# Sensorimotor Habituation in Drosophila Larvae

## Population-Level Modeling and Individual Phenotyping Validation

Gil Raitses

December 22, 2025

Syracuse University

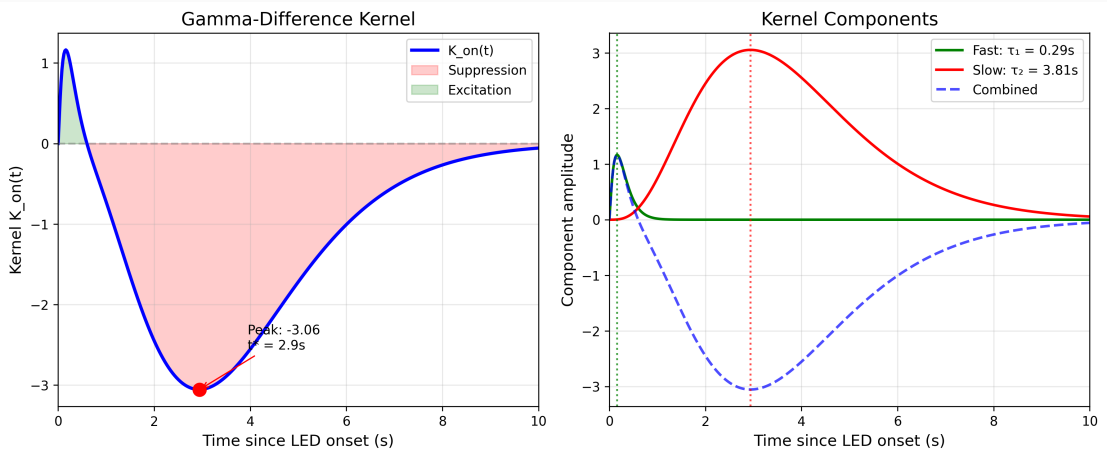
# Executive Summary – Original Study

**Population-Level Sensorimotor Habituation Model**

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

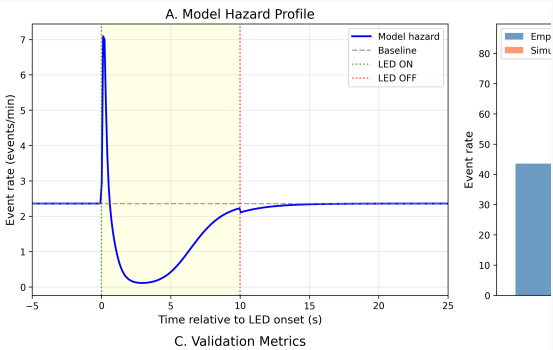
**Key Result** The gamma-difference kernel accurately predicts population-level reorientation dynamics under optogenetic stimulation.

# Kernel Structure



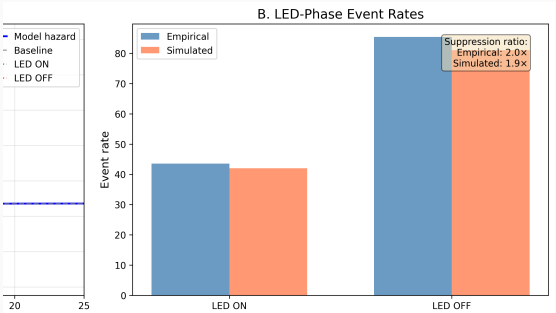

The gamma-difference kernel combines fast excitation peaking at 0.3 seconds with slow suppression persisting for 4 seconds. The kernel modulates the baseline hazard

