


NeurIPS 2026

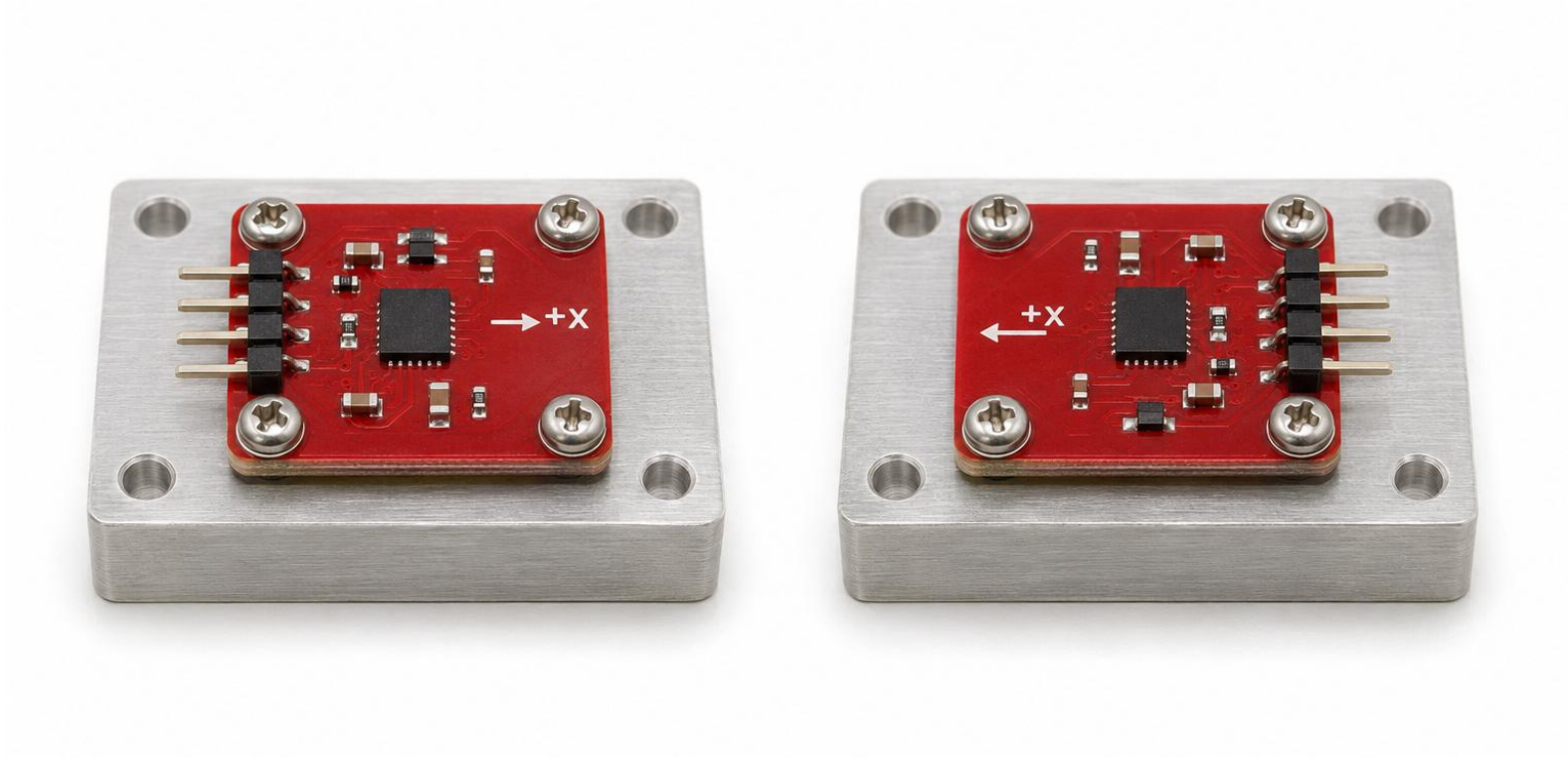
Can Metadata Fix the Gauge?

