Revealing unexpected complex encoding but simple decoding mechanisms in motor cortex via separating behaviorally relevant neural signals

Yangang Li, Xinyun Zhu, Yu Qi*, Yueming Wang*

ZHEJIANG UNIVERSITY

1 INTRODUCTION

Goal: Understanding how motor cortex encodes and decodes behavioral information.

Barrier: Behaviorally-irrelevant signals.

Raise long-standing concern: Whether irrelevant signals could conceal some critical facts?

The confusions caused by irrelevant signals

About Encoding	About Decoding
They may mask some small neural components → lead to a partial understanding of neural mechanisms	They may complicate the information readout → hinder the discovery of the true readout mechanism
Smaller R^2 neurons and **smaller variance PC signals** are considered **useless** and **discarded.**	**Linear readout** (biologically plausible, widely used) **or** **Nonlinear readout** (better performance) ?
Truly useless or **Appear useless** ?	performance gap → Irrelevant signals → **Culprits** ?
What's the **role** of these ignored signals ?	

One possible solution: Separating behaviorally-relevant and irrelevant signals at the single-neuron level (such methods remain elusive).

Challenge: The ground truth of relevant signals is unknown, making their definition, extraction, and validation difficult.

Our approach: Propose **a complete framework** to define, extract and validate relevant signals.

2 SEPARATION FRAMEWORK (d-VAE)

Definition : (2 requirements)

1) Similar to raw signals to preserve the underlying neuronal properties.
2) Contain behavioral information as much as possible.

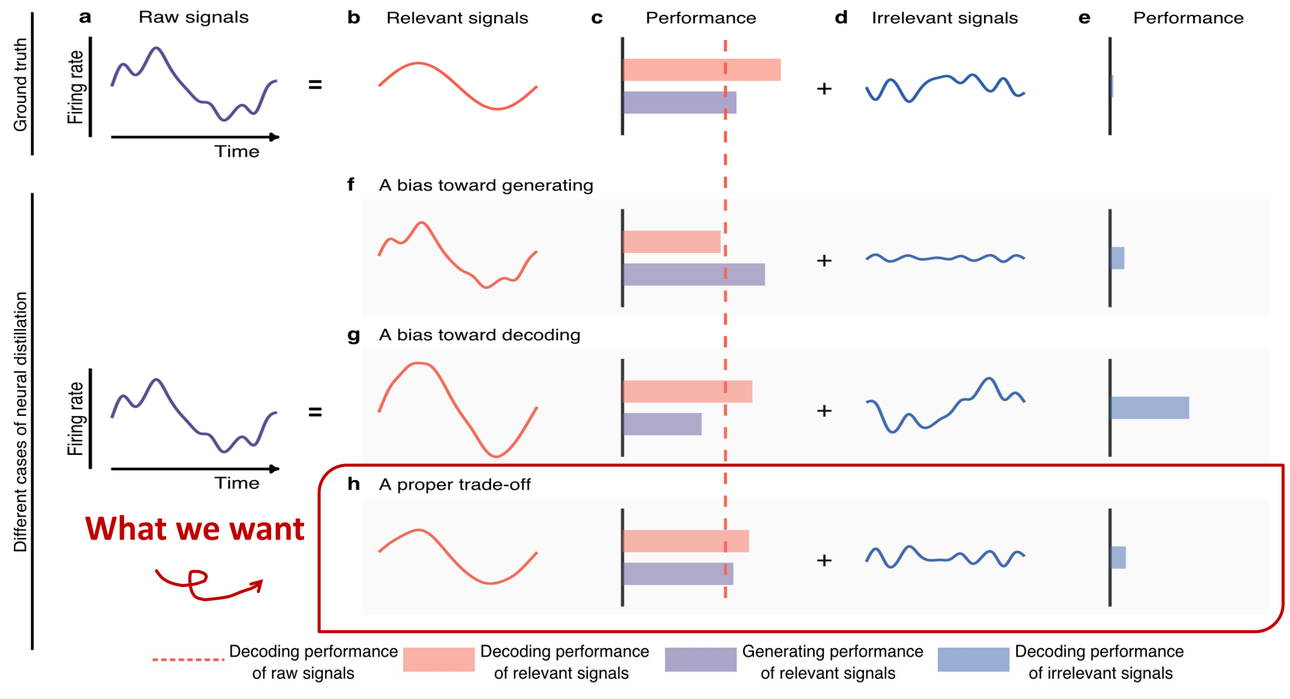
Extraction :

✓ **Assumption:** Irrelevant signals are **noise** relative to relevant signals, which degrade the decoding generalization of generated signals.

