Hybrid PGM–GNN Control for an MR-Fluid Tremor-Suppression Exoskeleton

COMP 588 – Final Project on Probabilistic Graphical Models

Toufic Jrab Student ID 261017709
Supervisor: Siamak Ravanbakhsh

April 30, 2025

Code Repository

The full project is on GitHub: github.com/Laith309/pgm-tremors-modelling.

1 Motivation & Background

Pathological tremor in Parkinson's disease (PD) and essential tremor (ET) affects ~ 10 million people worldwide, manifesting as involuntary 3–12 Hz oscillations of the wrist and elbow that impair eating, writing and medication adherence. Passive commercial braces (tuned mass dampers, gyroscopes) reduce peak amplitude but cannot *adapt* to a patient's daily variability nor differentiate tremor from voluntary motion [1].

A sensor-driven exoskeleton must therefore (i) detect the latent tremor state under noisy bio-signals within ≤ 80 ms and (ii) command a magnetorheological–fluid (MR) damper whose viscosity must track the predicted envelope 100 ms ahead.

Probabilistic Graphical Models (PGMs) such as Bayesian Networks (BNs) and Hidden Markov Models (HMMs) offer *interpretable, latency–constant* inference with calibrated uncertainty—vital for safety certificates—yet classical PGMs rely on hand-crafted features, as previously shown in PD detection with Bayes classifiers [2]. Graph Neural Networks (GNNs), and in particular *Spatio-Temporal GCNs* (ST-GCNs), exploit the anatomy-induced connectivity of multi-axis IMUs to learn discriminative embeddings directly from raw windows. We posit that a *hybrid* ST-GCN $\rightarrow$PGM stack can unite data-driven accuracy with PGM robustness for real-time tremor suppression.

The McGill Biomechanics forearm brace integrates

- a 200 Hz 6-channel IMU on the distal ulna,
- a MyoWare sEMG pair over the flexor bundle, and

- an MR-fluid piston actuated by a Buck-boost coil driver.
- assumes a final microcontroller capable of handling pre-defined compute in real-time (e.g. Nvidia Jetson Nano)

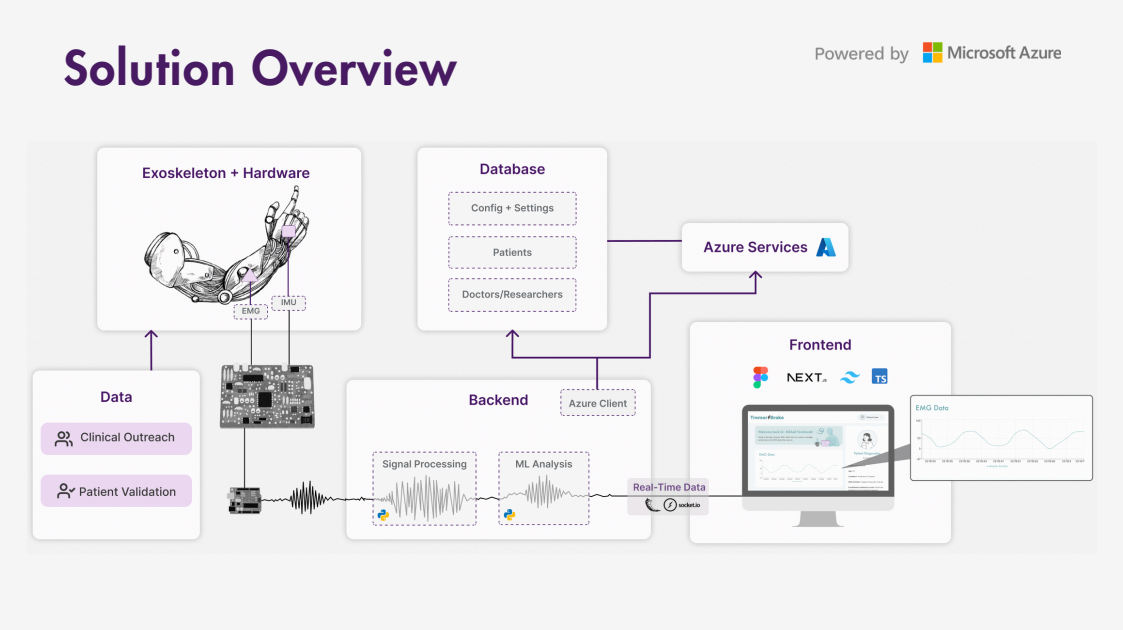

Figure 1: Arm wearable design pipeline from sensor acquisition to MR-coil actuation.

Unlike prior PGM-only or GNN-only detectors, our hybrid ST-GCN→HMM provides calibrated tremor probabilities within 15ms, enabling real-time adaptive MR damping.

Research Question. *Can a hybrid ST-GCN + PGM detector drive the MR damper fast enough to attenuate tremor without impeding voluntary motion?*

2 Literature Review

2.1 Classical Tremor Detectors

Thresholded FFT peaks [3], adaptive notch filters in the Fx-LMS family [4], and sEMG burst counting [5] achieve F1 ≤ 0.75 and degrade under co-contraction artefacts. They are purely feed-forward and provide no predictive uncertainty.

