Supplementary Information:
Neuromorphic Energy-Aware Learning for Adaptive Deep Brain Stimulation

Contents

1 Detailed Neuronal Dynamics and Gating Kinetics

This section provides the complete set of algebraic equations governing the voltage-dependent gating variables for the biophysical network model described in the Main Text Methods.

1.1 Subthalamic Nucleus (STN) Neurons

The STN neuron model includes current contributions from sodium (I_{Na}), potassium (I_K), A-type potassium (I_A), L-type calcium (I_L), T-type calcium (I_T), and calcium-dependent potassium (I_{CaK}) channels. The steady-state activation/inactivation functions (x_∞) and time constants (τ_x) for a gating variable x are defined below. Voltage V is in mV, and $[Ca]$ represents intracellular calcium concentration.

- **Sodium Current (I_{Na}):** m^3h kinetics

$$m_\infty(V) = \frac{1}{1+\exp\left(\frac{-(V+40)}{8}\right)}, \quad \tau_m(V) = 0.2 + \frac{3}{1+\exp\left(\frac{-(V+53)}{-0.7}\right)} \tag{S1}$$

$$h_\infty(V) = \frac{1}{1+\exp\left(\frac{V+45.5}{6.4}\right)}, \quad \tau_h(V) = \frac{24.5}{\exp\left(\frac{-(V+50)}{-15}\right)+\exp\left(\frac{-(V+50)}{16}\right)} \tag{S2}$$

- **Potassium Current (I_K):** n^4 kinetics

$$n_\infty(V) = \frac{1}{1+\exp\left(\frac{-(V+41)}{14}\right)}, \quad \tau_n(V) = \frac{11}{\exp\left(\frac{-(V+40)}{-40}\right)+\exp\left(\frac{-(V+40)}{50}\right)} \tag{S3}$$

- **A-type Potassium Current (I_A):** a^2b kinetics

$$a_\infty(V) = \frac{1}{1+\exp\left(\frac{-(V+45)}{14.7}\right)}, \quad \tau_a(V) = 1 + \frac{1}{1+\exp\left(\frac{-(V+40)}{-0.5}\right)} \tag{S4}$$

$$b_\infty(V) = \frac{1}{1+\exp\left(\frac{V+90}{7.5}\right)}, \quad \tau_b(V) = \frac{200}{\exp\left(\frac{-(V+60)}{-30}\right)+\exp\left(\frac{-(V+40)}{10}\right)} \tag{S5}$$

- **L-type Calcium Current (I_L):** $c^2d_1d_2$ kinetics

$$c_\infty(V) = \frac{1}{1+\exp\left(\frac{-(V+30.6)}{5}\right)}, \quad \tau_c(V) = 45 + \frac{10}{\exp\left(\frac{-(V+27)}{-20}\right)+\exp\left(\frac{-(V+50)}{15}\right)} \tag{S6}$$

$$d_{1\infty}(V) = \frac{1}{1+\exp\left(\frac{V+60}{7.5}\right)}, \quad \tau_{d1}(V) = 400 + \frac{500}{\exp\left(\frac{-(V+40)}{-15}\right)+\exp\left(\frac{-(V+20)}{20}\right)} \tag{S7}$$

$$d_{2\infty}([Ca]) = \frac{1}{1+\exp\left(\frac{[Ca]-0.1}{0.02}\right)}, \quad \tau_{d2}(V) = 130 \quad \text{(constant)} \tag{S8}$$

- **T-type Calcium Current (I_T):** p^2q kinetics

$$p_\infty(V) = \frac{1}{1+\exp\left(\frac{-(V+56)}{6.7}\right)}, \quad \tau_p(V) = 5 + \frac{0.33}{\exp\left(\frac{-(V+27)}{-10}\right)+\exp\left(\frac{-(V+102)}{15}\right)} \tag{S9}$$

$$q_\infty(V) = \frac{1}{1+\exp\left(\frac{V+85}{5.8}\right)}, \quad \tau_q(V) = \frac{400}{\exp\left(\frac{-(V+50)}{-15}\right)+\exp\left(\frac{-(V+50)}{16}\right)} \tag{S10}$$