Calibration Turns Sparse Multi-Source Learning from Sparse PCA into Sparse Mean Recovery

Yirong Xiang[1] Yibo Zhou[2*]
[1] Air Force Medical Center [2] Beihang University

* Correspondence: ybzhou@buaa.edu.cn

A reversed axis can cancel a shared signal



Source A: +V **Source B: −V**

Illustration

Source labels

Which device?

Calibration

Which orientation?

Naive pooling can erase the mean

Source labels leave a sparse PCA experiment

$$W_{\ell r} = \theta R_\ell \sigma_\ell V + \varepsilon_{\ell r}, \qquad n = Ls, \quad \|V\|_0 = k$$

L sources, s samples each. V is k-sparse. σ is an unknown source sign.

Zero mean across source signs

$$X_\ell \sim N\left(0, I_d + \theta^2 s V V^\top\right)$$

Sparse PCA

Signal remains in the covariance

Exact statistical equivalence, including arbitrary feature-based estimators

Calibration restores a common orientation

Gauge–Synchronize–Linearize (GSL)

1 Synchronize

Same sign +1
Opposite sign −1

Relative records
fix source weights

2 Average

$$X_q = \frac{\sum_{\ell,r} q_{\ell r} W_{\ell r}}{\|q\|_2}$$

Alignment restores
a first-order signal

3 Keep top k

Select k coordinates with largest magnitude

The shared
sparse direction

Calibration uses no downstream features or labels

Aligned mass measures restored first-order signal

$$M_q(g) = \frac{(q^\top g)^2}{\|q\|_2^2} = n\cos^2(q, g)$$

q: metadata weights g: hidden signs

Best expected aligned mass

Effective first-order sample size

Source labels only

s

Insufficient for high-F1 recovery

$$\theta^4 ns \lesssim K, \qquad \theta^2 s \ll 1$$

Perfect relative calibration

n

Sufficient for recovery

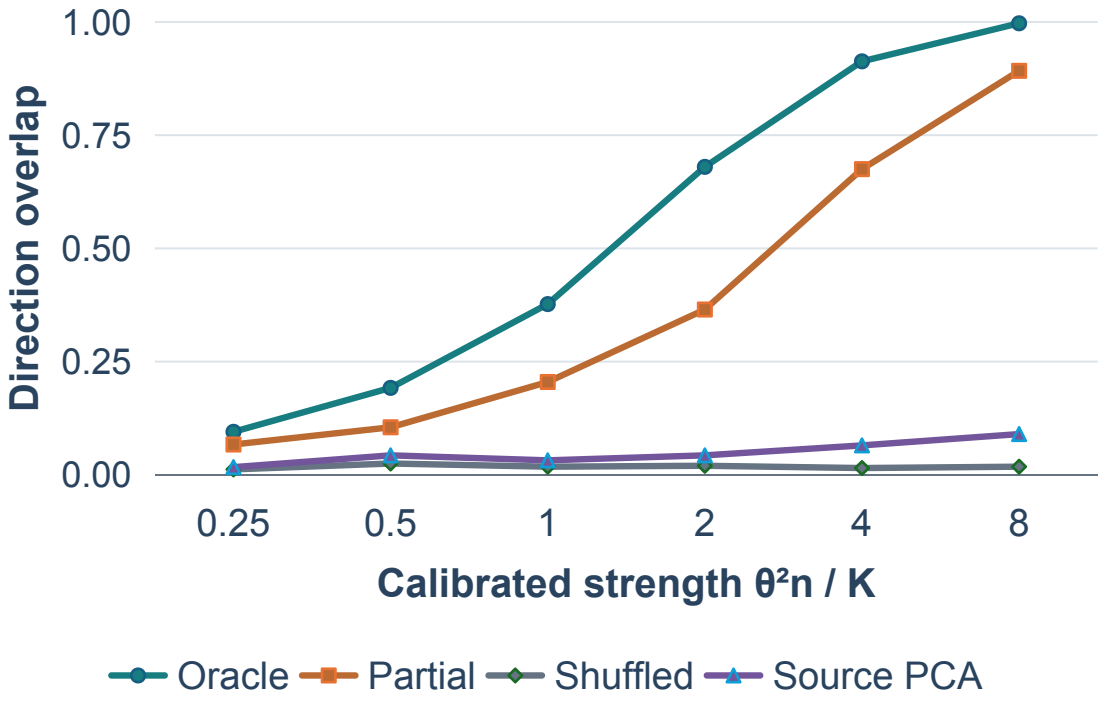
$$\theta^2 \beta^2 n \gtrsim K, \qquad K = k\log(ed/k)$$

β² is synchronization overlap. Enough aligned sources are also required.

