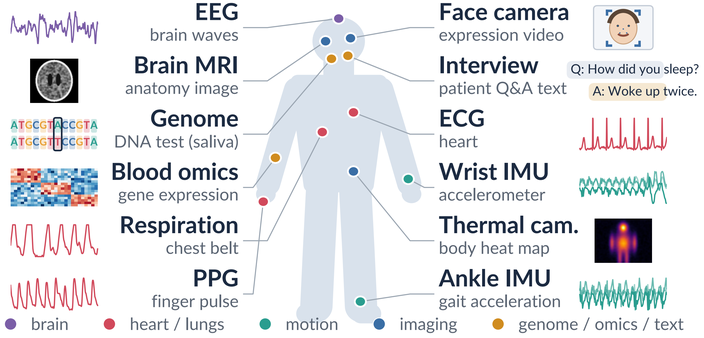
GRAB-MDM: Generalized Robust Adaptive-Bandwidth Multi-View Manifold Learning in High Dimensions with Noise

Xiucai Ding[1] · Chao Shen[2] · Hau-Tieng Wu[3]

[1]Department of Statistics and Data Science, University of California, Davis [2]Amazon [3]Courant Institute of Mathematical Sciences, New York University

NeurIPS 2026 Journal Track
Ann. Statist. 54(5): 2637–2664 (2026)
arXiv:2602.10530

Motivation: one system, many noisy views



each sensor / modality = a **view** $\mathbf{y}^\ell \in \mathbb{R}^{p_\ell}$ of the **same hidden state**
different size, scale & noise ⇒ **goal:** recover the *shared* structure

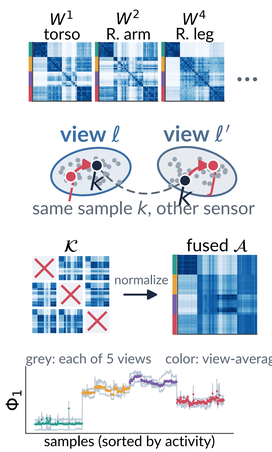
GRAB-MDM in 4 steps

1 $W^\ell(i,j) = K\left(\frac{\|\mathbf{y}_i^\ell - \mathbf{y}_j^\ell\|^2}{\epsilon_\ell}\right)$
a neighbor graph **inside each view**

2 $W^\ell W^{\ell'}, \quad \ell \neq \ell'$
hop in view ℓ, then in view ℓ'

3 $\mathcal{K} = \left[W^\ell W^{\ell'}\right]_{\ell \neq \ell'}$
$\mathcal{A} = D^{-1}\mathcal{K}$
zero diagonal blocks, then normalize

4 $\Phi = \frac{1}{K}\sum_\ell (\text{top eigvecs of } \mathcal{A})^{(\ell)}$
spectral embedding, **averaged over views**



Take-aways

≈ clean data: on par with the best

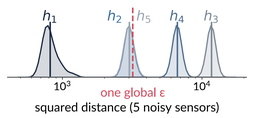
+19 ACC pts, heavy noise: over best baseline

+15 ACC pts, heavy noise: multi-view > concat-view

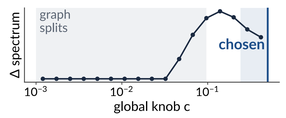
Paper Ding, Shen, Wu. *Ann. Statist.* 54(5): 2637–2664 (2026). doi:10.1214/26-AOS2664 · arXiv:2602.10530
Refs Lindenbaum et al., Multi-view diffusion maps, *Inf. Fusion* 2020 · Coifman & Lafon, Diffusion maps, *ACHA* 2006 · Rodosthenous et al. (mSNE), *PeerJ CS* 2024 · Do & Canzar (mUMAP), *Genome Biol.* 2021
Data UCI Daily & Sports Activities · Sleep-EDF (PhysioNet)

Bandwidth: per-view scale × one global knob

$$\epsilon_\ell = c \cdot h_\ell$$

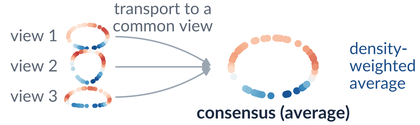


h_ℓ: a percentile of view ℓ's distances — one ϵ cannot fit all



c: where the spectrum of $\mathcal{A}$ stops changing

Theory, in pictures

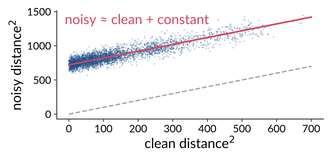


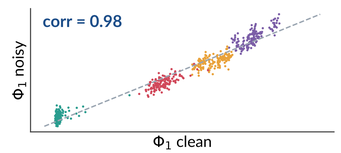
What does $\mathcal{A}$ learn?