- **Calcium-dependent Potassium Current (I_{CaK}):** r^2 kinetics

$$r_\infty([Ca]) = \frac{1}{1 + \exp\left(\frac{-([Ca]-0.17)}{0.08}\right)}, \quad \tau_r(V) = 2 \quad \text{(constant)} \tag{S11}$$

- **Intracellular Calcium Dynamics ($[Ca]$):**

$$\frac{d[Ca]}{dt} = -\alpha(I_L + I_T) - k_{Ca}[Ca] \tag{S12}$$

where $\alpha = 1/(ZF) = 5.18 \times 10^{-6}$ is a conversion factor describing the ion flux per unit current ($Z = 2$ is the calcium ion valence and $F = 96{,}485$ C/mol is Faraday's constant), and the calcium extrusion rate is $k_{Ca} = 2 \times 10^{-3}$ ms^{-1}.

1.2 Other Neuronal Populations

The neuronal characteristics and gating variables for the Globus Pallidus (GPe/GPi), Thalamus (Th), and Striatal (D1/D2) neurons are grounded in the established Hodgkin-Huxley style models originally developed by Rubin and Terman [1]. However, the specific equations, parameters, and Cortical Izhikevich neuron implementations used in this simulation are adopted from the comprehensive rat cortico-basal ganglia-thalamic (CBGT) network model described by Kumaravelu et al. [2].

1.3 Synaptic Dynamics

Inter-population synapses are modeled with conductance-based kinetics. The synaptic current for a given connection is:

$$I_{syn}(t) = g_{max} S(t)(V(t) - E_{rev}) \tag{S13}$$

where g_{max} is the maximal synaptic conductance, E_{rev} is the synaptic reversal potential, and the synaptic gating variable $S(t)$ follows second-order kinetics with distinct rise (τ_r) and decay (τ_d) time constants. This introduces realistic temporal latencies (e.g., N-methyl-D-aspartate (NMDA) $\tau_{decay} \approx 67$–90 ms) that the neuromorphic controller must navigate.

1.4 Multi-Taper Point-Process Spectral Analysis

For offline validation and spectral analysis, we computed power spectral densities (PSDs) directly from the discrete GPi population spike times using the multi-taper point-process method. This approach evaluates spectral variance by averaging independent estimates obtained from a set of K orthogonal data tapers evaluated at exact spike times. Evaluating the spectrum at frequency f for a set of N_{sp} spike times $\{t_j\}$ is defined as:

$$J_k(f) = \sum_{j=1}^{N_{sp}} h_k(t_j) e^{-i2\pi f t_j} - \frac{N_{sp}}{N\Delta t} H_k(f) \tag{S14}$$

$$S(f) = \frac{1}{K} \sum_{k=1}^{K} |J_k(f)|^2 \tag{S15}$$

where $h_k(t_j)$ represents the k-th orthogonal Slepian taper (discrete prolate spheroidal sequence) interpolated to the spike time t_j, and the second term of $J_k(f)$ centers the point process by subtracting the mean rate component, with $H_k(f)$ defining the Fourier transform of the taper over the total duration $N\Delta t$. We selected a time-bandwidth product of $NW = 3$, which mathematically allows for a maximum of $K = 2NW - 1 = 5$ strongly concentrated Slepian tapers. By averaging across these $K = 5$ independent tapers, this approach ensures robust, low-variance estimation of oscillatory power independent of synaptic filtering assumptions or discrete time binning artifacts.

2 Network Parameters

The intrinsic parameters for each neuron type used in the simulation are listed in Table S1.

Table S1: Maximal conductances (mS/cm^2) and reversal potentials (mV) for the biophysical model.

Parameter	**Thalamus**	**STN**	**GPe**	**GPi**	**Striatum (D1/D2)**
Leak Conductance (g_{Leak})	0.05	0.35	0.1	0.1	0.1
Leak Reversal (E_{Leak})	-70	-60	-65	-65	-67
Sodium (g_{Na})	3.0	49.0	120.0	120.0	100.0
Potassium (g_K)	5.0	57.0	30.0	30.0	80.0
A-Type Potassium (g_A)	–	5.0	–	–	–
T-type Calcium (g_T)	5.0	5.0	0.5	0.5	–
L-type Calcium (g_L)	–	15.0	–	–	–
High-Thresh Ca (g_{Ca})	–	–	0.15	0.15	–
Ca-dependent K (g_{CaK})	–	1.0	–	–	–
AHP Conductance (g_{AHP})	–	–	10.0	10.0	–

AHP: afterhyperpolarization conductance.

