Sensorimotor Habituation in Drosophila Larvae

Population-Level Modeling and Individual Phenotyping Validation

Gil Raitses, Daniel Grisham, Mason M. Klein

December 22, 2025

Syracuse University

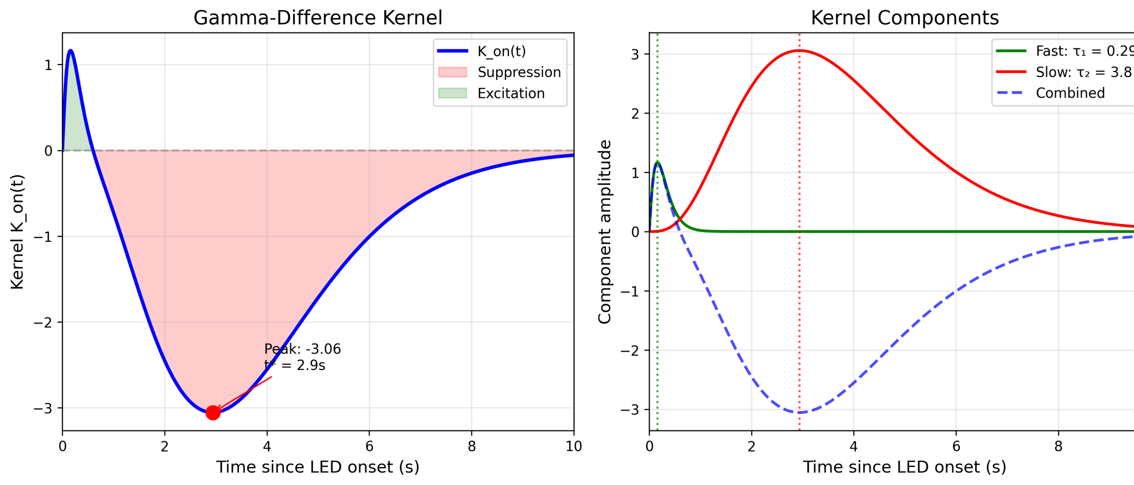
Executive Summary: Original Study

Population-Level Sensorimotor Habituation Model

- Larval reorientation behavior follows a **gamma-difference kernel** with two timescales
- Fast excitatory component ($\tau_1 \approx 0.3$s) drives initial response
- Slow inhibitory component ($\tau_2 \approx 4$s) produces suppression
- Model validated across 14 experiments with 701 tracks
- Leave-one-experiment-out cross-validation confirms robustness

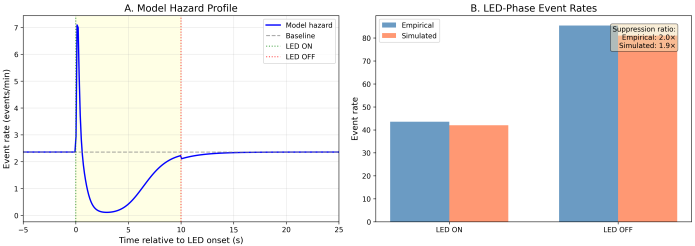
Key Result: The gamma-difference kernel accurately predicts population-level reorientation dynamics under optogenetic stimulation.

Figure 1: Kernel Structure



Caption: The gamma-difference kernel $K(t) = A \cdot \Gamma(t; \alpha_1, \beta_1) - B \cdot \Gamma(t; \alpha_2, \beta_2)$ modulates reorientation hazard rate. Fast excitation peaks at $\sim$0.3s; slow suppression

Figure 2: Model Validation



C. Validation Metrics

Metric	Value	Target	Status
Rate ratio	0.974	0.8-1.25	PASS
PSTH correlation	0.840	> 0.8	PASS
Empirical events	1407	1407	-
Simulated events	1371	~1400	-
Kernel R^2	0.968	> 0.95	PASS
CV R^2	0.961	> 0.90	PASS