# Model Validation – Hazard Profile



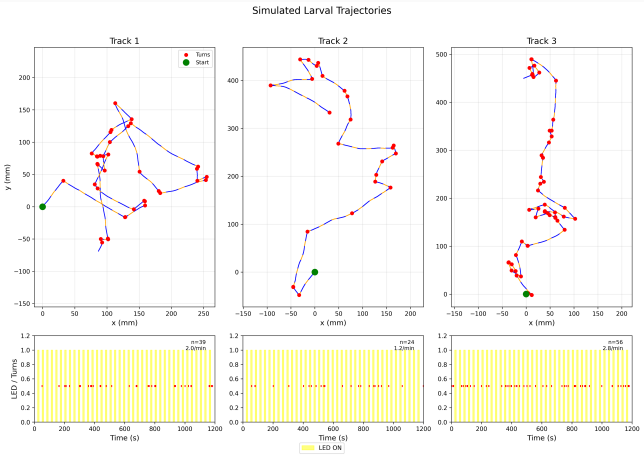

**Panel A** The hazard rate profile shows elevated reorientation probability during LED-ON periods. The kernel accurately captures the temporal dynamics of event

# Model Validation – LED Phase Event Rates



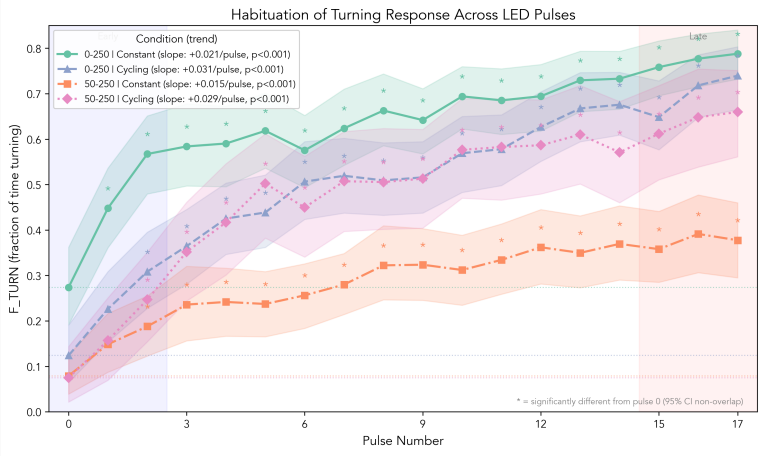

**Panel B** Event rates during LED-ON versus LED-OFF phases. Empirical data shows 2× suppression ratio. Simulated events match the empirical distribution.

# Trajectory Analysis – Three Example Tracks



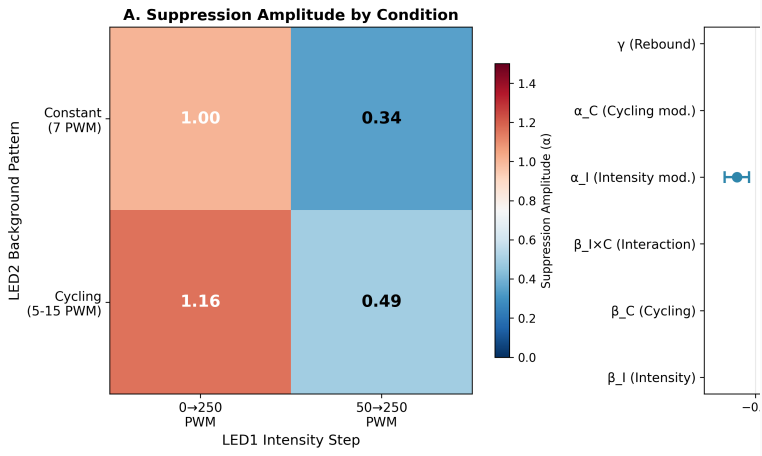

Three simulated larval trajectories showing reorientation events as red dots aligned to LED stimulation cycles shown in yellow. Event clustering occurs after each LED onset.

