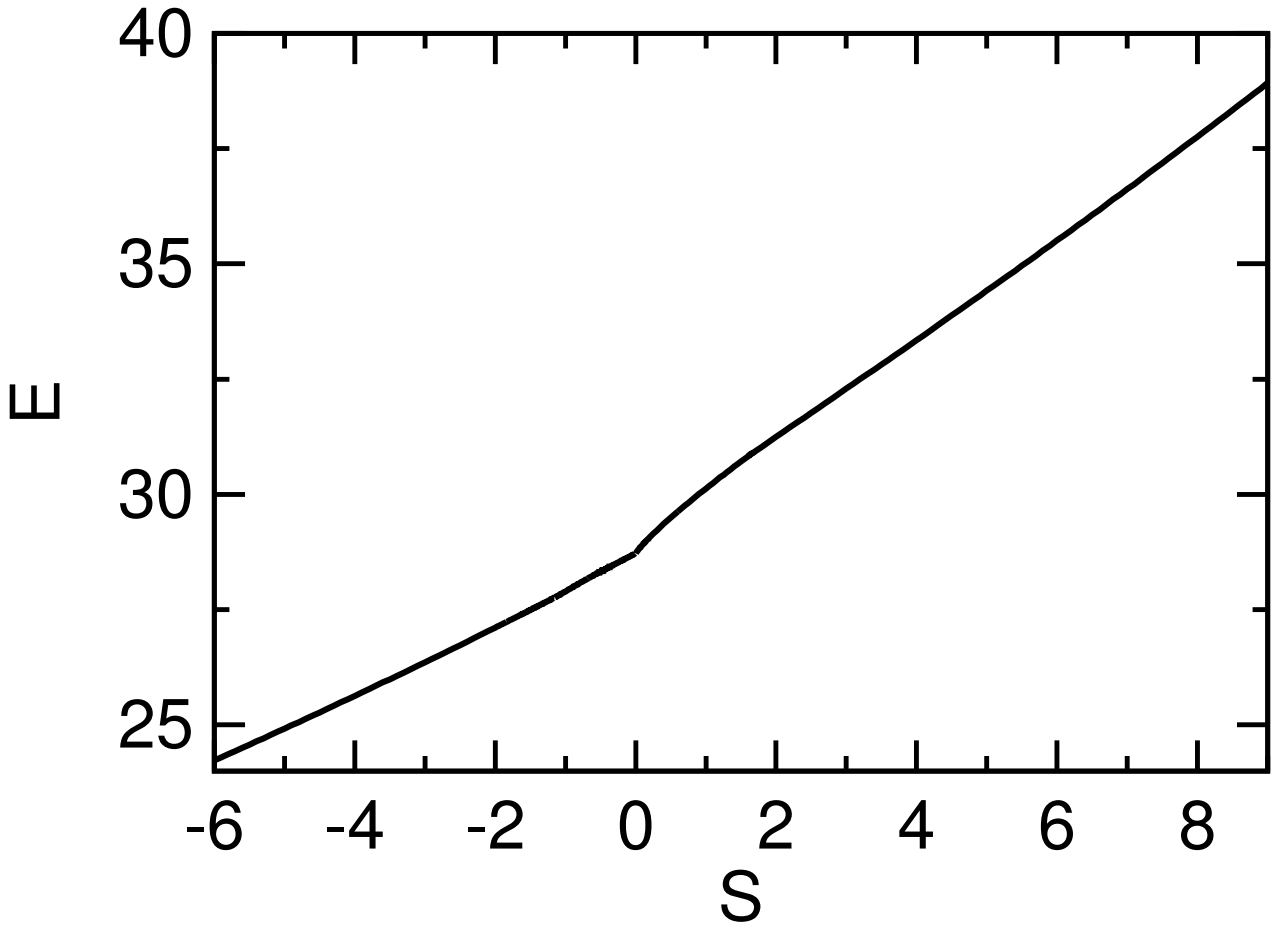
Zhen-Ye Huang[1,2†], Ruyi Zhou[3†], Miao Huang[1,2†], and Hai-Jun Zhou[1,2,4*]

[1]*Key Laboratory for Theoretical Physics, Institute of Theoretical Physics, Chinese Academy of Sciences, Beijing 100190, China;*
[2]*School of Physical Sciences, University of Chinese Academy of Sciences, Beijing 100049, China;*
[3]*School of Optics and Photonics, Beijing Institute of Technology, Beijing 100081, China;*
[4]*Minjiang Collaborative Center for Theoretical Physics, Minjiang University, Fuzhou 350108, China*

Received January 12, 2024; accepted February 4, 2024; published online May 7, 2024

Lateral predictive coding is a recurrent neural network that creates energy-efficient internal representations by exploiting statistical regularity in sensory inputs. Here, we analytically investigate the trade-off between information robustness and energy in a linear model of lateral predictive coding and numerically minimize a free energy quantity. We observed several phase transitions in the synaptic weight matrix, particularly a continuous transition that breaks reciprocity and permutation symmetry and builds cyclic dominance and a discontinuous transition with the associated sudden emergence of tight balance between excitatory and inhibitory interactions. The optimal network follows an ideal gas law over an extended temperature range and saturates the efficiency upper bound of energy use. These results provide theoretical insights into the emergence and evolution of complex internal models in predictive processing systems.

Entropy versus Energy



If E(S) is not smooth, or changes convexity, phase transition!

Obtain minimum energy E for each entropy S.