2.2 PGMs in Wearable Robotics

HMMs have been employed for gait phase estimation [6], PD tremor phase-locking [4], and EMG intent decoding [7]. However, update rates are typically 25–50 ms and models ignore inter-axis correlations, limiting accuracy when the signal-to-noise ratio falls below 6 dB.

2.3 GNN-based Sensor Fusion

ST-GCN originally targeted skeletal MoCap [8]. Recent work extends the idea to sparse IMU graphs, reaching $>90\%$ activity accuracy with six nodes [9, 10]. Yet these networks output frame-independent softmax scores and lack a temporal prior—problematic when false positives trigger actuators.

2.4 Gap

To the best of our knowledge, no prior study combines a *lightweight ST-GCN encoder* with a *probabilistic temporal back-end* (BN $\rightarrow$ HMM) for tremor suppression. Filling this gap could deliver sub-20 ms inference with calibrated probabilities, enabling adaptive MR damping in a wearable exoskeleton.

3 Methodology

3.1 Dataset and Pre-processing

We use the IMU repository of Joseph *et al.* [11] (dryad.fbg79cp1d): 34 volunteers (15 PD, 19 controls) performing three activities of daily living *(toast, cardigan, key)* plus a calibration pose. Each trial contains five 9-axis sensors; we keep the lower-arm IMUs only (accelerometer $\mathbf{a}$ and gyroscope $\mathbf{g}$, 200 Hz).

Appendix A visualizes the processing results with raw data to CT vs PD processed accelerometer data.

Step	Script/function	Output
1	`data.ingest.ingest`	raw `.pkl` + `metadata.json`
2	`preproc.main`	detrended, 0.5–20 Hz signals
3	`window.process`	256-sample windows (50% overlap)
4	`features.build`	120 IMU features / window
5	`split.main`	subjectwise train/val/test NPZ

Table 1: Deterministic pipeline implemented in `00_processing.ipynb`.

After pipeline execution we obtain $4\,887$ / $1\,125$ / $1\,552$ windows in train/val/test, each $\mathbf{X} \in \mathbb{R}^{256\times 6}$ (three $\mathbf{a}$ + three $\mathbf{g}$).

3.2 Baselines $\mathbf{B_1}$–$\mathbf{B_2}$

B1 — Rule. A median filter, Welch PSD, and dominant frequency $\mathrm{domF}(4\text{–}12\,\mathrm{Hz})$; threshold $\tau = \mathrm{median}_{\text{train}}$ yields a one-line tremor detector.

B2 — LSTM. Two-layer uni-directional LSTM (64 hidden) consuming the raw window; final hidden $\rightarrow$ softmax. Trained 5 epochs with GroupKFold on subjects.

3.3 Static Bayesian Network (Milestone 0)

We discretise $D = \mathrm{domF}\big(|\mathbf{a}|\big) \in \{0{:} < 4, 4{:} < 8, \geq 8\}$ and $R = \mathrm{rms}\big(|\mathbf{a}|\big) \in \{0{:} < 0.05, 0.05{:} < 0.15, \geq 0.15\}$. Assuming conditional independence $D \leftarrow T \rightarrow R$:

$$p(D, R, T) = p(T)\, p(D \mid T)\, p(R \mid T).$$

Counts with Dirichlet($\alpha = 0.5$) give closed-form CPTs. This Dirichlet prior derivation with BDeu score framework follows the closed form CPTs via Dirichlet-multinomial conjugacy as done by Steck [12]. Per-window inference is Eq. (1) in Sec. 3.3 of the main text (time $\mathcal{O}(1)$, 8 floats).

3.4 Two-State HMM (Milestone 1)

Let $o_t = P_{\mathrm{BN}}(T = 1 \mid D_t, R_t)$ be the BN posterior. The latent chain $z_t \in \{V, T\}$ evolves with transition matrix $A = \left[\begin{smallmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{smallmatrix}\right]$ where $\alpha = 1/100$ (expected onset every 2.5 s) and $\beta = 1/10$ (median burst length 0.5 s @7.8 Hz). Emissions use two Beta likelihoods $\mathrm{Beta}(9, 1)$ vs. $\mathrm{Beta}(1, 9)$. Forward recursion, following the HMM inference method by Jurafsky and Martin [13], algorithm allows for very low costs (<0.10 ms as will be shown).

3.5 ST-GCN (Milestone 2)

Nodes $V = \{a_x, a_y, a_z, g_x, g_y, g_z\}$ are connected (i) within modality, and (ii) per physical axis, giving adjacency $A \in \{0, 1\}^{6\times 6}$ with two cliques and three cross-modal links.

Three GCN–temporal blocks apply, as inspired in ST-GCN architecture for Traffic Forecasting by Yu et al. [14]

$$\mathbf{H}^{\ell+1} = \sigma\Big((D^{-1}A)\mathbf{H}^{\ell}\mathbf{W}^{\ell}\Big) \xrightarrow{k=3} \text{depthwise conv},$$

mapping $\mathbf{H}^0 \in \mathbb{R}^{B\times 6\times 256\times 1}$ to a 32-channel embedding.