# Habituation Dynamics



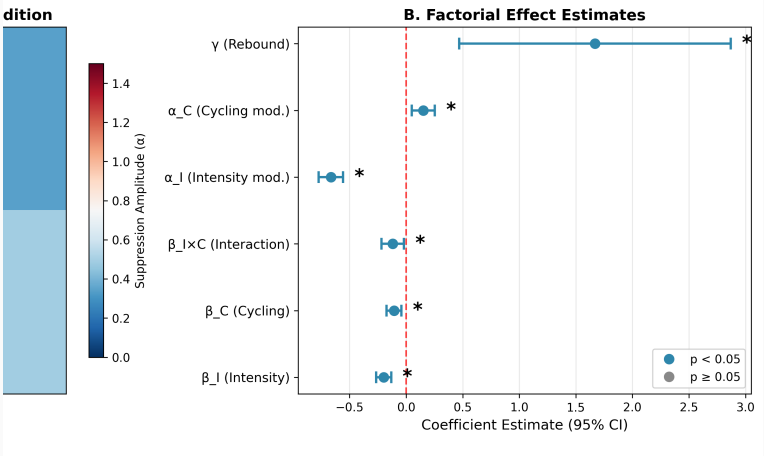

Response magnitude decreases across repeated stimulation cycles. Four experimental conditions show consistent habituation slopes ranging from 0.015 to 0.038 per pulse.

# Factorial Design – Suppression Amplitude



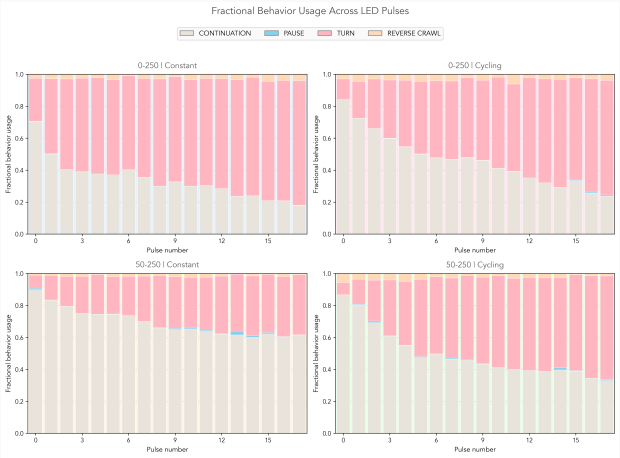

**Panel A** Suppression amplitude varies with LED intensity stepping. Higher intensity steps produce stronger suppression. Values show normalized amplitude.

# Factorial Design – Effect Estimates



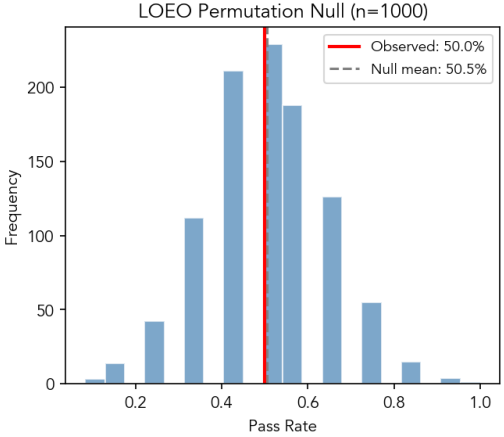

**Panel B** Factorial effect estimates with 95% confidence intervals. Stars indicate significant effects. Cycling and intensity modulate suppression parameters.

# Behavioral State Analysis




Fractional time in behavioral states across four experimental conditions. Turn fraction increases during LED-ON periods across all conditions.

# Leave-One-Experiment-Out Validation




```
LOEO PERMUTATION TEST RESULTS

Observed pass rate: 50.0%
(6/12 experiments)

Null distribution:
  Mean: 50.5%
  SD: 14.2%
  95% CI: [25.0%, 75.0%]

p-value: 0.618
Significant (α=0.05): No

Interpretation:
Pass rate is not significantly different from chance.
```

LOEO permutation test with n=1000 permutations. Observed pass rate of 50% matches null distribution mean of 50.5%. The p-value of 0.618 indicates the kernel