Outline

1. **Motivation**
2. **Problem settings**
3. **Theoretical framework**
4. **Detecting one non-Gaussian feature**
5. Detecting and separating two non-Gaussian features
6. Conclusion

Task 1:
Detecting a single non-Gaussian feature

$$\boldsymbol{s} = \underbrace{a_1}_{\text{non-Gaussian}} \boldsymbol{\phi}_1 + \sum_{j=2}^{N} \underbrace{b_k}_{\text{Gaussian}} \boldsymbol{\phi}_k$$

$$\langle a_i \rangle = 0\,, \quad \langle b_j \rangle = 0$$

$$\langle a_i^2 \rangle = 1\,, \quad \langle b_j^2 \rangle = 1$$

$\boldsymbol{\phi}_i,\ \boldsymbol{\phi}_j,\ \ldots$ orthogonal random bases

$\langle \boldsymbol{s}\boldsymbol{s}^\top \rangle = \boldsymbol{I}\,,$ Intractable by PCA

Probability of non-Gaussian coefficient

$$q(a_1) = \begin{cases} p_0 \,, & a_1 = 0 \\ \frac{1-p_0}{2} \,, & a_1 = \frac{1}{\sqrt{1-p_0}} \\ \frac{1-p_0}{2} \,, & a_1 = -\frac{1}{\sqrt{1-p_0}} \end{cases}$$

Order parameter

Projection of non-Gaussian feature direction on unit ***i***:

$$\mu_i = \left[\frac{\boldsymbol{I}}{\boldsymbol{I}+\boldsymbol{W}} \, \boldsymbol{\phi}_1 \right]_i$$

$$Q = \max_i \frac{|\mu_i|}{\sqrt{\sum_j \mu_j^2}}$$

$Q \approx 1$: selective response by a single unit

Discontinuous phase transitions of feature detection in lateral predictive coding

Zhen-Ye Huang,[1,2,*] Weikang Wang,[1,†] and Hai-Jun Zhou[1,2,3,‡]

[1]*Institute of Theoretical Physics, Chinese Academy of Sciences, Beijing 100190, China*

[2]*School of Physical Sciences, University of Chinese Academy of Sciences, Beijing 100049, China*

[3]*MinJiang Collaborative Center for Theoretical Physics, MinJiang University, Fuzhou 350108, China*

(Received 28 January 2025; accepted 19 August 2025; published 2 September 2025)

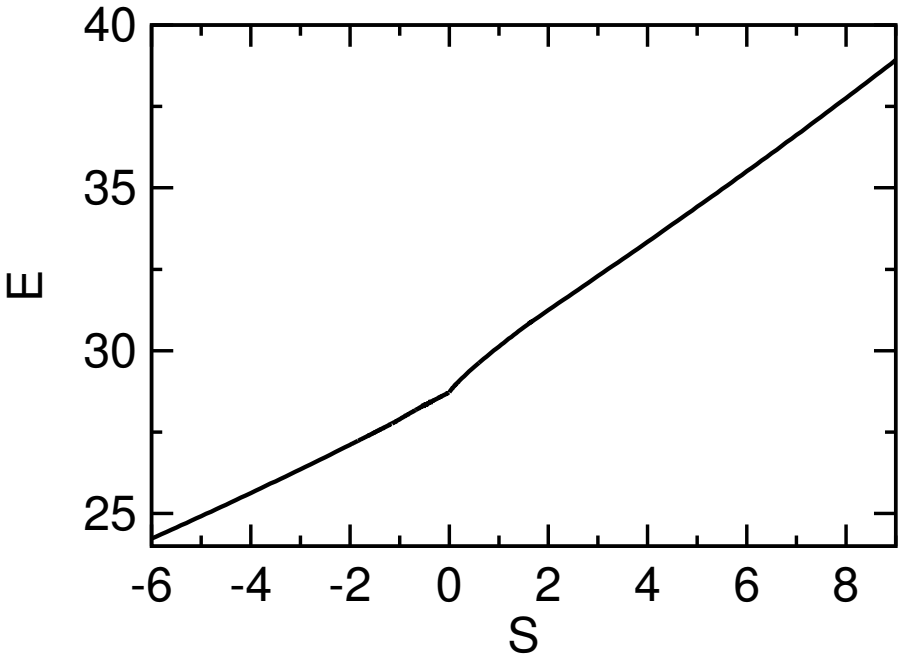
The brain may adopt the strategy of lateral predictive coding (LPC) to construct optimal internal representations for salient features in input sensory signals, reducing the energetic cost of information transmission. Here we first consider the task of detecting one non-Gaussian signal by LPC from Gaussian background signals of the same magnitude, which is intractable by principal component decomposition. We study the emergence of the feature detection function from the perspective of tradeoff between energetic cost E and information robustness, and implement this tradeoff by a thermodynamic free energy. We define E as the mean L_1-norm of the internal state vectors, and quantify the level of information robustness by an entropy measure S. There are at least three types of optimal LPC matrices, one type with very weak synaptic weights and $S \approx 0$, and two functional types either with low energy E or with high entropy S in which one single unit selectively responds to the non-Gaussian input feature. The energy-information tradeoff induces two discontinuous phase transitions between these three types of optimal LPC networks. We then extend the discussion to detecting and distinguishing between two non-Gaussian input features and observe similar discontinuous phase transitions.