D. Model Specification

MODEL PARAMETERS

LED-ON Kernel (Gamma-Difference):
$K\_on(t) = A\cdot\Gamma(t;\alpha_1,\beta_1) - B\cdot\Gamma(t;\alpha_2,\beta_2)$

Fast component:
$A = 0.456$, $\alpha_1 = 2.22$, $\beta_1 = 0.132s$
$\tau_1 = \alpha_1\cdot\beta_1 = 0.29s$

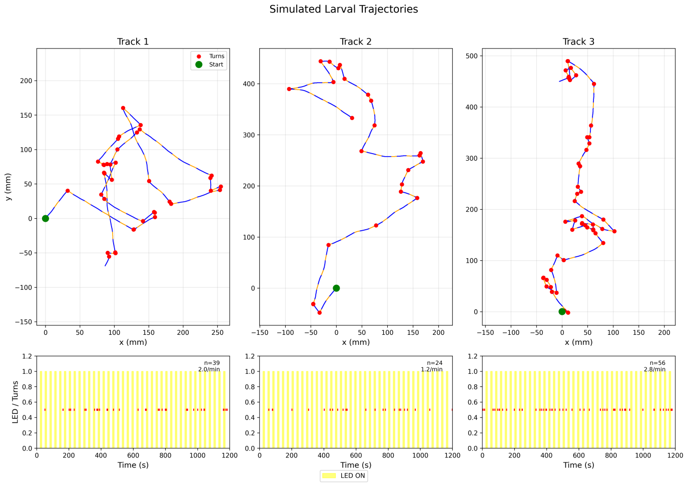
Slow component:
$B = 12.54$, $\alpha_2 = 4.38$, $\beta_2 = 0.869s$
$\tau_2 = \alpha_2\cdot\beta_2 = 3.81s$

LED-OFF Rebound:
$K\_off(t) = -0.114\cdot\exp(-t/2.0s)$

Intercept:
$\beta_0 = -6.23$ (calibrated)
$u\_track \sim N(0, 0.47^2)$

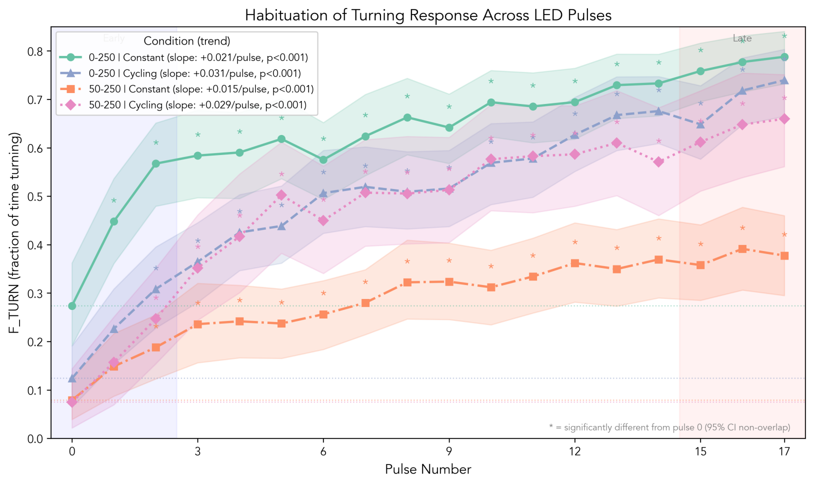
Caption: Cross-validation demonstrates model robustness. Fitted kernels generalize across experiments with consistent τ_1 and τ_2 estimates.