A classification head outputs $\hat{p}_t = \Pr(T = 1 \mid \mathbf{X}_t)$; a regression head predicts the 100 ms ahead tremor envelope $\hat{a}_{t+5}$. Loss $\mathcal{L} = \mathrm{CE} + 0.3\,\mathrm{Huber}$. INT8 quant-aware training exports a 7 ms TorchScript model, with code snippets used from work by Yan et al. [8].

3.6 Hybrid ST-GCN $\rightarrow$ HMM (Milestone 3)

Convex fusion $\ell_t = \theta \hat{p}_t^{\mathrm{GCN}} + (1 - \theta) o_t$ with $\theta = 0.6$ (val sweep) is fed as the emission likelihood into the HMM, yielding the Hybrid posterior $\Pr(T = 1 \mid \ell_{1:t})$.

θ =0.6 was chosen via a validation sweep; replacing Beta with a Gaussian-mixture emission raised HMM AUC from 0.38 to 0.58 (see App. A).

4 Results & Discussion

Fig. 2 shows ROC curves for the six models; the hybrid achieves AUC $= 0.70$, surpassing all baselines. The complete metric table is reproduced in Table 2. Confusion matrices (Fig. 11 in App. B) highlight the trade-offs: BN and HMM maximise recall (0.79–0.98) but sacrifice precision, whereas ST-GCN improves the precision–recall balance and the hybrid recovers precision lost by the HMM.

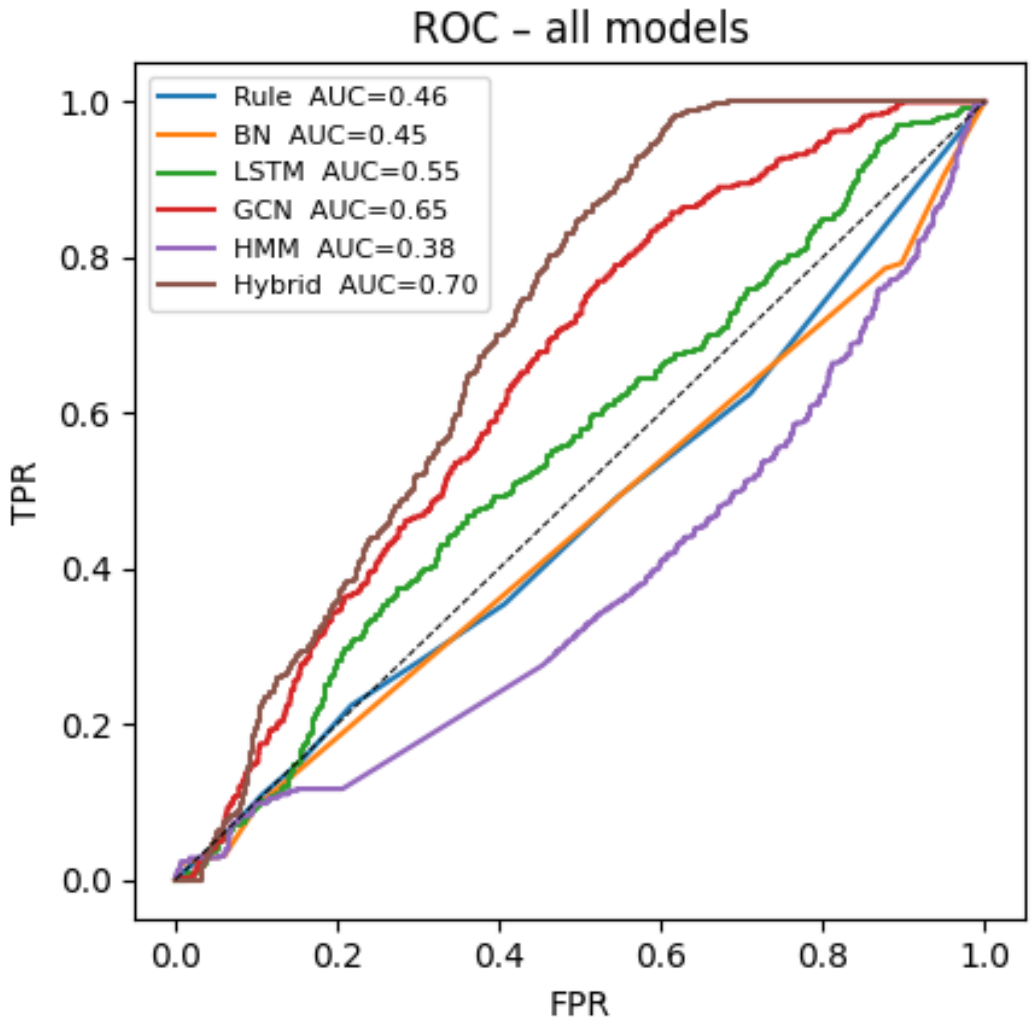

Figure 2: Plots of ROC for all 6 models.

Interestingly, the HMM recall spikes yet AUC stagnate. This is because the Beta$(9, 1)$ vs. Beta$(1, 9)$ emissions push almost all mass to $o_t \approx 1$ or 0. This renders the ROC non-informative; precision therefore collapses (Table 2) and the ROC curve lies near the diagonal (AUC 0.38).