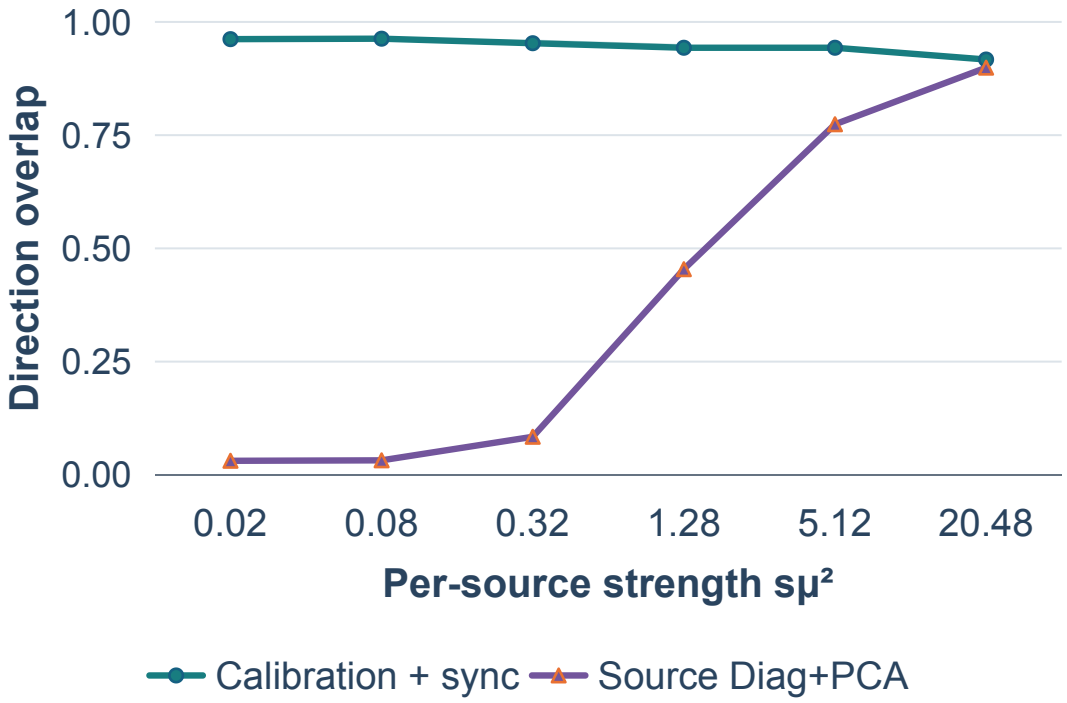
Statistical separation is unconditional. Computational separation assumes sparse-PCA hardness.