$\mathcal{A}f \approx$ **consensus** of views $+ \epsilon\cdot$(mixed Laplacian terms)

views are transported to a common view and averaged (Thm 4.2)

Why is noise harmless? high-dim noise ≈ a **constant shift** of all distances; at high SNR the kernels barely change:
$\|\mathcal{A}_{\text{noisy}} - \mathcal{A}_{\text{clean}}\| \lesssim \frac{1}{\text{SNR}\cdot c} \Rightarrow$ leading eigenvectors stay (Thm 5.2, Prop 5.3)



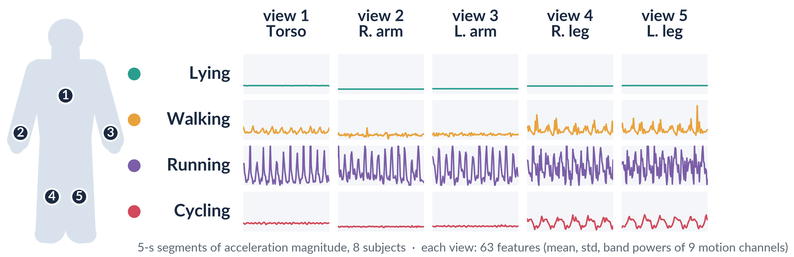


How to set c? too small: graph falls apart; pick c where the spectrum is stable (bandwidth panel)



Real data 1: Daily & Sports Activities (UCI)

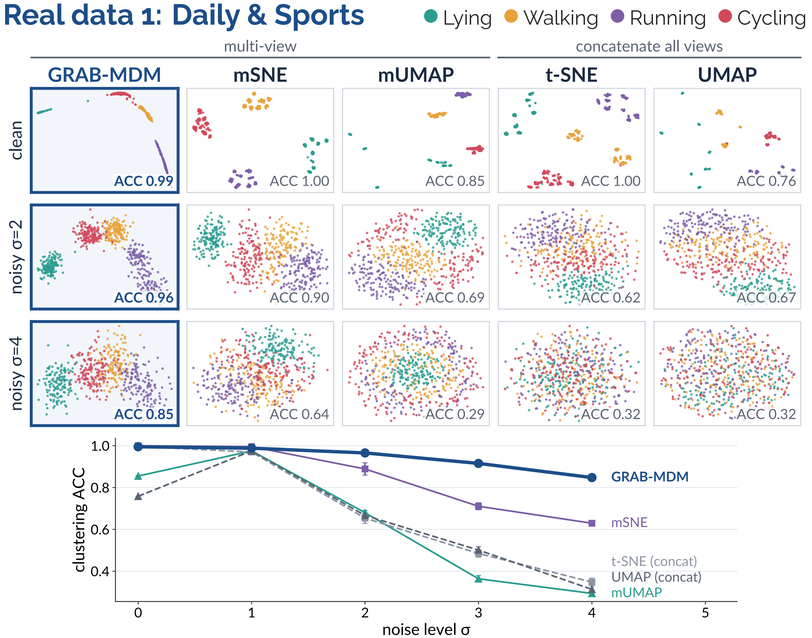
5 body-worn motion sensors = 5 views · 4 activities



5-s segments of acceleration magnitude, 8 subjects · each view: 63 features (mean, std, band powers of 9 motion channels)

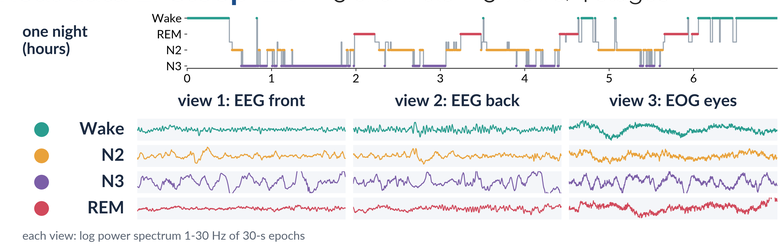
Clean: comparable. Noisy: GRAB-MDM wins big

Real data 1: Daily & Sports

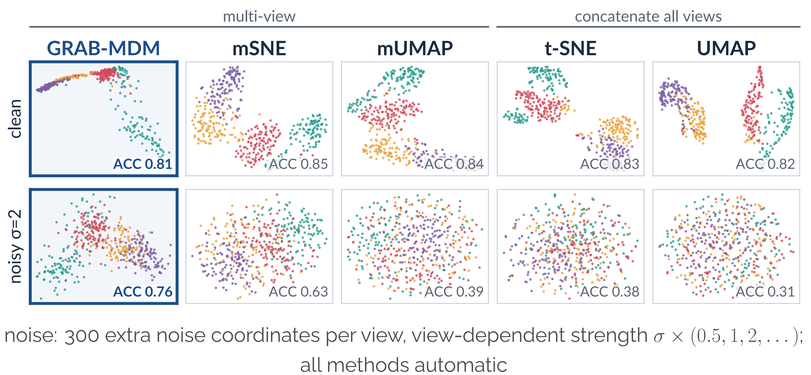


Real data 2: Sleep-EDF

3 channels = 3 views, 4 stages



each view: log power spectrum 1-30 Hz of 30-s epochs



noise: 300 extra noise coordinates per view, view-dependent strength $\sigma \times (0.5, 1, 2, \ldots)$; all methods automatic