✓ **Strategy:** Utilize the **trade-off** between decoding and generation.

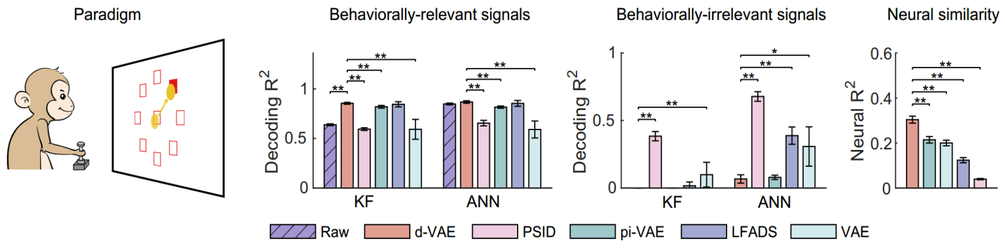
Validation : (3 criteria)

1) **Decoding performance:** relevant signals > raw signals
2) **Decoding performance:** irrelevant signals ≈ 0
3) **Generation performance:** similar to raw signals



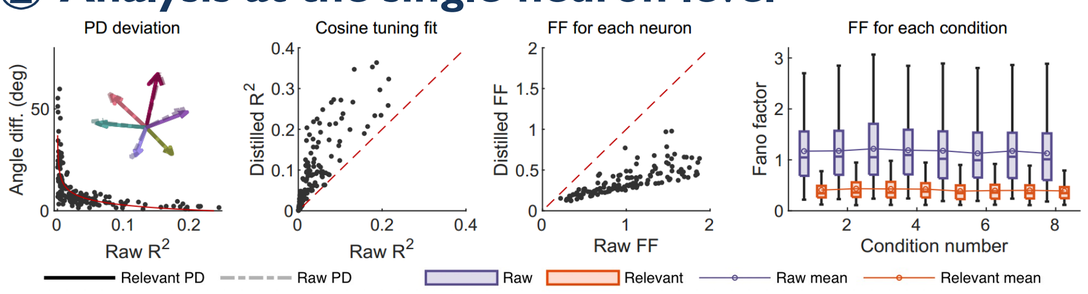
3 RESULTS

① Effectiveness of d-VAE

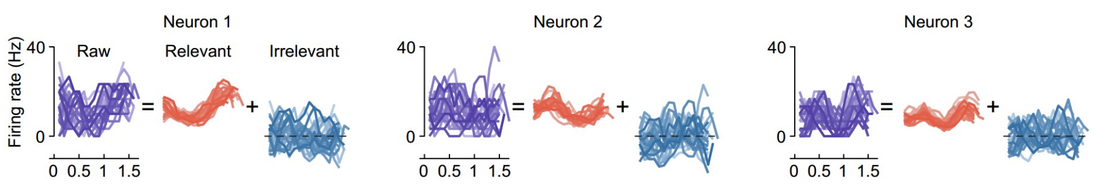


- d-VAE successfully separates the behavioral-relevant and irrelevant signals.

② Analysis at the single-neuron level

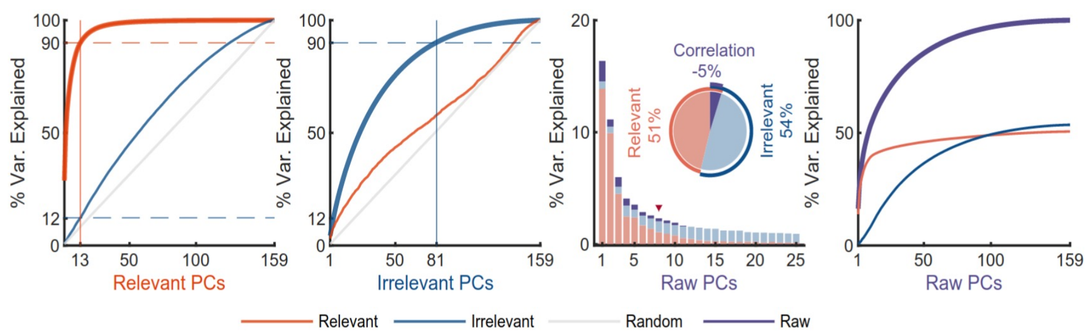


- Irrelevant signals have a large effect on weakly-tuned neurons.
- Irrelevant signals significantly contribute to neuronal variability.



