

Exploring the Shared Low-Dimensional Subspace of EEG and EOG Signals in Sleep Stages

Ilia Golub¹ and Xueqing Li¹

¹*Gonda Multidisciplinary Brain Research Center, Bar-Ilan University, Ramat Gan, Israel*

1 Abstract

EEG and EOG signals are fundamental components of sleep monitoring, yet their joint dynamics remain underexplored. While these modalities are often treated separately, recent evidence suggests overlapping signal content and potential interactions that reflect underlying neurophysiological coordination. Understanding how this coordination evolves across sleep stages may reveal new insights into brain-eye communication during sleep. **Aim:** To quantify how strongly EEG and EOG signals couple across human sleep stages.

Overnight PSG from 29 adults were segmented into Wake, N1, N2, N3 and REM. Canonical correlation analysis (CCA) extracted the first two shared dimensions (ρ_1, ρ_2) in static stage-wise blocks and in 30 s sliding windows.

Stage-wise ρ_1 and ρ_2 increased from Wake to N3, with REM and N1 showing intermediate coupling. Time-resolved analysis confirmed greater strength and stability in deeper stages, while coupling entropy was highest in Wake and REM.

EEG and EOG share a state-dependent low-dimensional subspace that is strongest in N3, high in N2, moderate in N1 and REM, and minimal in Wake. These findings refine our understanding of brain-eye interactions and support stage-aware sensing and analysis in sleep research.

Keywords: EEG, EOG, CCA

2 Introduction

Sleep stages unfolds through non-rapid eye movement (NREM) sleep, comprising stages N1 to N3, and rapid-eye-movement (REM) sleep, classically scored with polysomnography that monitors electroencephalographic (EEG) activity and electrooculographic (EOG) signals related to eye movements [1]. Growing evidence shows these two channels are not independent: ocular potentials contaminate frontal EEG, while EOG electrodes sample cortical rhythms, and exploiting this overlap improves automatic detection of REM and drowsiness [3,4]. Such findings imply that brain and eye signals cohabit a low-dimensional "communication subspace" whose geometry may vary with different state. Yet the strength and dynamics of this coupling across all stages remain unquantified.

In this study, we seek to quantify this shared subspace across sleep stages by extracting the primary dimensions of joint EEG-EOG variance and systematically assessing how their correlation strength and stability vary as a function of sleep stages. We operationalize this coupling using canonical correlation analysis (CCA), allowing us to track not only average synchrony but also the variability and dimensional richness of shared activity across stages. Using the public Apnea Positive Pressure Long-term Efficacy Study (APPLES) overnight PSG dataset [2], we applied canonical correlation analysis (CCA) to EEG and EOG signals to extract dominant joint components and assess their coupling strength during wake, N1, N2, N3, and REM stages, testing the hypothesis that coupling peaks in REM and diminishes in deeper NREM sleep.

3 Methods

3.1 Dataset

Data for this study were drawn from 29 adult participants in the APPLEES dataset. For each subject, we analyzed four bipolar EEG channels (C3-M2, C4-M1, O1-M2, O2-M1) alongside two EOG channels (LOC and ROC), with all recordings segmented according to the five sleep stages: Wake (W), N1, N2, N3 and REM (R).

3.2 Preprocessing

Preprocessing was carried out in MNE-Python by first loading the EEG and EOG channels from EDF recordings and then importing the corresponding sleep stage annotations. We converted annotation timestamps to seconds relative to each recording’s start time to ensure precise alignment with the signal data and excluded or corrected any segments with alignment issues. Finally, we extracted continuous EEG and EOG epochs for each labeled stage (W, N1, N2, N3, REM) to serve as inputs for our canonical correlation analyses.

3.3 Dimensionality Reduction and Method Selection

We initially compared EEG and EOG subspaces using PCA and ICA followed by subspace angle computation, but both yielded near-zero or ill-defined angles, offering little physiological insight.