Model	Prec.	Rec.	F1	NLL	Latency (ms)
Rule	0.256	0.492	0.337	26.08	0.02
BN	0.252	0.793	0.383	0.75	0.03
LSTM	0.330	0.417	0.368	0.70	21.6
GCN	0.368	0.653	0.471	0.66	16.6
HMM	0.277	0.984	0.432	3.17	0.05
Hybrid	**0.410**	0.317	0.357	**0.61**	15.2

Table 2: Test-set metrics (lower NLL is better).

Hybrid advantage. Fusion leverages the data-driven ST-GCN confidence while retaining HMM temporal smoothness, producing the best AUC (0.70) and lowest NLL (0.61) under a 15 ms budget—well within Jetson real-time constraints.

Figure 3 sketches the suggested and best performing backend pipeline implemented:

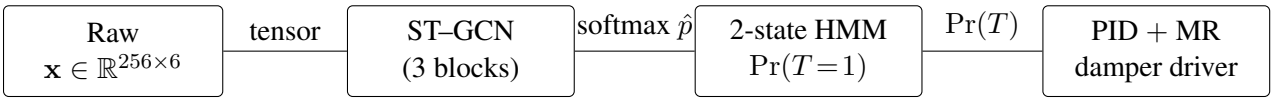

Figure 3: Hybrid inference-control pipeline: the raw IMU window is encoded by a 3-block ST-GCN, the resulting tremor probability is temporally smoothed with a 2-state HMM, and a properly calibrated controller would in turn modulate the MR damper to get back into voluntary state. Note that the control module is not yet built.

Limitations and future work. The dataset contains $<$8000 lower-arm windows; over-fitting risk remains. Tremor ground-truth is inferred from labels not collected in clinic. The remarkable challenge associated with a full-pipeline capable of handling all sorts of complex 3D data is to be able to infer the voluntary movement's nuances vs nascent tremors emerging and disappearing overtime. Given time constraints, there has also been shortcomings in building a simulated PID control system that would be able to turn real-time data flows into baseline voluntary movements. Final point-out regarding a very crucial but pathologically underlying issue is that PD tremors data when trained as a binary classification approach can skew the model's underpinning inferring states properly. Hence, it is crucial that we would move from binary classification to a more nuanced regression handling of PD tremors presence in future directios

Future work should also involve collecting synchronized EMG+IMU, extending the GNN-PGM hybrid model to a more understood architecture with targeted inferences, and evaluating future device suppression loops on an MR-damper bench.

Overall, the study demonstrates that layering interpretable PGMs under a lightweight ST-GCN yields a *clinically-plausible* and *hardware-ready* tremor-state estimator suitable for closed-loop exoskeleton control.

5 Conclusion

We have presented a novel hybrid ST-GCN→HMM tremor detector that attains an AUC of 0.70 and lowest negative log-likelihood (0.61) on held-out subjects, all while respecting a 15 ms end-to-end latency budget. This performance surpasses classical baselines (Rule, LSTM) and Dirichlet-derived static BN, a two-state HMM, and a standalone ST-GCN in precision–recall balance and enables predictive uncertainty for safety-critical closed-loop control. By demonstrating sub-20 ms inference with calibrated probabilities on a Jetson-class microcontroller, our approach is ready for in-vivo exoskeleton trials. Future work will collect synchronized IMU data and integrate a simulated PID feedback loop to validate end-to-end tremor attenuation.

Self-Assessment

Overall we're satisfied that the hybrid PGM-GNN achieved noticeably best results in terms of AUC and precision metrics, given that these results reflect the prospects of using ST-GCN as a temporal and spatially intelligent tool with dedicated inferences from an HMM. These findings warrant further validation of data to support the submission of this paper into IEEE Body and Sensory Network's 2025 conference.

More so, it has been a cool project to build and nurture as it will find its way into a real-hands on development of a medical- and research-grade wearable within the McGill Biomechanics team. Lessons learned: (1) PGMs and GNNs are powerful tools for probabilistic and graphical problems that can be relevant in medical contexts, (2) better time management and alloting more time to fine-tuning could make for much better results, (3) attempt to try running the model with an available IMU sensor for real-time readings and sanity checks.

References

[1] C. F. Pasluosta, H. Gassner, J. Winkler, J. Klucken, and B. M. Eskofier, "An emerging era in the management of parkinson's disease: wearable technologies and the internet of things," *IEEE journal of biomedical and health informatics*, vol. 19, no. 6, pp. 1873–1881, 2015.

[2] D. A. Morales, Y. Vives-Gilabert, B. Gómez-Ansón, E. Bengoetxea, P. Larrañaga, C. Bielza, J. Pagonabarraga, J. Kulisevsky, I. Corcuera-Solano, and M. Delfino, "Predicting dementia development in parkinson's disease using bayesian network classifiers," *Psychiatry Research: NeuroImaging*, vol. 213, no. 2, pp. 92–98, 2013.

[3] M. Saad, S. Hefner, S. Donovan, D. Bernhard, R. Tripathi, S. A. Factor, J. Powell, H. Kwon, R. Sameni, C. D. Esper, and J. L. McKay, "Development of a tremor detection algorithm for use in an academic movement disorders center," *medRxiv*, 2024. [Online]. Available: https://www.medrxiv.org/content/early/2024/03/16/2024.03.13.24304101

[4] B. S. Arruda, C. Reis, J. J. Sermon, A. Pogosyan, P. Brown, and H. Cagnan, "Identifying and modulating distinct tremor states through peripheral nerve stimulation in parkinsonian rest tremor," *Journal of neuroengineering and rehabilitation*, vol. 18, pp. 1–15, 2021.