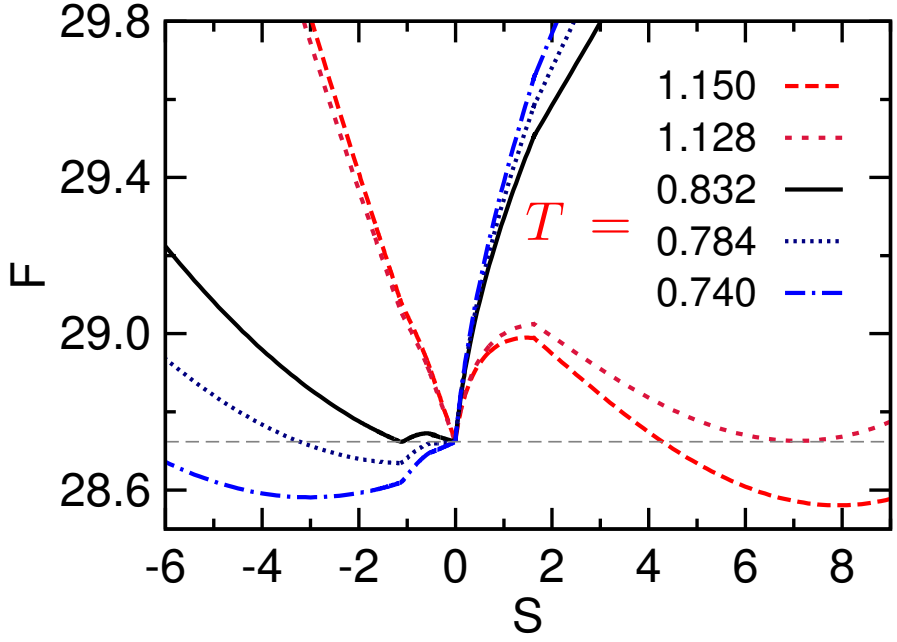
DOI: 10.1103/3jk3-x177

Representative results

$$N = 36\,, \quad p_0 = 0.7$$

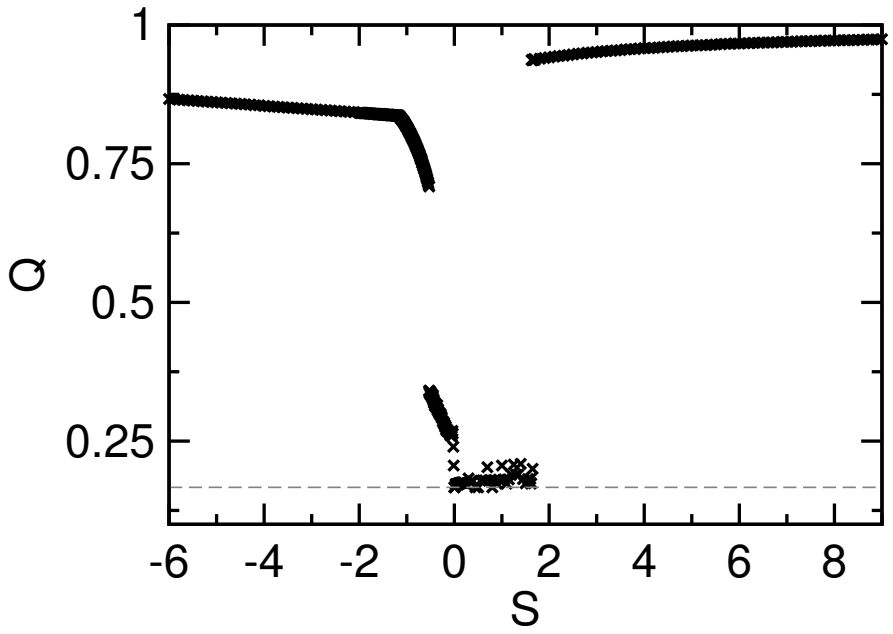
Energy and free energy





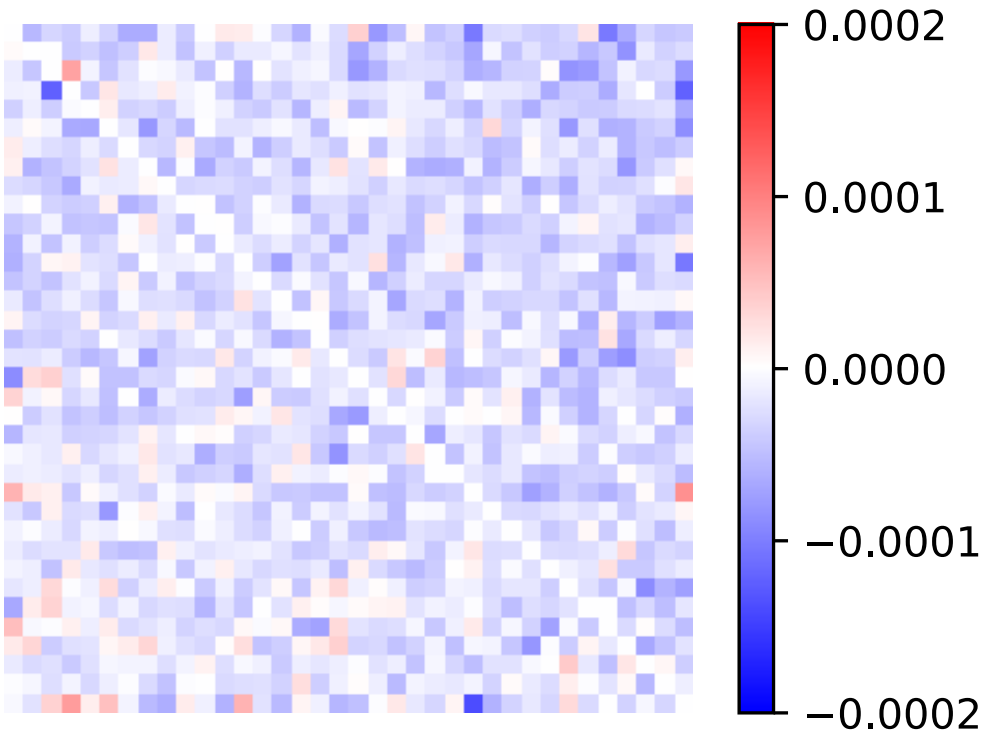
Free energy global minimum changes suddenly at certain critical *T* values.