3.4 Static CCA Analysis (Per-Stage Aggregation)

We adopted Canonical Correlation Analysis (CCA) for its robustness and interpretability in capturing cross-modal dependence. After mean-centering, a two-component CCA was applied. The first and second canonical correlations (ρ_1 and ρ_2) were obtained by correlating the paired canonical variates. To reduce serial dependence, the canonical time series were resampled at 1 Hz before computing stage-wise summary statistics.

To assess the representational compactness of EEG-EOG coupling, we computed the proportion of signal variance explained by each canonical variate. For each subject and sleep stage, we calculated the variance of the downsampled canonical projections (X_c , Y_c), and expressed each component’s variance as a fraction of the total variance within that domain (EEG or EOG). This provided an estimate of how efficiently the CCA components capture the shared structure across modalities.

3.5 Time-Resolved CCA Analysis

To characterize the temporal evolution of EEG–EOG coupling across sleep stages, the same two-component CCA was run in sliding windows of 30 s with a 15 s step (50% overlap). To better capture long-range trends in coupling dynamics, we averaged CCA values into non-overlapping 10-minute bins across the night, enabling visualization of stage-specific coupling profiles with reduced high-frequency variability. These values were analysed in three ways:

- Stage-wise aggregation: mean and dispersion of ρ_1 and ρ_2 across all windows within each stage.
- Overnight trajectories: 10-min bin averages to visualise stage-specific coupling trends across the recording.
- Distributional complexity: per subject and stage, Shannon entropy, skewness, and kurtosis of the ρ_1/ρ_2 distributions to assess variability and deviation from normality.

To validate the assumption of local stationarity within each 30-second window, we applied both the Augmented Dickey-Fuller (ADF) and KPSS tests to the time-resolved canonical correlation components across subjects. Out of 266 segments, 85% passed the ADF test and 61% passed the KPSS test, with 56% showing agreement across both. Only 10.5% of windows were classified as non-stationary by both tests, supporting the use of 30-second windows for tracking coupling dynamics. In some cases, the KPSS test statistic fell outside the tabulated p-value range, returning a clipped estimate; these were retained to ensure consistency across the dataset.

Statistical comparisons across sleep stages used one-way ANOVA implemented in Python (SciPy.stats).

4 Results

4.1 Static Canonical Correlation Analysis (CCA)

Stage-wise CCA on full-length segments showed a clear depth effect. The first canonical correlation (ρ_1) rose from 0.55 ± 0.14 in Wake to 0.86 ± 0.06 in N3; the second (ρ_2) increased from 0.32 ± 0.15 to 0.61 ± 0.09 (1). One-way ANOVAs confirmed significance (ρ_1 : $F(4, 140)=24.4$, $p<0.001$; ρ_2 : $F(4, 140)=16.7$, $p<0.001$), indicating stronger EEG-EOG synchrony in deeper NREM sleep.

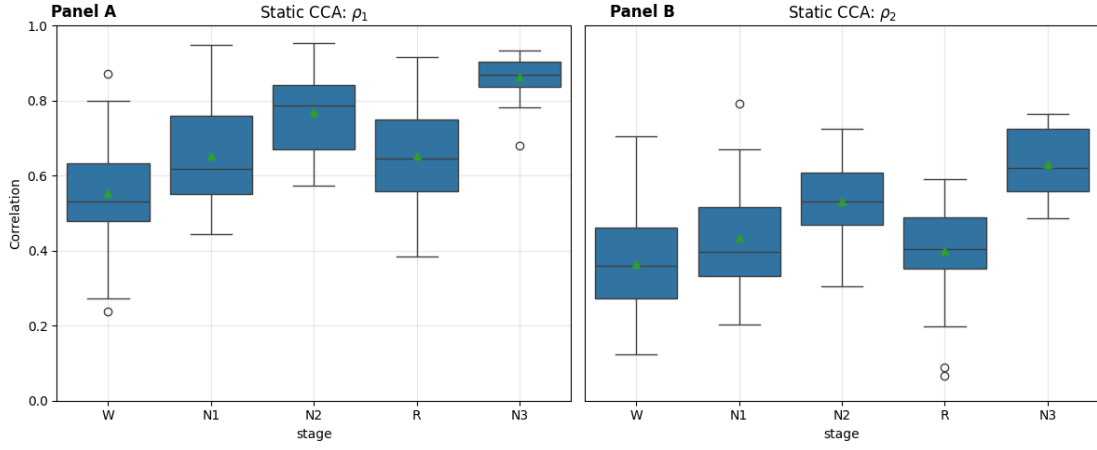


Figure 1: Static canonical correlation analysis (CCA) across sleep stages. Panel A shows the first canonical correlation (ρ_1), and Panel B shows the second (ρ_2), both increasing from Wake to N3.