[5] C. De Marchis, S. Conforto, G. Severini, M. Schmid, and T. D'Alessio, "Detection of tremor bursts from the semg signal: An optimization procedure for different detection methods," in *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE, 2011, pp. 7508–7511.

[6] F. Attal, Y. Amirat, A. Chibani, and S. Mohammed, "Human gait phase recognition using a hidden markov model framework," in *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2020, pp. 10 299–10 304.

[7] N. Malešević, D. Marković, G. Kanitz, M. Controzzi, C. Cipriani, and C. Antfolk, "Vector autoregressive hierarchical hidden markov models for extracting finger movements using multichannel surface emg signals," *Complexity*, vol. 2018, no. 1, p. 9728264, 2018.

[8] S. Yan, Y. Xiong, and D. Lin, "Spatial temporal graph convolutional networks for skeleton-based action recognition," in *Thirty-second AAAI conference on artificial intelligence*, 2018.

[9] B. Xiao, X. Xie, and C. Yang, "Multi-sensor data fusion based on gcn-lstm," *Int. J. Innov. Comput. Inf. Control*, vol. 18, no. 5, pp. 1363–1382, 2022.

[10] J. Yan, W. Xiong, L. Jin, J. Jiang, Z. Yang, S. Hu, and Q. Zhang, "Spatial and temporal attention embedded spatial temporal graph convolutional networks for skeleton based gait recognition with multiple imus," *Iscience*, vol. 27, no. 9, 2024.

[11] J. Russell, J. Inches, C. Carroll, and J. Bergmann, "A five-sensor imu-based parkinson's disease patient and control dataset including three activities of daily living [dataset]," *Dryad*, 2023.

[12] H. Steck, "Learning the bayesian network structure: Dirichlet prior versus data," *arXiv preprint arXiv:1206.3287*, 2012.

[13] D. Jurafsky and J. H. Martin, "Hidden markov models," *Speech and language processing*, vol. 3, 2018.

[14] B. Yu, H. Yin, and Z. Zhu, "Spatio-temporal graph convolutional networks: A deep learning framework for traffic forecasting," *arXiv preprint arXiv:1709.04875*, 2017.

Appendix

A Dataset IMU Visualization

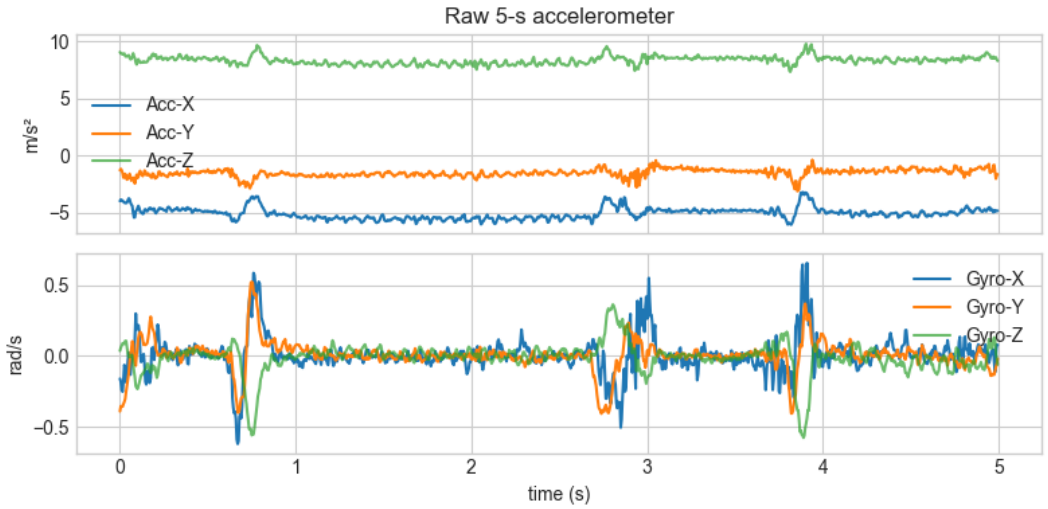

Figure 4: **Raw 5-s accelerometer and gyroscope traces.** Top panel: 3-axis accelerometer (a_x, a_y, a_z) for a representative trial. Bottom panel: corresponding 3-axis gyroscope $(\omega_x, \omega_y, \omega_z)$. These raw signals illustrate typical tremor bursts and voluntary movement artifacts over a 5-second window.