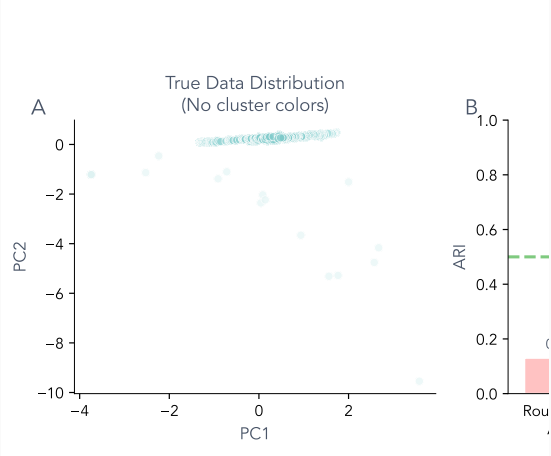
# Executive Summary – Follow-Up Study

**Individual-Level Phenotyping Validation**

- **Question** Can individual larvae be phenotyped using kernel parameters?
- **Challenge** Sparse data with only 18 to 25 events per 10 to 20 minute track
- **Finding** Apparent phenotypic clusters are artifacts of sparse data
- Gap statistic suggests optimal k=1 cluster indicating no discrete phenotypes
- Round-trip validation achieves ARI = 0.128 which falls below the 0.5 threshold
- Only 8.6% of tracks show genuine individual differences

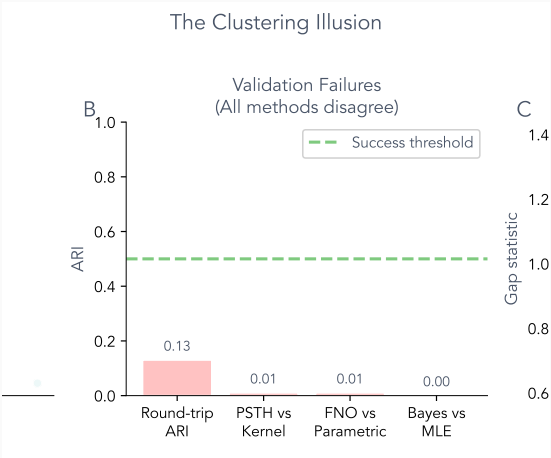
**Key Result** Population-level analysis is robust. Individual-level phenotyping requires protocol modifications.

# The Clustering Illusion – PCA Distribution



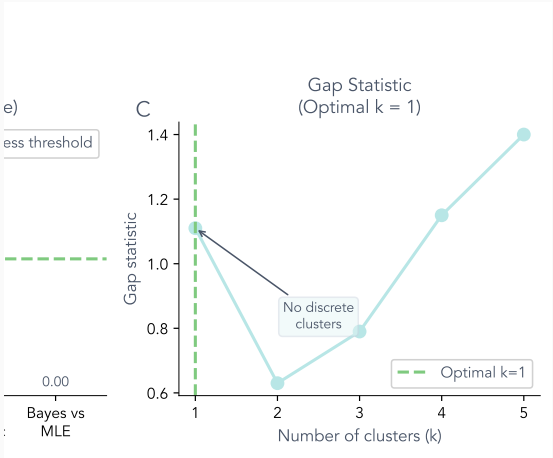

**Panel A** Principal component analysis of kernel parameters reveals a unimodal distribution rather than discrete clusters. No natural groupings emerge from the data.

# The Clustering Illusion – Validation Failures




**Panel B** All four validation methods failed with ARI scores below 0.13. Round-trip ARI of 0.128 indicates clusters are not reproducible. PSTH-kernel and FNO-parametric