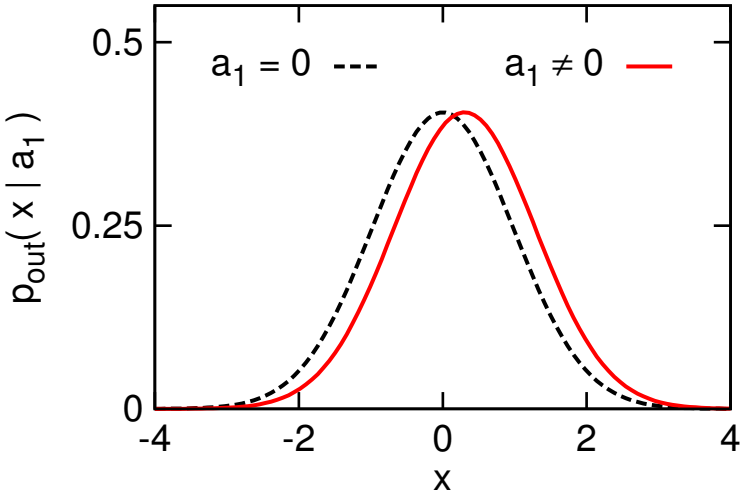
Feature detection power



Feature detection is successful both for
(1) S much larger than 0 (S > 2);
(2) S much lower than 0 (S < -1).

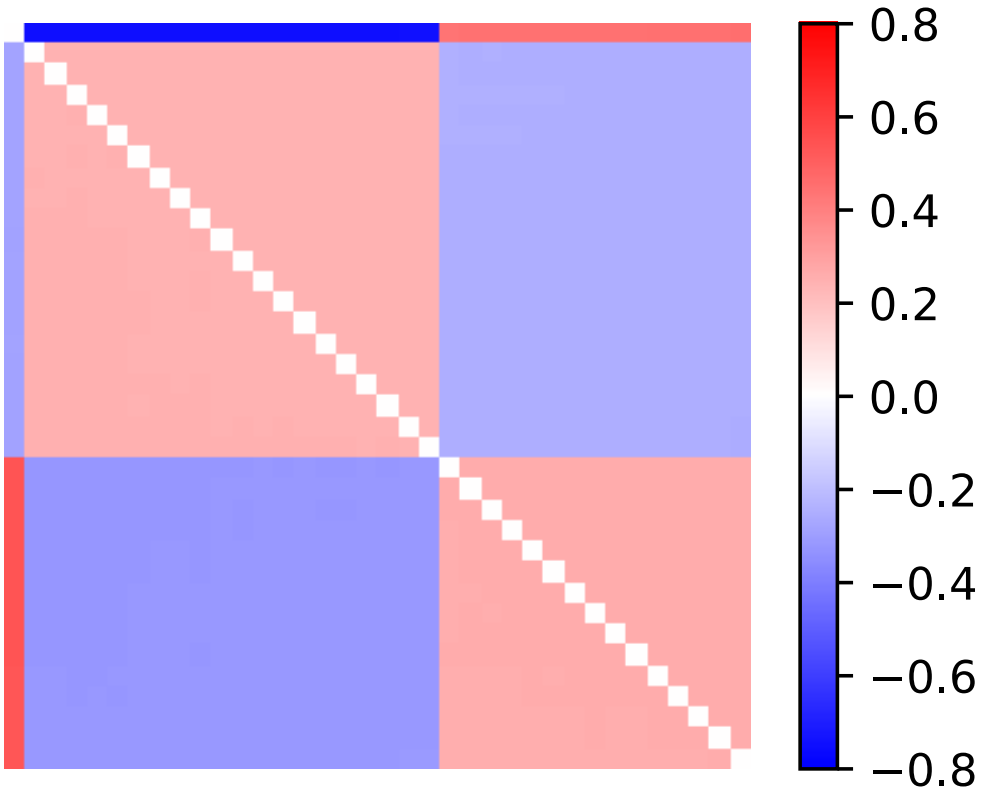
Weight matrix with *S=0*

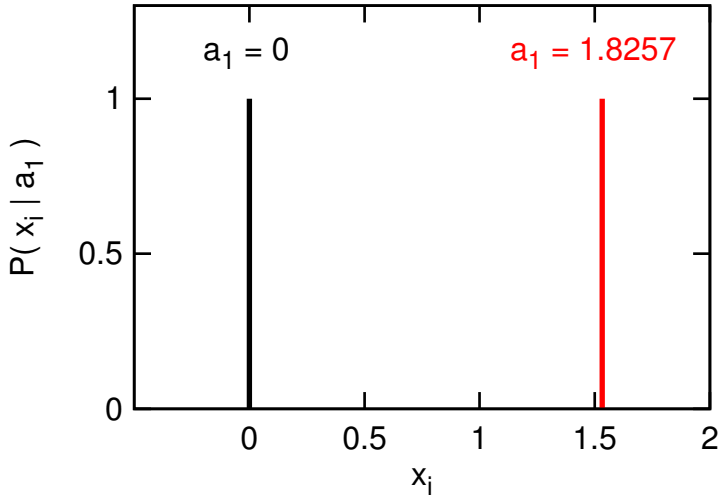




All synaptic weights are weak.
Similar response in the presence /absence of non-Gaussian feature.