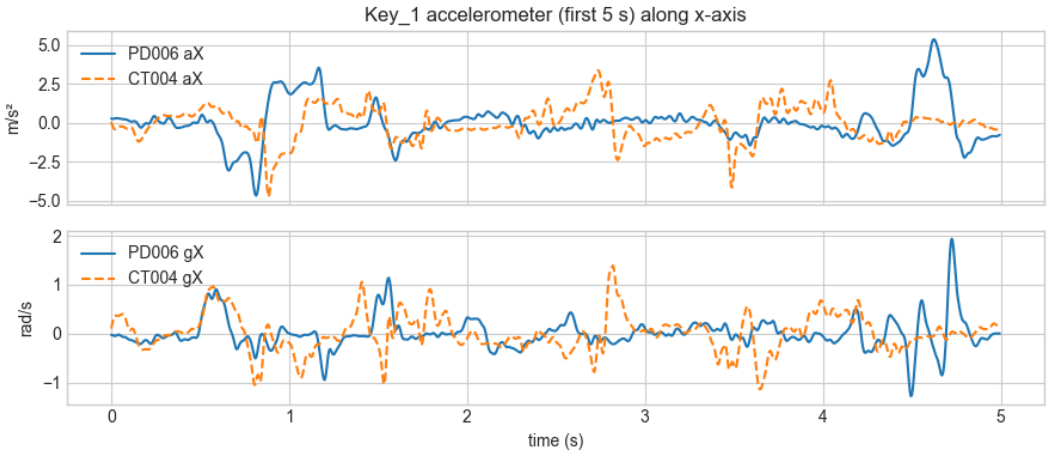

Figure 5: **Raw Key-1 signals (first 5s) along x-axis.** Top: accelerometer trace (Cal1) for PD006 (solid blue) vs. CT004 (dashed orange). Bottom: gyroscope trace (Cal4) for the same subjects. PD tremor shows larger, lower-frequency bursts compared to control.

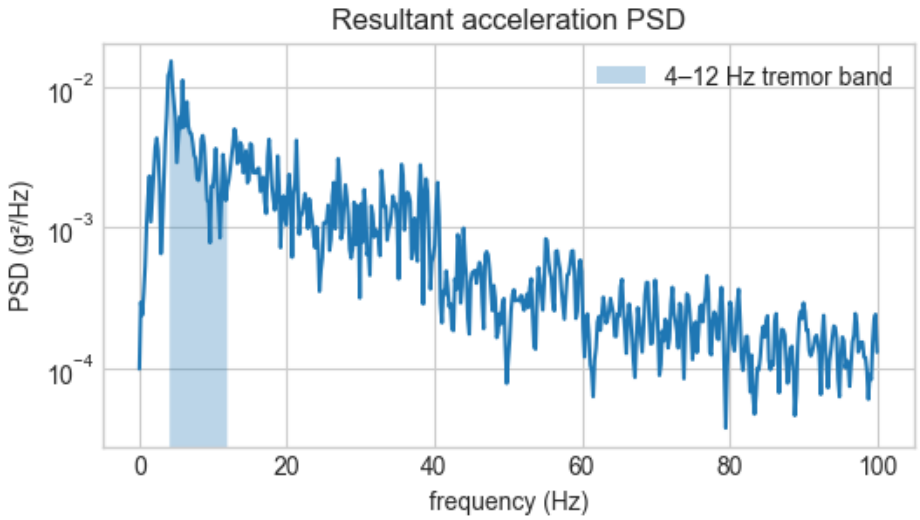

Figure 6: **Resultant acceleration PSD.** Power spectral density via Welch's method, with the 4–12Hz tremor band shaded.

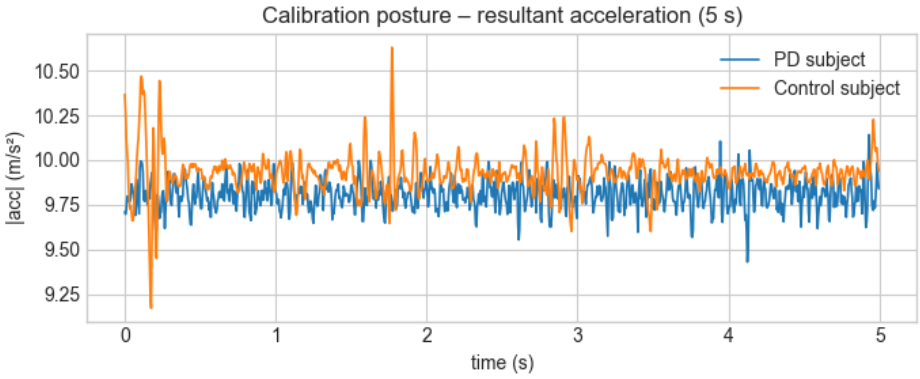

Figure 7: **Calibration posture – resultant acceleration (5s).** Magnitude of acceleration $|\mathbf{a}|$ for a PD vs. control subject holding a static pose. Both center around 9.8m/s², but the control subject exhibits larger transients.

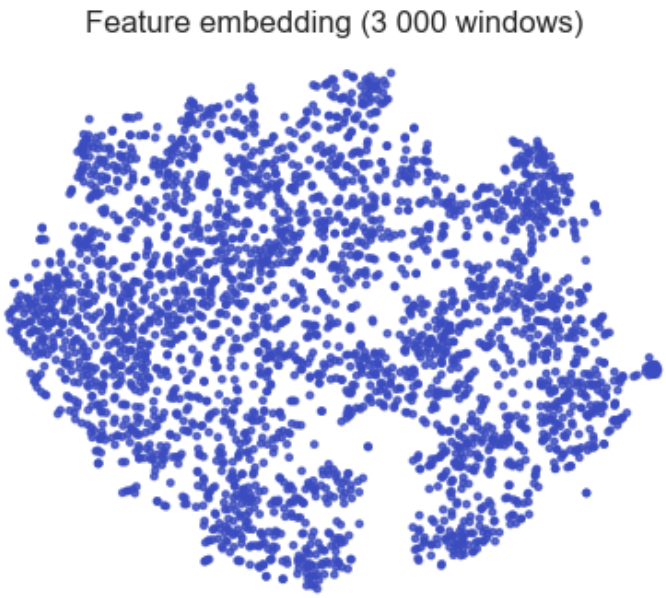

Figure 8: **t-SNE of 120-D feature vectors (3 000 windows).** Two-dimensional embedding of the time- and frequency-domain features, revealing clusters corresponding to different activity segments.