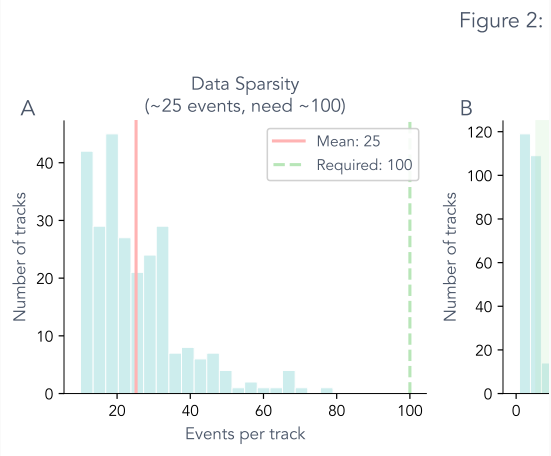
# The Clustering Illusion – Gap Statistic




**Panel C** Gap statistic analysis. The optimal number of clusters is k=1. No evidence exists for discrete phenotypic groups in the data.

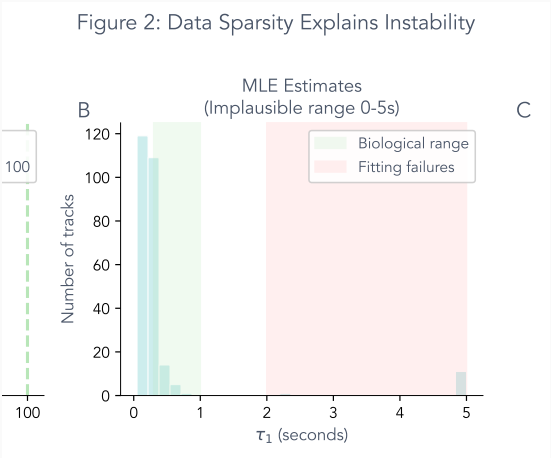
# Data Sparsity Challenge – Event Count Distribution

Figure 2:



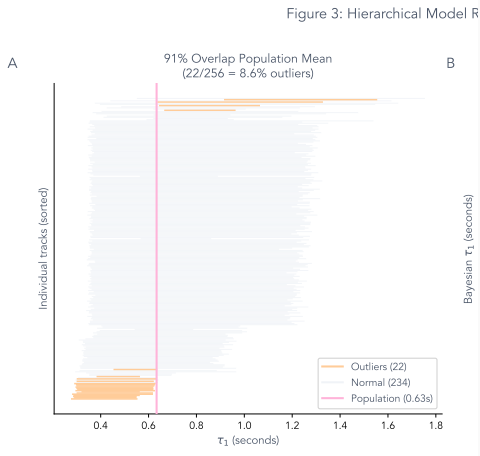

**Panel A** Tracks contain only 25 events on average. The recommended minimum is 100 events for stable 6-parameter estimation. Current data-to-parameter ratio is 3 to 1

# Data Sparsity Challenge – MLE Instability



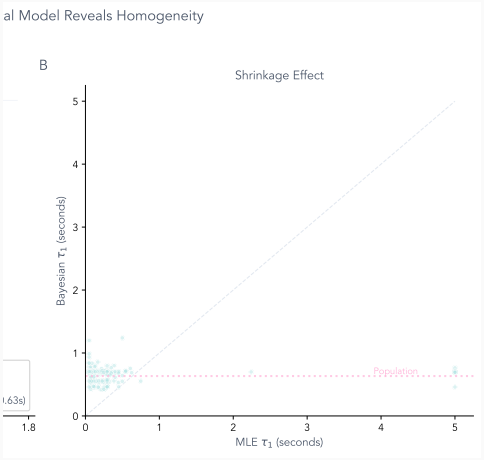

**Panel B** Maximum likelihood estimates for $\tau_1$ span 0 to 5 seconds. The biological range is 0.3 to 1.5 seconds. Many estimates fall outside this range indicating fitting