Table S2: Synaptic parameters used in the network simulation. τ_{rise} and τ_{decay} govern the synaptic gating variable $S(t)$. Delays (t_d) represent axonal propagation time.

Source → Target	**Receptor**	**Delay (t_d ms)**	**τ_{rise} (ms)**	**τ_{decay} (ms)**
Cortex → STN	AMPA	5.9	0.5	2.49
Cortex → STN	NMDA	5.9	2.0	90.0
Cortex → Striatum	AMPA	5.1	–	5.0*
STN → GPe	AMPA	2.0	0.4	2.5
STN → GPe	NMDA	2.0	2.0	67.0
STN → GPi	AMPA	1.5	–	5.0*
GPe → STN	GABA	4.0	0.4	7.7
GPe → GPi	GABA	3.0	–	5.0*
GPe → GPe	GABA	1.0	–	5.0*
GPi → Thalamus	GABA	5.0	–	5.0*
Striatum (D2) → GPe	GABA	5.0	–	5.0*
Striatum (D1) → GPi	GABA	4.0	–	5.0*
Thalamus → Cortex	AMPA	5.0	–	5.0*

*Modeled using single exponential decay (α-function) where τ_{decay} dominates dynamics. AMPA: α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid; NMDA: N-methyl-D-aspartate; GABA: γ-aminobutyric acid.

3 EA-DSQN: Full Implementation Details

This section expands Algorithm 1 from the Main Text, providing the complete implementation-level procedure required for exact reproducibility. The key additions relative to the main-text version are: (i) the 3-head action decoding over independent output sub-populations; (ii) explicit Q-value readout via membrane-potential accumulation; (iii) per-head Bellman targets; (iv) the sparsity regularisation loss applied to the two hidden LIF layers; and (v) gradient clipping via `clip_grad_value_(100)` rather than norm clipping.

Note on gradient clipping. The implementation uses `clip_grad_value_(100)`, which clips each individual gradient element to $[-100, 100]$, rather than the more common `clip_grad_norm_`. This choice prevents outlier gradients from destabilising early training when the SNN membrane potentials are randomly initialized.

Table S3: EA-DSQN hyperparameters used in all reported experiments.

Hyperparameter	**Symbol**	**Value**
Discount factor	γ	0.99
Soft update coefficient	τ_{tgt}	0.005
Learning rate	η	1×10^{-3}
Replay buffer size	$\lvert\mathcal{D}\rvert_{\max}$	100,000
Mini-batch size	B	128
ϵ initial	ϵ_0	0.9
ϵ final	$\epsilon_{\min}$	0.05
ϵ decay steps	—	2,000
Energy-balance weight	α	0.5
Base therapeutic reward	τ_{reward}	3000
Suppression penalty	κ	30
β-power threshold	τ_β	150 μV^2
Max stimulation energy	$E_{\max}$	250.0 μA
Sparsity weight	λ	0.0 (teacher/baseline); grid-searched in KD
Target sparsity	ρ	0.0 (teacher/baseline); grid-searched in KD
Gradient clip value	—	100
SNN time steps per obs.	T_{win}	100

Algorithm S1 EA-DSQN — Full Implementation

Require: Environment $\mathcal{E}$; policy SNN $\mathcal{S}_\theta$; target SNN $\mathcal{S}_{\theta'}$
Require: Replay buffer $\mathcal{D}$; batch size B; discount γ; soft-update coefficient τ_{tgt}
Require: Exploration ϵ; β-threshold τ_β; energy-balance weight α; suppression penalty κ
Require: Sparsity weight λ; target sparsity ρ