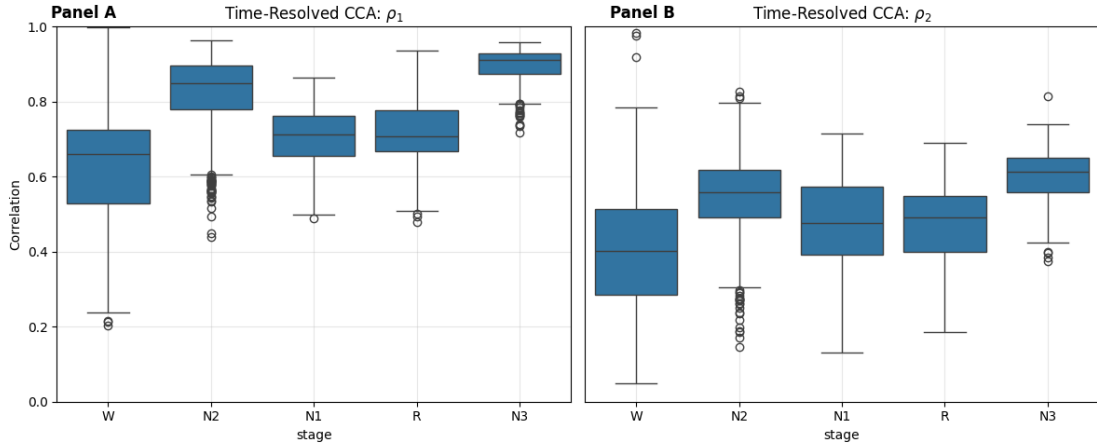


Figure 2: Time-resolved CCA reveals stronger and more stable EEG-EOG coupling in deeper sleep stages.

4.2 Time-Resolved CCA Analysis

Sliding windows (30 s, 50% overlap) reproduced the pattern with finer resolution. Aggregated ρ_1 ranged from 0.73 ± 0.13 (Wake) to 0.87 ± 0.07 (N3); ρ_2 from 0.45 ± 0.17 to 0.65 ± 0.11 (2). Variability diminished with depth, and 10-min bins revealed stable plateaus in N2/N3 but fluctuating profiles in REM and Wake. Transition phases (e.g., N1) showed inconsistent coupling patterns.

Local stationarity analysis confirmed that the majority of 30-second windows met standard weak stationarity criteria. Specifically, 85% of the time-resolved segments passed the ADF test and 61% passed the KPSS test, with both tests agreeing on stationarity in 56% of cases. Only 10.5% of windows were classified as non-stationary by both tests. These results justify the use of fixed-length sliding windows for estimating coupling dynamics.

4.3 Distributional Complexity of Coupling Dynamics

Entropy was highest in Wake and REM and lowest in N3 (3). Skewness (between -0.5 and 0.5) remained near zero, but kurtosis was elevated in Wake/REM, suggesting occasional extreme coupling events.

4.4 Projection Statistics and Downsampled Components

Mean and variance of the individual canonical projections showed no stage effect (all ANOVA $p > 0.17$; 4). Thus, stage differences stem from correlation structure rather than projection amplitude.

We also evaluated how much signal variance was captured by each canonical variate. On average, the first component explained 52-57% of EEG variance and 52-63% of EOG variance across sleep stages. In deeper stages such as N3, variance was more evenly distributed across components, while in REM and Wake, the first component dominated. These findings indicate a low-dimensional, structured coupling that varies with brain state.