Calibration matters most when each source is weak

Calibration quality



Weak-source separation



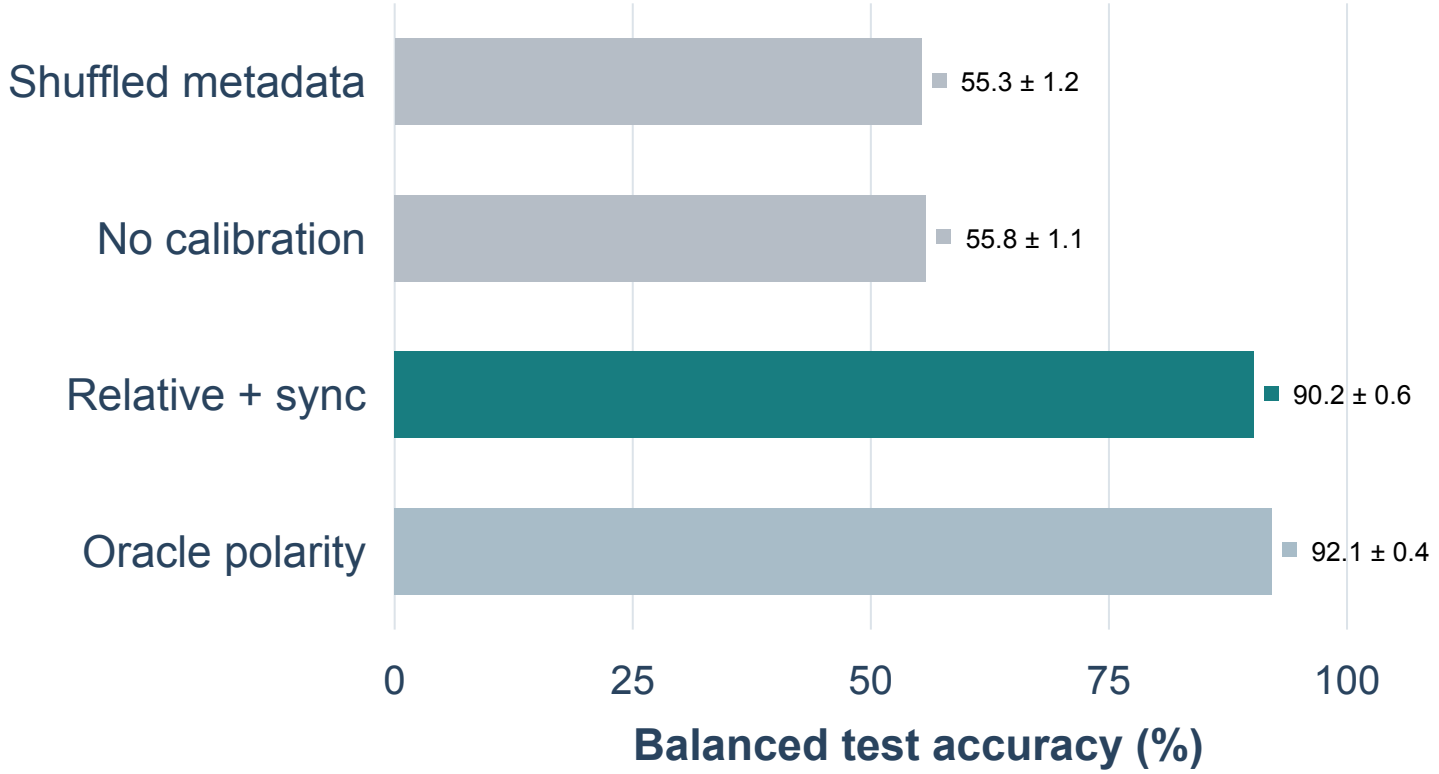
1 = exact direction recovery. Means over 30 runs. Log-spaced x. $\theta^2 = 2\mu^2$.

One calibration supports four physical tasks

24 device mountings
240 reference trials
4 downstream tasks



Illustrative sensor, SparkFun
CC BY 2.0, not the study apparatus



The same alignment is reused unchanged

± source-cluster bootstrap SE

Calibration can change the learning problem

Source labels group observations.
Calibration recovers orientation.

A source convention can be calibrated once and reused across tasks

Theory: sign gauges, sparse signals and Gaussian noise

ybzhou@buaa.edu.cn

arxiv.org/abs/2610.25663

Assumptions behind the separation

Task independence

$$P(\mathfrak{M} \mid V, R, D_{\text{task}}, \sigma, E) = P(\mathfrak{M} \mid \sigma, E)$$

Signal strength and source count

$$M_q(g) = \beta^2 n, \qquad N_{\text{src}} = \beta^2 L$$

$\beta^2 L \gtrsim \log(1/\delta)$ ensures amplitude concentration

Gaussian source coefficients and noise. Equal source sizes. Sparse sign directions.
Computational claims additionally assume sparse-PCA hardness.

Controlled real-feature checks

Real features and labels, with imposed source signs and calibration

Method	Digits accuracy (%)	Breast-cancer accuracy (%)
Clean, no gauge	96.8 ± 0.1	93.3 ± 0.3
Oracle calibration	96.8 ± 0.1	93.3 ± 0.3
Relative + sync	**96.8 ± 0.1**	**93.2 ± 0.3**
No calibration	50.0 ± 0.4	50.1 ± 0.5
Shuffled metadata	53.3 ± 0.7	55.7 ± 0.7

Mean ± SE over 30 seeds