Weight matrix with $S=8$

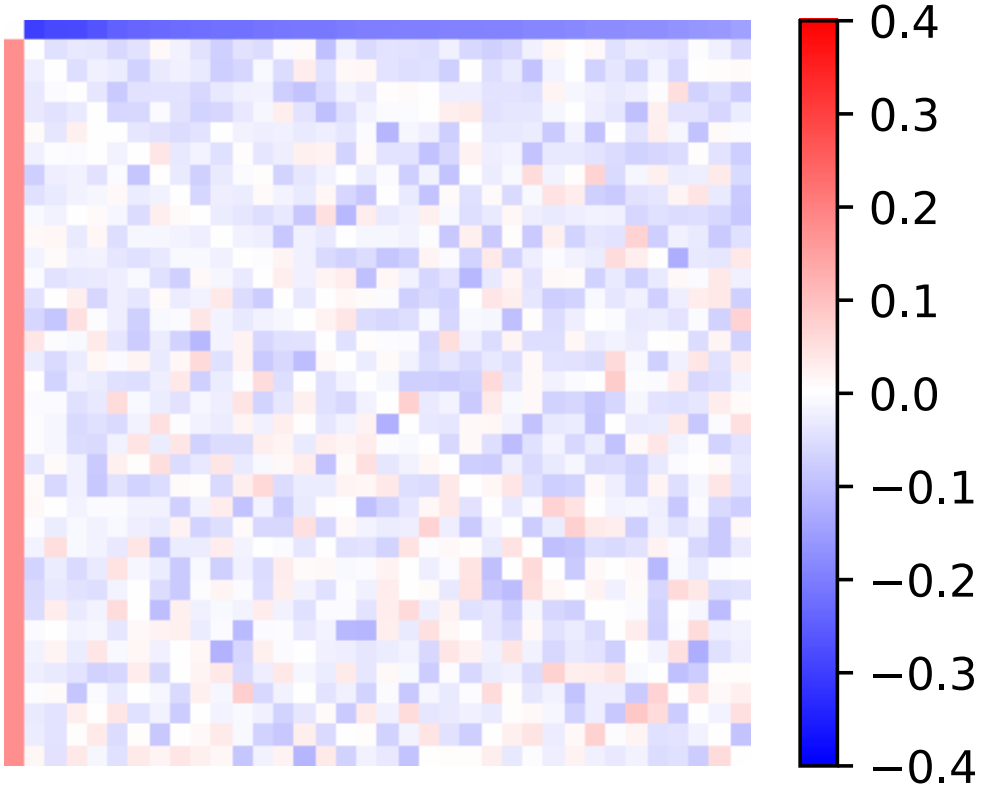


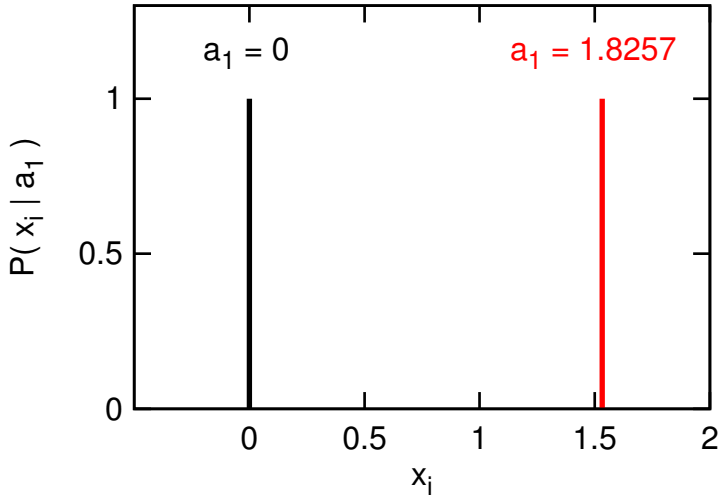


Three subgroups of units, strong excitation and strong inhibition.

One unit selective to non-Gaussian feature.

Weight matrix with $S=-2$



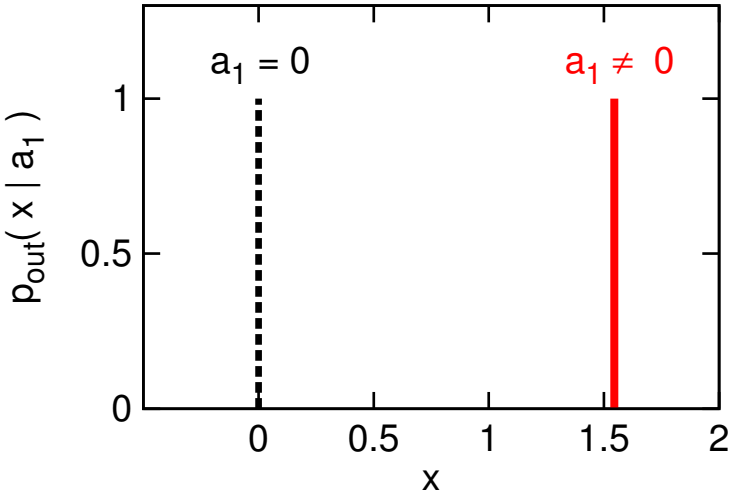


One unit strongly interact with all the others.
This unit selective to non-Gaussian feature.

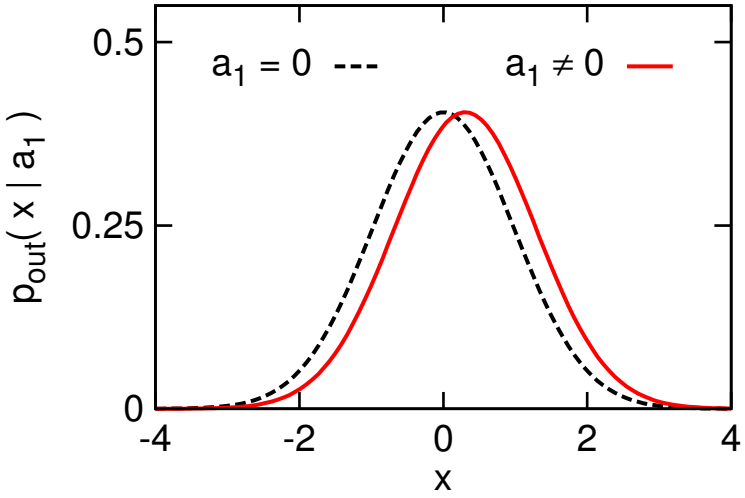
Through energy-information trade-off, the system achieves feature detection function.

A single neuron is selectively responding to the non-Gaussian feature.