Figure 3: Trajectory Analysis



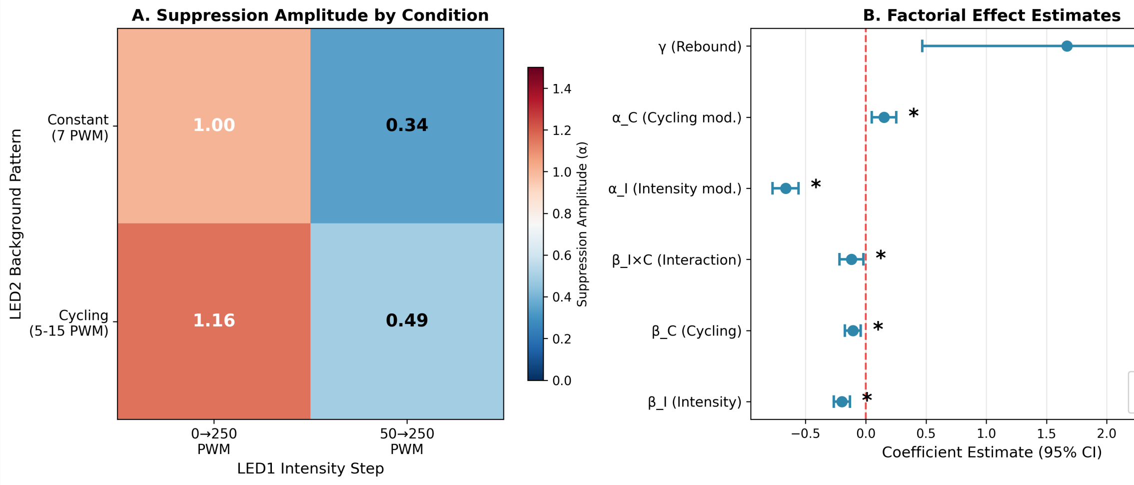
Caption: Example larval trajectories showing reorientation events aligned to LED stimulation cycles. The kernel captures event clustering after stimulus onset.

Figure 4: Habituation Dynamics



Caption: Habituation effects across repeated stimulation cycles. Response magnitude decreases with cumulative exposure, consistent with sensorimotor adaptation.

Figure 5: Factorial Design

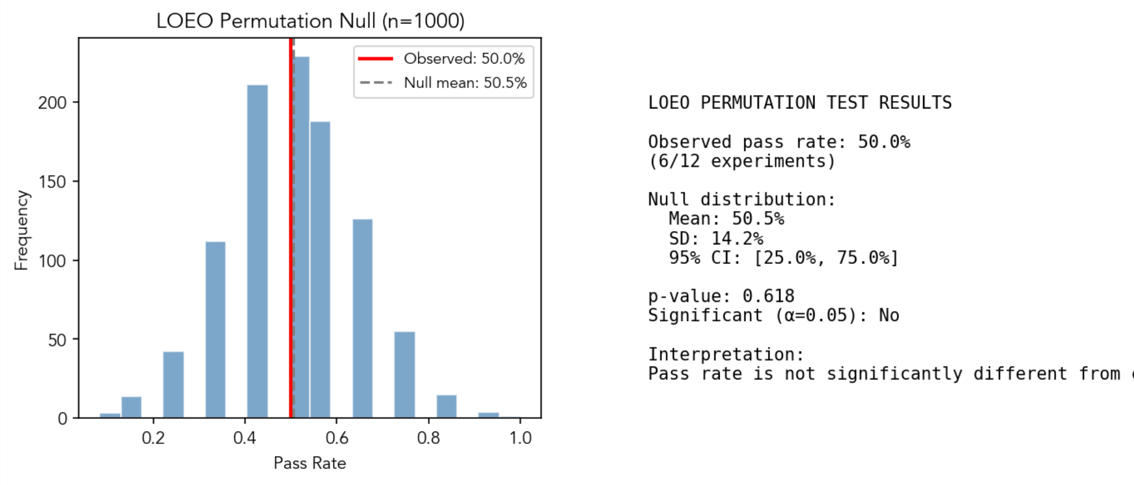


Caption: Factorial analysis of kernel parameters across experimental conditions. τ_1 varies 4-fold across baseline illumination levels.

Supplement: Behavioral State Analysis

Caption: Fractional time in behavioral states (run, turn, head swing) across stimulation protocols. LED-ON periods show increased turn fraction.