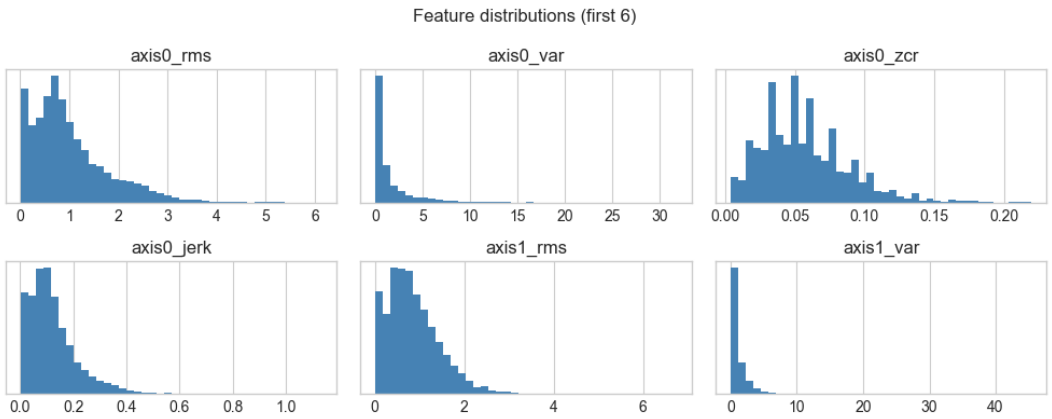

Figure 9: **Distributions of first six features.** Histograms of `axis0_rms`, `axis0_var`, `axis0_zcr`, `axis0_jerk`, `axis1_rms`, and `axis1_var`, showing skewed, long-tailed statistics prior to normalization.

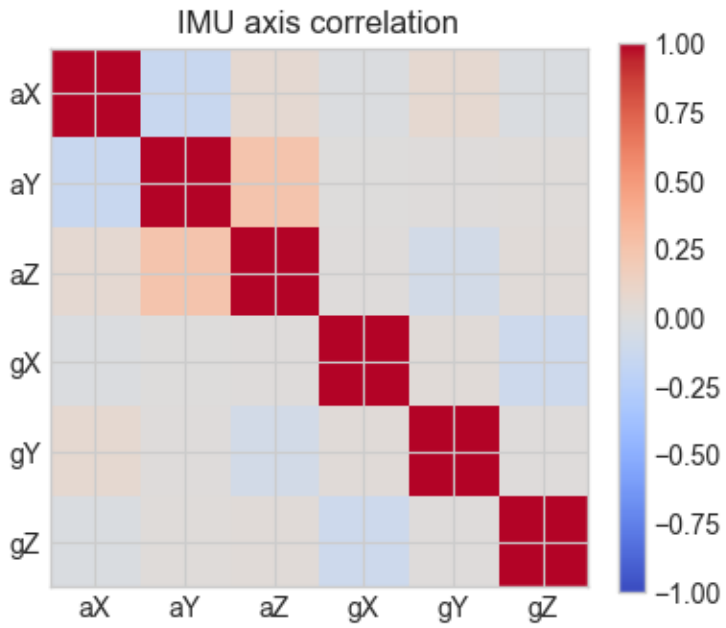

Figure 10: **IMU axis correlation heat-map.** Pearson correlation matrix for the six IMU channels $\{a_x, a_y, a_z, g_x, g_y, g_z\}$. Strong intra-modality and cross-axis couplings are visible.