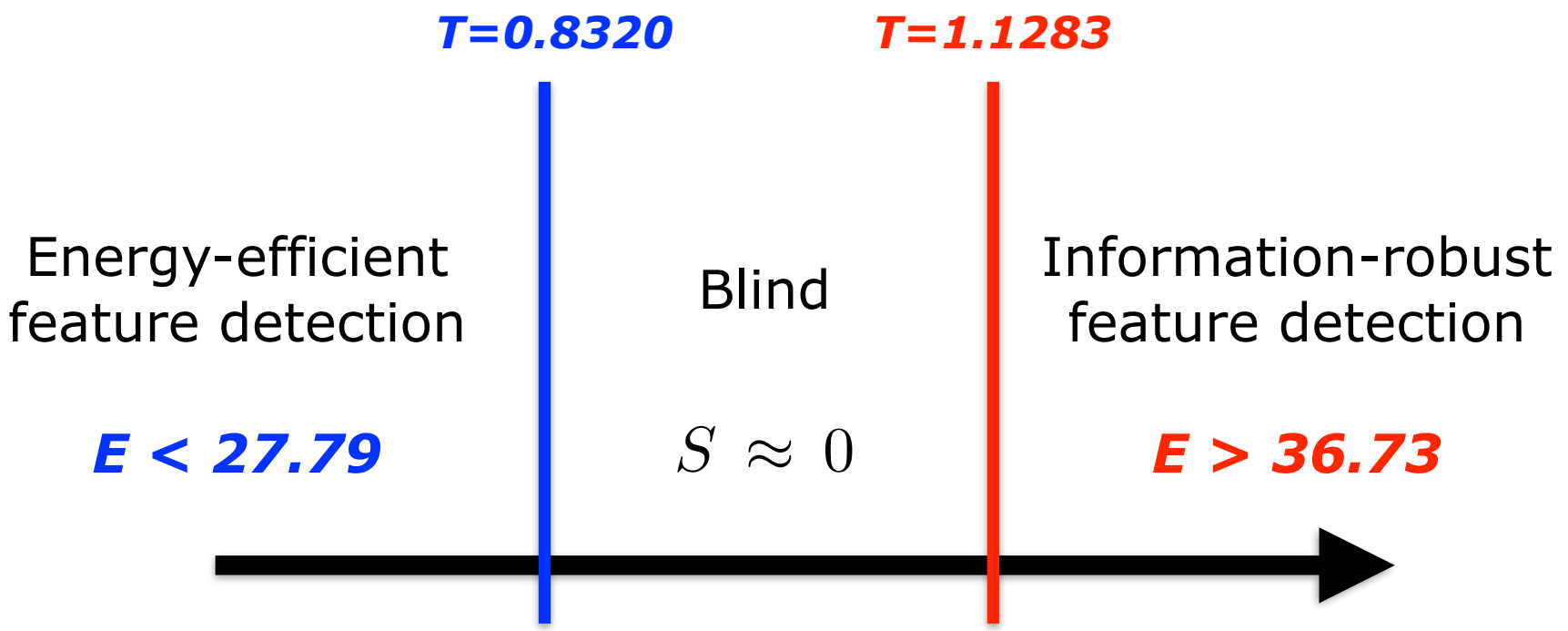
one neuron



all other neurons



Phase diagram in *T*



Outline

1. **Motivation**
2. **Problem settings**
3. **Theoretical framework**
4. **Detecting one non-Gaussian feature**
5. **Detecting and separating two non-Gaussian features**
6. Conclusion

$$\boldsymbol{s} = \underbrace{a_1}_{\text{non-Gaussian}} \left[\cos\frac{\theta}{2}\,\boldsymbol{\phi}_1 + \sin\frac{\theta}{2}\,\boldsymbol{\phi}_2\right] + \underbrace{a_2}_{\text{non-Gaussian}} \left[\cos\frac{\theta}{2}\,\boldsymbol{\phi}_1 - \sin\frac{\theta}{2}\,\boldsymbol{\phi}_2\right] + \sum_{j=3}^{N} \underbrace{b_k}_{\text{Gaussian}}\,\boldsymbol{\phi}_k$$

$$\hat{\boldsymbol{\phi}}_1 = \cos\frac{\theta}{2}\boldsymbol{\phi}_1 + \sin\frac{\theta}{2}\boldsymbol{\phi}_2 ,$$
$$\hat{\boldsymbol{\phi}}_2 = \cos\frac{\theta}{2}\boldsymbol{\phi}_1 - \sin\frac{\theta}{2}\boldsymbol{\phi}_2 .$$

$$\theta = \frac{\pi}{2} : \quad \text{orthogonal} ,$$
$$\theta = \frac{\pi}{4} : \quad \text{partially aligned} .$$

Order parameters

Projection of non-Gaussian feature directions on unit ***i***:

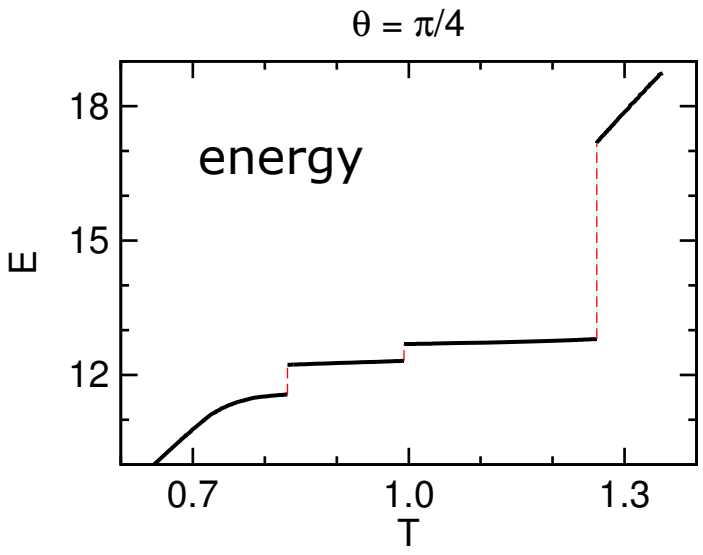
$$\hat{\mu}_i^{(1)} = \left[\frac{\boldsymbol{I}}{\boldsymbol{I}+\boldsymbol{W}}\,\hat{\boldsymbol{\phi}}_1\right]_i \qquad \hat{\mu}_i^{(2)} = \left[\frac{\boldsymbol{I}}{\boldsymbol{I}+\boldsymbol{W}}\,\hat{\boldsymbol{\phi}}_2\right]_i$$