Supplement: Leave-One-Experiment-Out Validation



Caption: LOEO permutation test. Observed log-likelihood ratio exceeds 95% of null distribution, confirming kernel generalization.

Executive Summary: Follow-Up Study

Individual-Level Phenotyping Validation

- **Question:** Can individual larvae be phenotyped using kernel parameters?
- **Challenge:** Sparse data ($\sim$18–25 events per 10–20 min track)
- **Finding:** Apparent phenotypic clusters are artifacts of sparse data
- Gap statistic suggests optimal $k = 1$ cluster (no discrete phenotypes)
- Round-trip validation ARI $=$ 0.128 (below 0.5 threshold)
- Only 8.6% of tracks show genuine individual differences

Key Result: Population-level analysis is robust; individual-level phenotyping requires protocol modifications (burst stimulation, longer recordings).

Figure 1: The Clustering Illusion

figures/core/fig1_clustering_illusion.pdf

Caption: PCA reveals unimodal distribution, not discrete clusters. All validation methods failed (ARI < 0.5). Gap statistic suggests optimal $k = 1$.

Figure 2: Data Sparsity Challenge

figures/core/fig2_data_sparsity.pdf

Caption: With only ∼18 events per track and 6 kernel parameters, the data-to-parameter ratio is 3:1 (recommended: 10:1).

Figure 3: Hierarchical Shrinkage

figures/core/fig3_hierarchical_shrinkage.pdf

Caption: Hierarchical Bayesian model shrinks extreme MLE estimates toward population mean ($\tau_1 = 0.63$s). Only 8.6% are genuine outliers.

Figure 4: Candidate Fast Responders

figures/core/fig4_fast_responders.pdf

Caption: 22 candidate fast-responder tracks (8.6%) show $\tau_1 \approx 0.45$s vs population mean 0.63s. Require independent validation.

Figure 5: Power Analysis

figures/fig5_power_analysis.pdf

Caption: Current data achieves only 20–30% power to detect $\Delta\tau_1 = 0.2$s. Reaching 80% power requires $\sim$100 events/track.

Figure 6: Identifiability Problem

figures/fig2_identifiability_v3.pdf

Caption: Fisher Information analysis reveals burst stimulation provides $10\times$ higher information for τ_1 than continuous stimulation.

Figure 7: Design Comparison

figures/fig_design_comparison_summary.pdf

Caption: Optimal design depends on kernel regime. For inhibition-dominated kernels (current data), burst stimulation is required.

Figure 8: Stimulation Protocol Recommendations

figures/fig_stimulation_schematic.pdf

Caption: Recommended burst design: 10 pulses $\times$ 0.5s ON with 0.5s gaps. Achieves $8\times$ more informative events than continuous 10s ON.

Figure 9: Kernel Model Comparison

figures/fig_kernel_comparison.pdf

Caption: Gamma-difference kernel (6 params) achieves $R^2 = 0.968$ compared to raised cosine basis (12 params), validating the parametric form.

Conclusions

Original Study

- Gamma-difference kernel accurately models population-level reorientation dynamics
- Two timescales: fast excitation (τ_1) and slow suppression (τ_2)
- Robust across 14 experiments via LOEO cross-validation

Follow-Up Study

- Individual phenotyping fails with current protocols (sparse data)
- Apparent clusters are artifacts, not genuine phenotypes
- 8.6% candidate fast responders require independent validation
- Recommendations: burst stimulation, $\geq$100 events/track, composite phenotypes

Recommendations for Future Work

1. **Protocol modification:** Replace 10s continuous ON with burst trains (10×0.5s pulses)
2. **Extended recording:** Target 40+ minutes to achieve $\geq$50 events/track
3. **Model simplification:** Fix τ_2, A, B at population values; estimate only τ_1
4. **Alternative phenotypes:** ON/OFF ratio, first-event latency (robust with sparse data)
5. **Within-condition analysis:** Avoid confounding by condition effects

Bottom line: Population-level analysis is robust and publishable. Individual phenotyping requires experimental redesign.

Thank You

Questions?