Initialize θ randomly; $\theta' \leftarrow \theta$
Initialize AdamW optimizer with learning rate η
for episode $= 1, \ldots, N$ **do**
  Reset $\mathcal{E}$; obtain initial state $\mathbf{s}_1$
  Reset all LIF membrane potentials in $\mathcal{S}_\theta$ and $\mathcal{S}_{\theta'}$
  for $t = 1, \ldots, T_{\max}$ **do**
   // Action selection (3-head ϵ-greedy)
    $\mathbf{U}^S, \mathbf{Z}^S \leftarrow \mathcal{S}_\theta(\mathbf{s}_t)$ ▷ Forward pass; $\mathbf{Z}^S$ are hidden spike trains
    $\hat{\mathbf{u}} \leftarrow \sum_k \mathbf{U}_k^S \in \mathbb{R}^9$ ▷ Accumulate membrane potentials over time steps
    With prob. ϵ sample random a_t; **else**
   $a_t^{(j)} \leftarrow \arg\max \hat{\mathbf{u}}_{[3j:3j+3]},\ j \in \{0,1,2\}$ ▷ j=0, 1, 2: f, δ, A
   // Environment step
    Execute a_t in $\mathcal{E}$; observe $\mathbf{s}_{t+1}$, β-power β_t, energy E_t
   // Reward (adaptive)
    if $\beta_t > \tau_\beta$ **then**
      $r_t \leftarrow -\kappa \cdot (\beta_t - \tau_\beta)$ ▷ Clinical suppression penalty
    else
      $\bar{E} \leftarrow \min(E_t / E_{\max},\ 1.0)$
      $r_t \leftarrow \tau_{\text{reward}} \cdot \big((1-\alpha) + \alpha\,(1-\bar{E})\big)$ ▷ Energy-savings gradient
    end if
    Store $(\mathbf{s}_t,\ a_t,\ r_t,\ \mathbf{s}_{t+1})$ in $\mathcal{D}$; $\mathbf{s}_t \leftarrow \mathbf{s}_{t+1}$; decay ϵ
   // Optimisation
    if $|\mathcal{D}| \geq B$ **then**
      Sample $\{(\mathbf{s}_i, a_i, r_i, \mathbf{s}_{i+1})\}_{i=1}^B \sim \mathcal{D}$
      $\mathbf{U}^S, \mathbf{Z}^S \leftarrow \mathcal{S}_\theta(\mathbf{s}_i)$ ▷ Policy forward pass
      $\hat{\mathbf{u}}_i \leftarrow \sum_k \mathbf{U}_k^S \in \mathbb{R}^{B\times 9}$
      $\mathbf{q}_i \leftarrow \hat{\mathbf{u}}_i\big[\,[0,3,6] + a_i^{(j)}\,\big] \in \mathbb{R}^{B\times 3}$ ▷ Gather per-head Q-values
      $\mathbf{U}^T, \mathbf{Z}^T \leftarrow \mathcal{S}_{\theta'}(\mathbf{s}_{i+1})$ ▷ Target forward pass
      $\hat{\mathbf{u}}_i^T \leftarrow \sum_k \mathbf{U}_k^T \in \mathbb{R}^{B\times 9}$
      $\mathbf{y}_i \leftarrow r_i + \gamma\, \max_{a'}\, \hat{\mathbf{u}}^T_{i,[3j:3j+3]}, \quad j \in \{0,1,2\}$ ▷ Per-head Bellman target $\in \mathbb{R}^{B\times 3}$
      $\mathcal{L}_{\text{TD}} \leftarrow \frac{1}{B}\sum_{i=1}^B \text{SmoothL1}(\mathbf{q}_i,\ \mathbf{y}_i)$
     // Sparsity regularisation on hidden LIF layers
      $\bar{z}^{(l)} \leftarrow \text{mean}(\mathbf{Z}^{S,(l)}), \quad l \in \{1,2\}$
      $\mathcal{L}_{\text{sp}} \leftarrow \lambda \sum_{l=1}^2 \text{KL}\big(\rho \parallel \bar{z}^{(l)}\big)$
      $\mathcal{L} \leftarrow \mathcal{L}_{\text{TD}} + \mathcal{L}_{\text{sp}}$
      $\theta \leftarrow \textsc{AdamW}(\theta,\ \nabla_\theta \mathcal{L},\ \texttt{clip\_grad\_value\_(100)})$
      $\theta' \leftarrow \tau_{\text{tgt}}\,\theta + (1-\tau_{\text{tgt}})\,\theta'$ ▷ Soft target update
    end if
  end for
end for