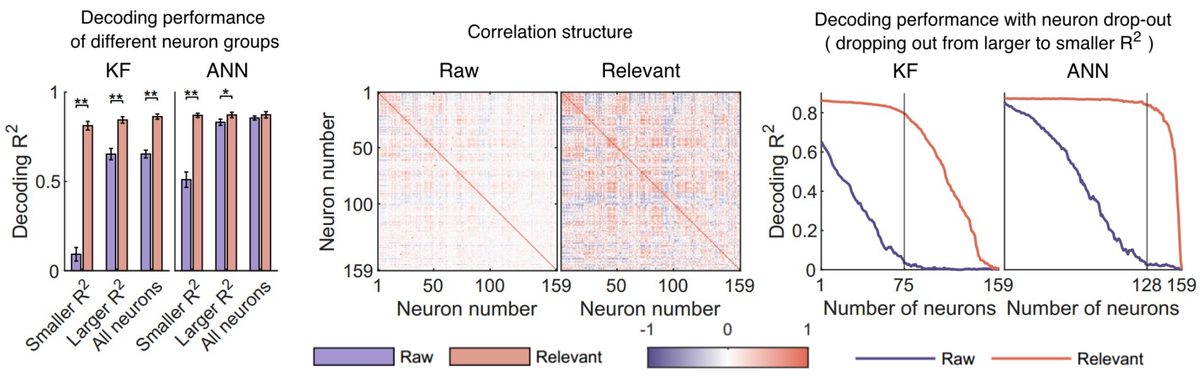
- Neuronal activity visualization : Raw = Relevant + Irrelevant .

③ Analysis at the population level



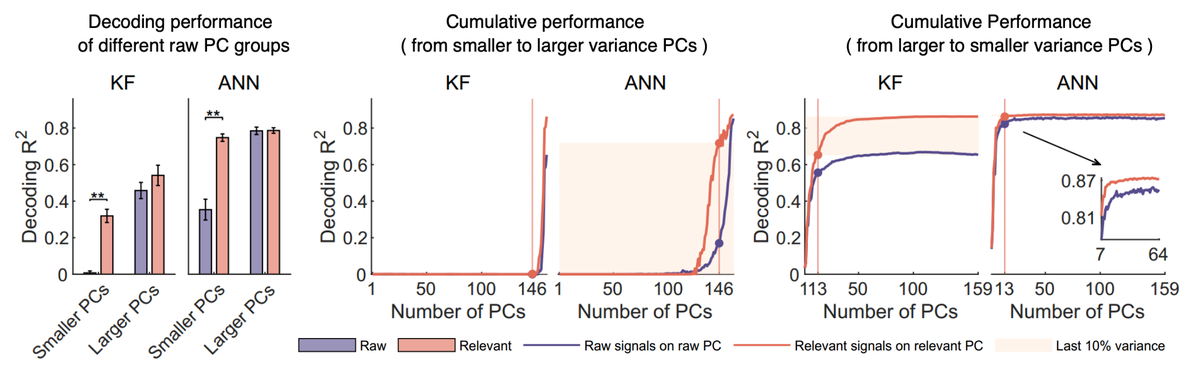
- The primary subspace of relevant signals is nearly orthogonal to irrelevant space.
- Irrelevant signals occupy the dominant proportions in the smaller variance raw PCs.

④ Smaller R^2 neurons actually encode rich information



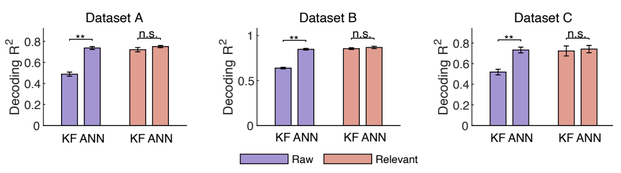
- Smaller R^2 neurons actually encode rich behavioral information, which can not be found before only because irrelevant signals mask their encoded information.
- Irrelevant signals would weaken neuronal correlation.
- Relevant signals are robust to the disturbance of neuron drop-out.

⑤ Smaller variance PCs actually encode rich information



- Smaller variance raw PC signals actually encode rich behavioral information.
- Smaller variance relevant PC signals encode rich information in a nonlinear way.
- The integration of smaller and larger variance PCs results in a synergistic effect, allowing the smaller variance PCs that can not be linearly decoded to significantly enhance the linear decoding performance.

⑥ Motor cortex may use a linear readout mechanism



- Considering the decoding performance and model complexity (Occam's razor), movement behaviors are more likely to be generated by the linear readout.