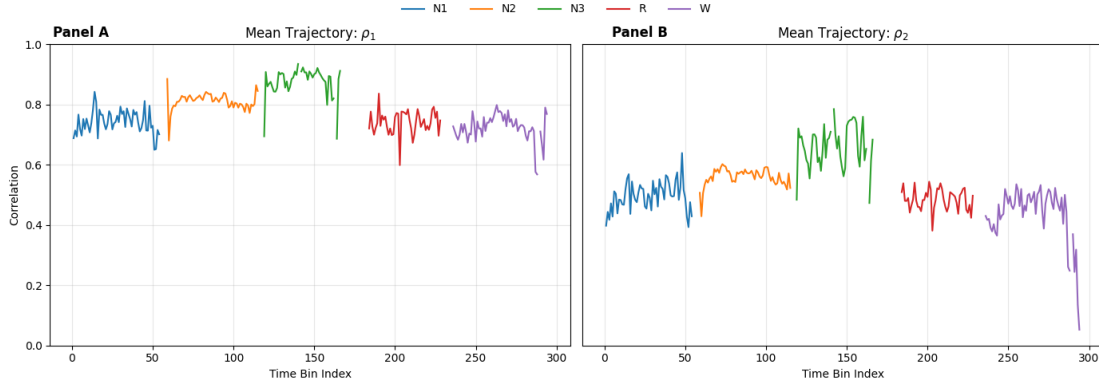


Figure 3: Coupling entropy is highest in Wake and REM, and lowest in N3.

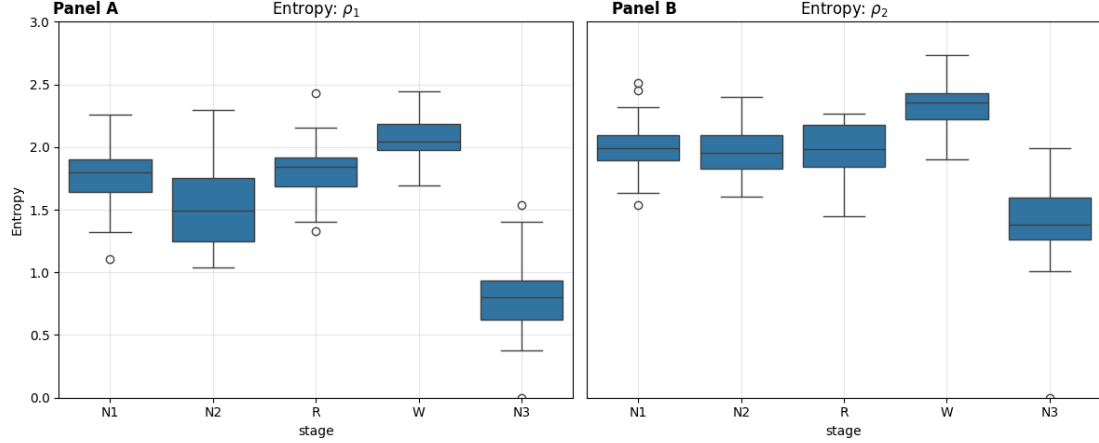


Figure 4: Mean canonical projections remain stable across stages.