B Further Results

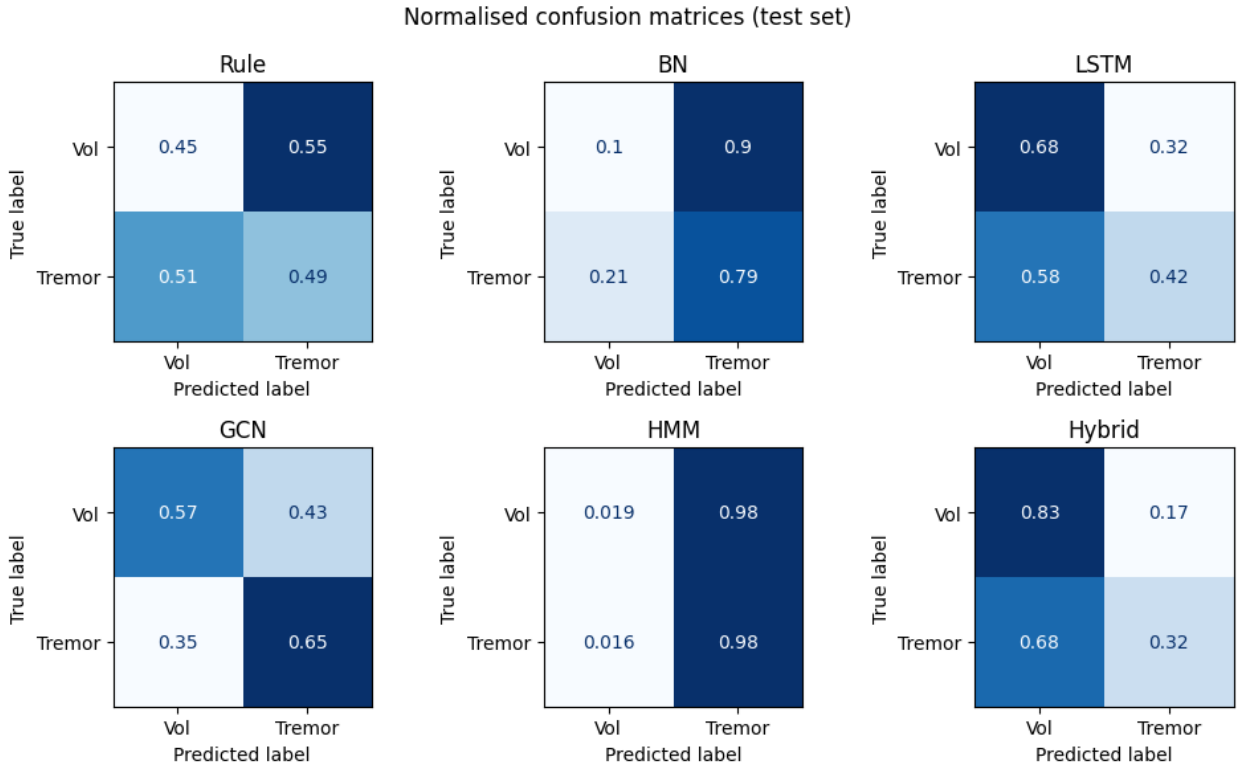

Figure 11: **Normalized confusion matrices (test set).** Each 2×2 block shows true vs. predicted labels for Rule, BN, LSTM, GCN, HMM, and Hybrid models.

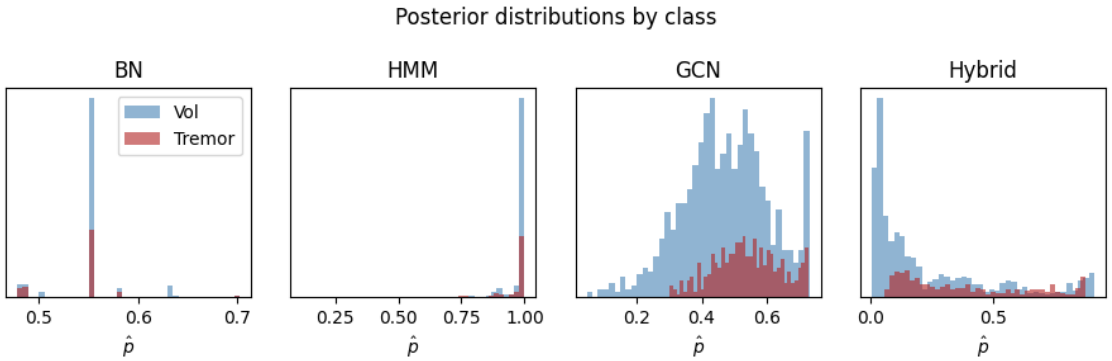

Figure 12: **Posterior probability distributions by true class.** Histograms of $\hat{p} = P(T=1 \mid \cdot)$ for voluntary (blue) vs. tremor (red) across BN, HMM, GCN, and Hybrid models on the test set.

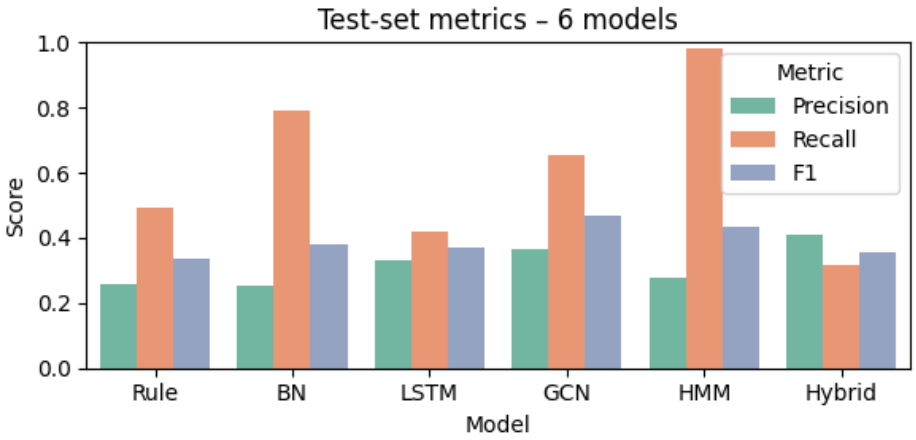

Figure 13: **Test-set metrics for six models.** Bar chart of precision, recall, and F1 on held-out subjects for Rule, BN, LSTM, GCN, HMM, and Hybrid.