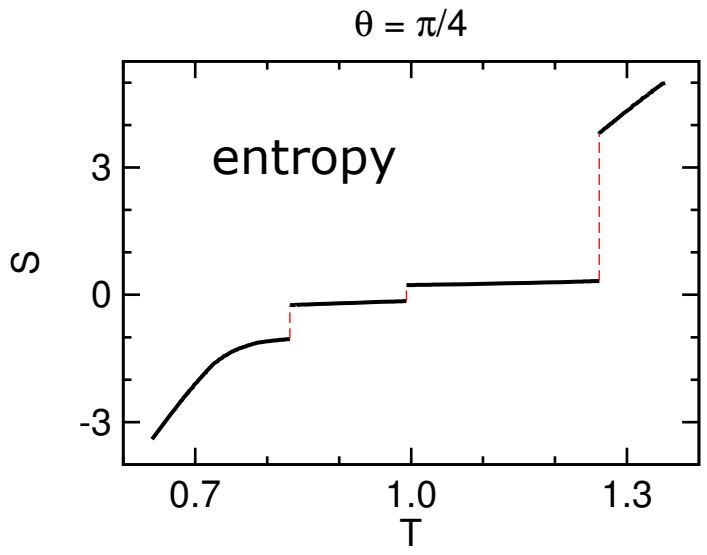
$$Q^{(1)} = \max_i \frac{|\hat{\mu}_i^{(1)}|}{\sum_j \left(\hat{\mu}_j^{(1)}\right)^2} \qquad Q^{(2)} = \max_i \frac{|\hat{\mu}_i^{(2)}|}{\sum_j \left(\hat{\mu}_j^{(2)}\right)^2}$$

$Q \approx 1$: selective response by a single unit

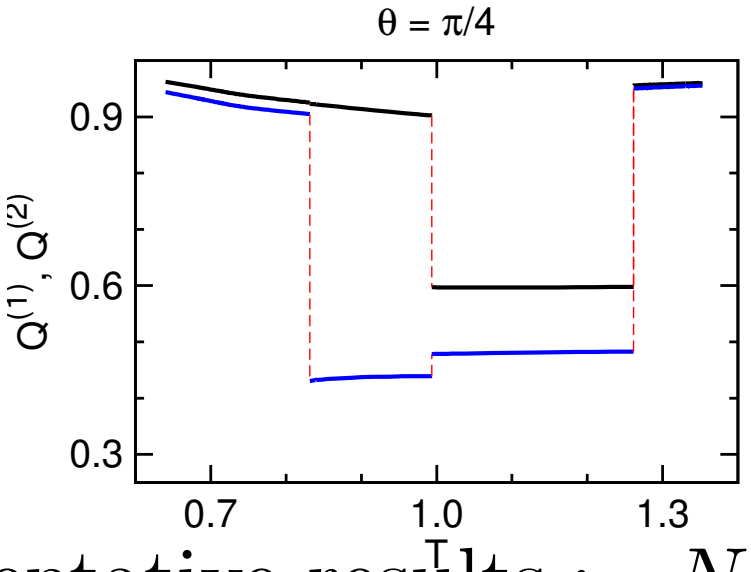
Orthogonal features

$$\theta = \frac{\pi}{2}$$





order parameter

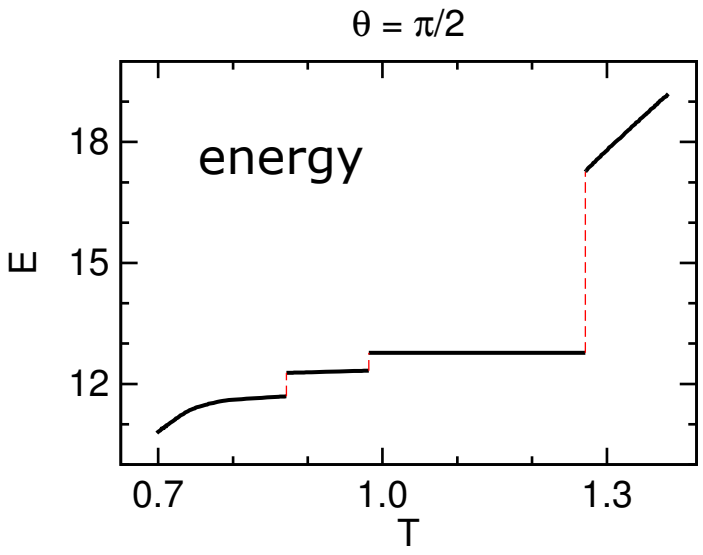


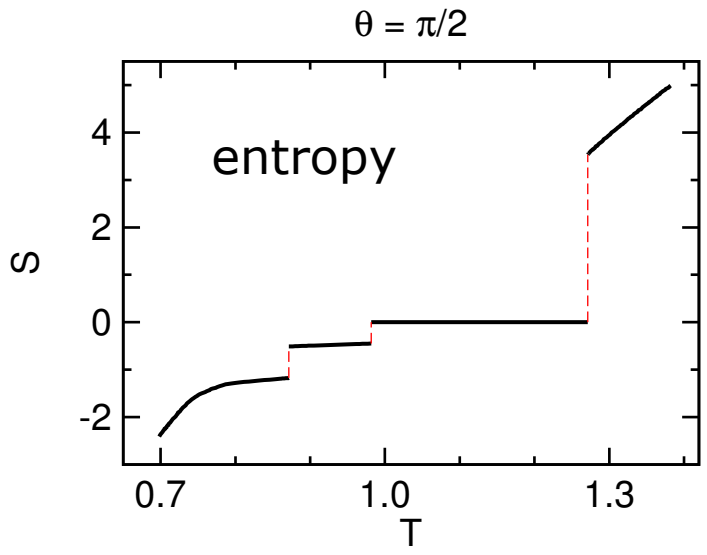
Again four phases, with different detection capability!