5 Discussion

This study quantified the low-dimensional communication subspace shared between EEG and EOG signals across sleep stages using CCA. Both static and time-resolved analyses revealed a robust increase in coupling strength with sleep depth: the first canonical correlation ρ_1 rose from 0.55 ± 0.14 in Wake to 0.86 ± 0.06 in N3, and the second ρ_2 from 0.36 ± 0.15 to 0.63 ± 0.10 . REM and N1 showed intermediate coupling, closer to Wake than to N2/N3 — a pattern replicated across time windows, further supporting the sleep-stage specificity of EEG-EOG coupling. Temporal trajectories showed stable coupling in N2/N3 and greater variability in Wake and REM as they showing lower correlation values and higher entropy, supported by entropy measures. Lower entropy in deep sleep suggests more consistent EEG-EOG coordination, potentially reflecting reduced sensory input and motor variability. In contrast, higher entropy in Wake and REM may reflect more complex and dynamic brain-eye interactions. The persistent contribution of ρ_2 across all stages, especially in Wake and REM, suggests that EEG-EOG interactions are not dominated by a single mode, but instead reflect multiplexed coupling mechanisms in more active states. These findings may have practical implications for both research and clinical applications. Given the sensitivity of EEG-EOG coupling to sleep depth and stage transitions, CCA-derived features could enhance automatic sleep staging algorithms, particularly in distinguishing REM from N1 or Wake. Furthermore, altered coupling profiles may serve as markers of disrupted sleep architecture in conditions such as obstructive sleep apnea or REM behavior disorder. Future studies should investigate how these coupling metrics vary in pathological sleep and whether they offer diagnostic or prognostic value. The analysis of explained variance provided additional insight into the geometry of the EEG-EOG subspace. The fact that just two canonical components captured nearly all relevant variance supports the hypothesis that EEG and EOG signals cohabit a compact, low-dimensional space. This compactness was most pronounced in N2 and N3, where variance was more evenly distributed across components, and less so in REM and Wake, which were dominated by a single component. These patterns mirror the stage-specific entropy and correlation results, reinforcing the interpretation that sleep depth modulates not only coupling strength, but also its complexity and dimensional shape. Furthermore, our validation of local stationarity across the majority of analysis windows supports the reliability of time-resolved CCA estimates. The rare non-stationary windows—primarily in Wake and REM—likely reflect

physiological transients or stage transitions rather than methodological artifacts. Together, these results strengthen the foundation for using CCA-based coupling metrics in sleep research.

Limitations This study has several limitations. First, although preprocessing was applied, residual artifacts (e.g., blinks, muscle activity) may have influenced coupling estimates, especially in wake and REM stages. Second, the APPLES dataset includes individuals with varying sleep pathologies (e.g., OSAS), but we did not stratify by clinical status due to metadata constraints, limiting generalizability. Third, our analysis focused on a subset of EEG (C3, C4) and EOG (LOC, ROC) channels; while representative, this may miss spatially distributed coupling patterns. Fourth, our use of 10-minute time bins improved stability but may have smoothed out transient events during REM or stage transitions. Finally, CCA captures only linear relationships; incorporating nonlinear approaches (e.g., kernel CCA or mutual information) may reveal richer coupling dynamics in future work.

Conclusion This study shows that EEG and EOG signals share a low-dimensional subspace whose strength and temporal stability vary systematically across sleep stages. Coupling was strongest and most consistent in N3 and N2, moderate in N1 and REM, and weakest in Wake. These findings suggest structured coordination between cortical and oculomotor systems that tracks sleep depth. By focusing on intermodality dynamics via CCA, our approach offers a compact, interpretable view of sleep architecture that may inform future studies of sleep disorders and altered states.

References

1. G.-R. Liu, T.-Y. Lin, H.-T. Wu, Y.-C. Sheu, C.-L. Liu, W.-T. Liu, M.-C. Yang, Y.-L. Ni, K.-T. Chou, C.-H. Chen, D. Wu, C.-C. Lan, K.-L. Chiu, H.-Y. Chiu, and Y.-L. Lo. Large-scale assessment of consistency in sleep stage scoring rules among multiple sleep centers using an interpretable machine learning algorithm. *Journal of Clinical Sleep Medicine*, 17(2):159–166, 2021.
2. R. Mueller. Sleep data—national sleep research resource—nsrr. <https://doi.org/10.25822/63pr-a591>, n.d. Accessed from <https://doi.org/10.25822/63pr-a591>.
3. D. Safieddine, A. Kachenoura, L. Albera, G. Birot, A. Karfoul, A. Pasnicu, A. Biraben, F. Wendling, L. Senhadji, and I. Merlet. Removal of muscle artifact from eeg data: Comparison between stochastic (ica and cca) and deterministic (emd and wavelet-based) approaches. *EURASIP Journal on Advances in Signal Processing*, 2012(1):127, 2012.
4. X. Xu, B. Zhang, T. Xu, and J. Tang. An effective and interpretable sleep stage classification approach using multi-domain electroencephalogram and electrooculogram features. *Bioengineering*, 12(3):286, 2025.