4 Sparsity-Constrained Knowledge Distillation Algorithm

Algorithm S2 Sparsity-Constrained Knowledge Distillation for Hardware-Constrained SNNs

Require: Trained teacher SNN $\mathcal{T}_{\theta_T}$; environment $\mathcal{E}$
Require: Temperature T_{KD}; target firing rate ρ^*; sparsity weight λ; learning rate η; epochs N
Initialize student SNN $\mathcal{S}_{\theta_S}$ with random weights
Initialize Adam optimizer with learning rate η
for $n = 1, \ldots, N$ **do**
    Reset $\mathcal{E}$; obtain raw spike train $\mathbf{X} \in \{0,1\}^{T \times C}$
    $\mathbf{x} \leftarrow \textsc{AvgPool}(\mathbf{X})$ ▷ Downsample C=80 → 16 channels

    $\mathbf{U}^T \leftarrow \mathcal{T}_{\theta_T}(\mathbf{x})$ ▷ Teacher forward pass (frozen)
    $\mathbf{q}^T \leftarrow \sum_t \mathbf{U}_t^T$ ▷ Temporal integration → Q-values

    $\mathbf{U}^S,\ \{\mathbf{s}^{(l)}\}_{l=1}^{L_h} \leftarrow \mathcal{S}_{\theta_S}(\mathbf{x})$ ▷ Student forward pass
    $\mathbf{q}^S \leftarrow \sum_t \mathbf{U}_t^S$

    $\hat{\mathbf{p}}^T \leftarrow \mathrm{softmax}\big(\mathbf{q}^T / T_{\mathrm{KD}}\big)$ ▷ Soft teacher targets
    $\hat{\mathbf{p}}^S_{\log} \leftarrow \log\text{-}\mathrm{softmax}\big(\mathbf{q}^S / T_{\mathrm{KD}}\big)$ ▷ Student log-probabilities
    $\mathcal{L}_{\mathrm{KD}} \leftarrow T_{\mathrm{KD}}^2\, D_{\mathrm{KL}}\big(\hat{\mathbf{p}}^S_{\log} \,\|\, \hat{\mathbf{p}}^T\big)$ ▷ Behaviour-matching loss

    for $l = 1, \ldots, L_h$ **do** ▷ Sparsity penalty over hidden LIF layers
        $\bar{\rho}_l \leftarrow \mathrm{mean}(\mathbf{s}^{(l)})$ ▷ Empirical mean firing rate
    end for
    $\mathcal{L}_{\mathrm{sparse}} \leftarrow \lambda \sum_{l=1}^{L_h} D_{\mathrm{KL}}^{\mathrm{Bern}}\big(\rho^* \,\|\, \bar{\rho}_l\big)$ ▷ Hardware-efficiency penalty

    $w \leftarrow \min\big(1,\ 2n/N\big)$ ▷ Warm-up: linear ramp over first 50%
    $\mathcal{L} \leftarrow \mathcal{L}_{\mathrm{KD}} + w \cdot \mathcal{L}_{\mathrm{sparse}}$ ▷ Total loss

    $\theta_S \leftarrow \textsc{Adam}(\theta_S,\ \nabla_{\theta_S}\mathcal{L},\ \text{clip\_norm=1.0})$ ▷ Update student weights
end for
return θ_S

5 Energy-Aware Stimulation Experimental Protocol

To quantify the stimulation energy savings of the energy-aware control policy, we conducted a cycling experiment comparing six conditions across alternating healthy and parkinsonian states. The simulation cycled through five 100-step blocks (10 s each): Healthy $\rightarrow$ PD $\rightarrow$ Healthy $\rightarrow$ PD $\rightarrow$ Healthy, for a total of 500 steps (50 s). Six conditions were evaluated:

1. **Unstimulated:** No stimulation applied (pathological baseline).
2. **Continuous DBS (cDBS):** Fixed stimulation at 130 Hz, 300 μA, 0.3 ms.
3. **Clinical adaptive DBS (aDBS; Dual-Threshold):** A dual-threshold heuristic modeled after the clinical protocol validated with the Medtronic Activa PC neurostimulator. The heuristic uses instantaneous GPi β-band power as its sole biomarker and applies a hysteresis band to prevent rapid stimulation cycling: stimulation activates at clinical parameters (130 Hz, 300 μA, 0.3 ms) when β-power exceeds an upper threshold ($\tau_{\text{upper}} = 160\ \mu V^2$) and deactivates when it falls below a lower threshold ($\tau_{\text{lower}} = 140\ \mu V^2$). Between the two bounds, the previous stimulation state is held.
4. **ANN Baseline (DQN):** A feedforward Deep Q-Network processing the spike matrix as a mean-pooled rate vector.
5. **RNN Baseline (GRU-DQN):** A recurrent GRU network processing the full 100-timestep observation through gated recurrent units.
6. **Adaptive SNN:** The energy-aware SNN controller, which continuously co-modulates all three stimulation parameters.

6 State-Dependent Sensory Ablation

To rule out the hypothesis that the Energy-Aware SNN controller converged to an open-loop strategy, we performed a sensory ablation study during continuous simulation. This experiment isolates the network's dynamically computed state-dependent actions from its ingrained structural baseline reactions by entirely severing the sensory input pathway.

The SNN was evaluated across seven 500-step simulation blocks, subjecting the controller to a symmetrical sequence of environmental states and sensory deprivation (Figure S1):

1. **Healthy (Block 1):** Natural physiological spiking from a healthy GPi.
2. **Parkinsonian State (Block 2):** Closed-loop control against high β-band oscillations typical of Parkinson's disease.
3. **Silent Brain (Block 3):** Sensory pathways are mathematically severed ($X = 0$), simulating total sensor failure in an underlying pathological state.
4. **Healthy Recovery (Block 4):** Sensory pathways are reconnected to healthy neural firing.
5. **Silent Brain (Block 5):** Second sensory blackout period.
6. **Parkinsonian Recovery (Block 6):** Immediate re-engagement against the pathological state.
7. **Healthy (Block 7):** Final return to the healthy physiological baseline.

During active closed-loop periods (Blocks 2 and 6), the SNN suppresses pathological GPi oscillations from a baseline of $\sim$325 μV^2 to 77.46 μV^2 and 82.12 μV^2, respectively, well below the clinical threshold of 150 μV^2. These values reflect the dynamics of this specific alternating-state protocol and are not directly comparable to the acute efficacy measurements in the main text; the purpose here is to characterise how the SNN's learned dynamics are retained and reactivated across state transitions, not to replicate a specific therapeutic outcome.

When the simulation transitions to the Silent Brain blocks (Blocks 3 and 5), sensory input vectors drop to zero. The SNN's internal membrane potentials and synaptic currents decay naturally in the absence of input spikes. Rather than outputting chaotic or over-stimulating parameters, the SNN safely defaults to 0.00 μA stimulation amplitude. Consequently, the untreated GPi network returns to its full pathological state ($\sim$320–325 μV^2). This fail-safe behavior validates that the SNN acts as a pure reactive controller that behaves in a stable, predictable manner when blinded.

The controller exhibits high resilience and rapid re-engagement following sensory deprivation. At the transition from Block 5 to Block 6, immediately after spending 500 steps in a completely blind state with 0.00 μA stimulation, the SNN is re-exposed to live pathological parkinsonian feedback. It instantly re-engages closed-loop control, rapidly ramping stimulation amplitude back to $\sim$60 μA and successfully restoring biomarker suppression. Finally, during the Healthy blocks directly following Silent blocks (Blocks 4 and 7), the mean local field potential (LFP) remains temporarily elevated ($\sim$210–220 μV^2) compared to the initial Healthy baseline. This hysteresis occurs because the biophysical GPi population, left untreated in a highly pathological state during the silent blocks, establishes a persistent resonant oscillation that takes time to naturally decay after the environment switches back to a healthy state without active pathological inputs to drive the SNN.