Representative results : $N = 16,\ p_0 = 0.6$

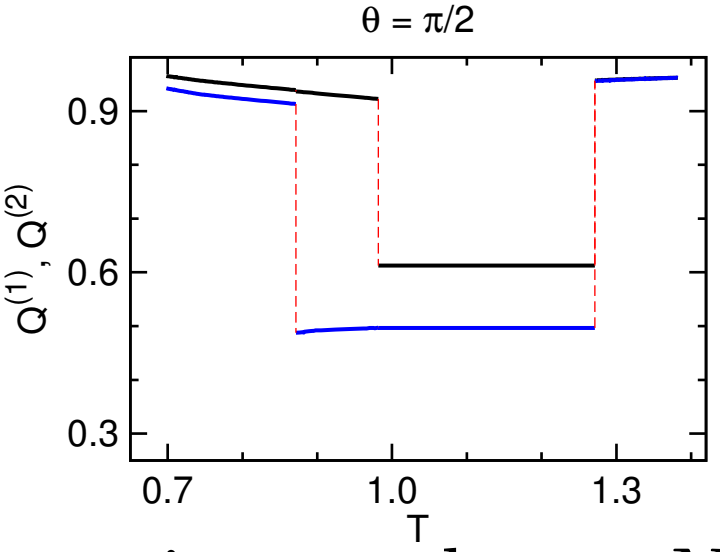
Non-orthogonal features

$$\theta = \frac{\pi}{4}$$





order parameter



Four phases, with different detection capability!

Representative results : $N = 16,\ p_0 = 0.6$

$$\theta = \frac{\pi}{4}$$

Phase diagram in *T*

T=0.8316 **T=0.9935** **T=1.2612**

Energy-efficient feature separation

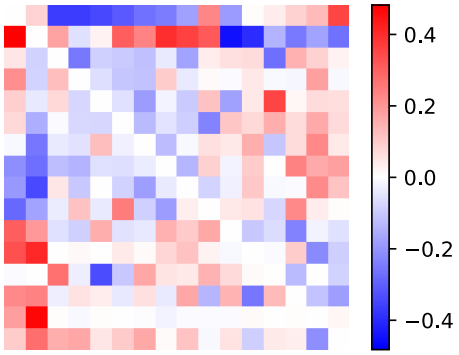
E < 11.57

One feature detected

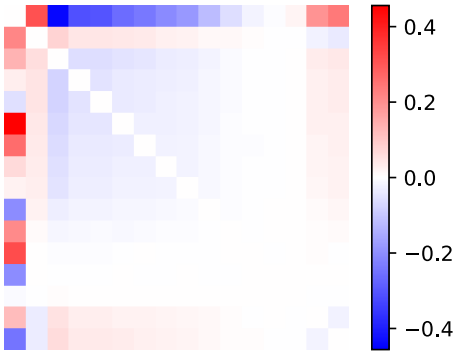
Blind

Information-robust feature separation

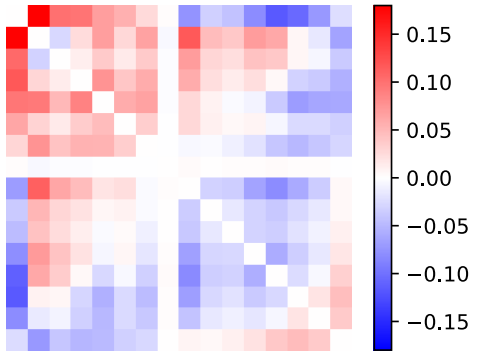
E > 17.17



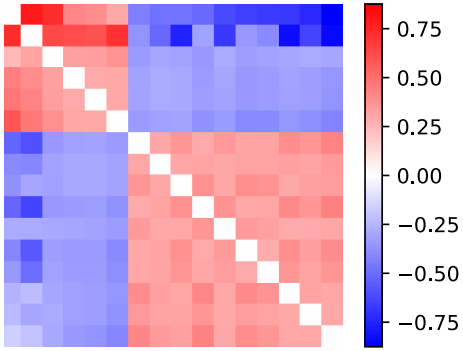
S=-2



S=-0.2



S=0.3



S=4

Outline

1. **Motivation**
2. **Problem settings**
3. **Theoretical framework**
4. **Detecting one non-Gaussian feature**
5. **Detecting and separating two non-Gaussian features**
6. **Conclusion**

Conclusive remarks

- Lateral, recurrent, interactions for predictive coding.
- Prediction-error minimization —> redundancy reduction, correlation reduction, differentiated response times.
- Energy-Information trade-off leads to highly nontrivial internal models, **spontaneous emergence of feature detection, extraction, and separation functions.**
- Asymmetry of synaptic weights, cyclic dominance, inhibition-excitation balance.
- Future work: Nonlinearity, multi-layers, noise, …

Future extensions

- Memory effect.

$$\tau_0 \frac{\mathrm{d}x_i(t)}{\mathrm{d}t} = s_i(t) - x_i(t) - \sum_{j \neq i} w_{ij} \; f[\{x_j\}]$$

$$f[\{x_j\}] = \int_{-\infty}^{t} \underbrace{K(t-t')}_{\text{memory kernel}} \; x_j(t') \; \mathrm{d}t'$$

- Non-linearity.

$$\tau_0 \frac{\mathrm{d}x_i(t)}{\mathrm{d}t} = s_i(t) - x_i(t) - \sum_{j \neq i} w_{ij} \; f_j$$

$$f_j = \min\big[\, 0,\, x_j(t) \,\big] \qquad \text{(ReLu)}$$

- Multiple layers.

Acknowledgements

Co-workers:
Zhen-Ye Huang (黄振业)
Weikang Wang (王维康)
Xin-Yi Fan (樊欣祎)
Jianwen Zhou (周健文)
Miao Huang (黄淼)
Ruyi Zhou (周如意)

Funds:
NSFC (国家自然科学基金委员会)
CAS (中国科学院先导培育项目)
JKW （国防科技前沿创新计划开放课题）