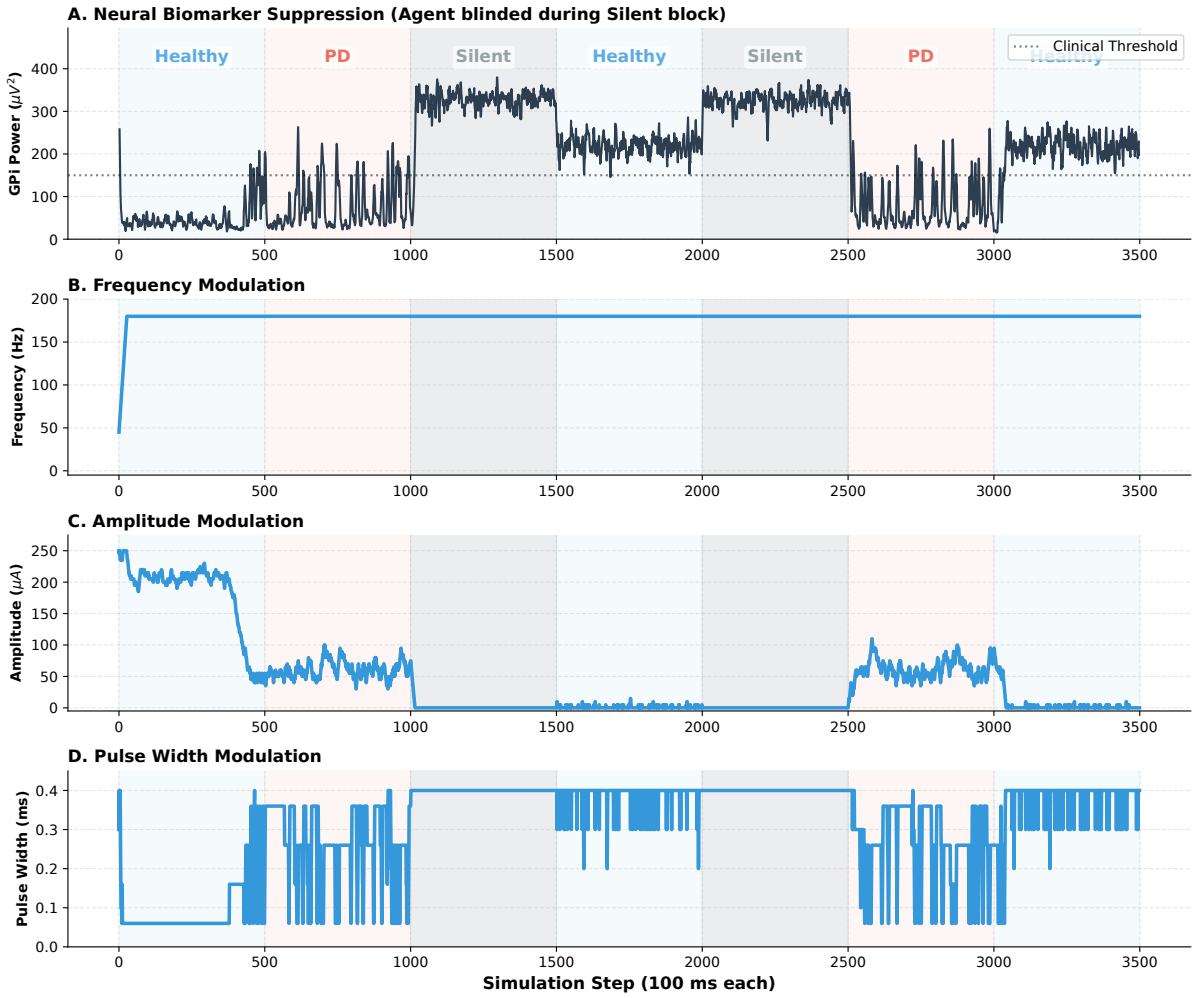

Figure S1: Sensory ablation study testing closed-loop behavior against a 3500-step sequence of Healthy, Parkinsonian, and Silent (sensory-deprived) states.

References

[1] Rubin JE, Terman D. High frequency stimulation of the subthalamic nucleus eliminates pathological thalamic rhythmicity in a computational model. *J Comput Neurosci.* 2004;16(3):211–235.

[2] Kumaravelu K, Brocker DT, Grill WM. A biophysical model of the cortex-basal ganglia-thalamus network in the 6-OHDA lesioned rat model of Parkinson's disease. *J Comput Neurosci.* 2016;40(2):207–229.