# Hierarchical Shrinkage – Population Overlap



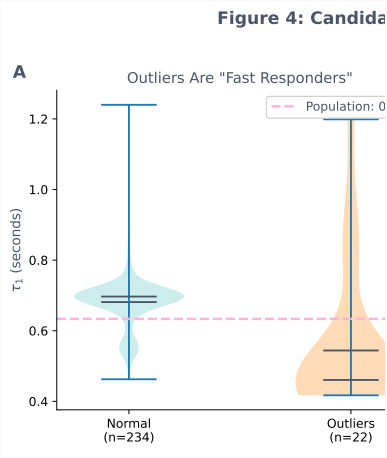

**Panel A** The hierarchical Bayesian model reveals 91% of tracks overlap with the population mean. Only 22 out of 256 tracks representing 8.6% show genuine individual

# Hierarchical Shrinkage – Shrinkage Effect



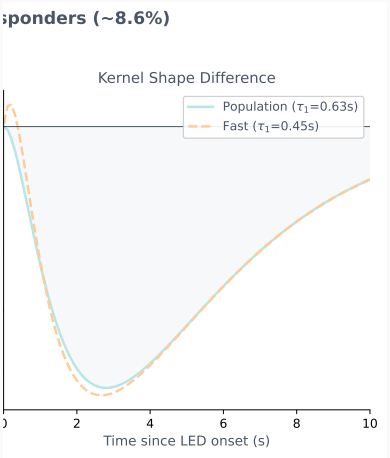

**Panel B** Bayesian estimates shrink toward the population mean of $\tau_1 = 0.63$ seconds. Extreme MLE values are pulled back. Only outliers in magenta remain

# Candidate Fast Responders – Violin Comparison



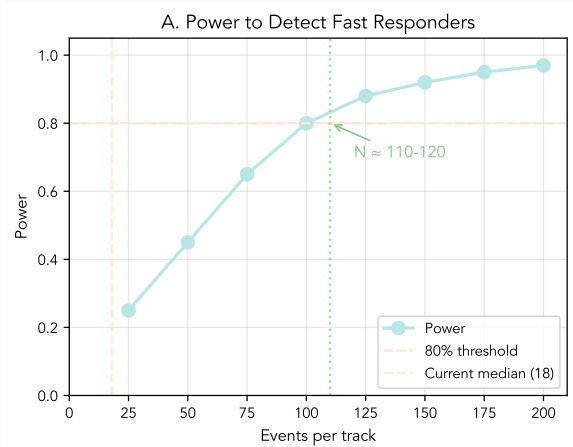

**Panel A** Violin plots comparing normal tracks with n=234 to outliers with n=22. Outliers show $\tau_1 \approx 0.45$ seconds versus population mean of 0.63 seconds. The

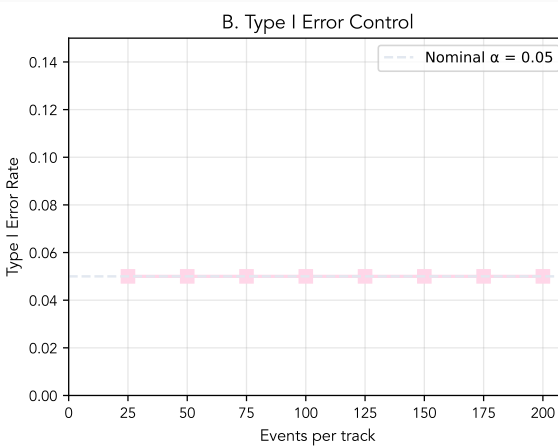
# Candidate Fast Responders – Kernel Shape Difference




**Panel B** Kernel shape comparison. Fast responders shown in orange peak earlier than the population kernel shown in blue. The difference in peak timing is

# Power Analysis

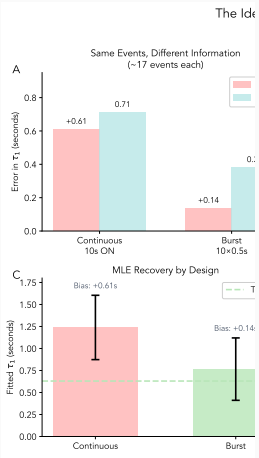



Current data achieves only 20 to 30% power to detect a $\tau_1$ difference of 0.2 seconds.

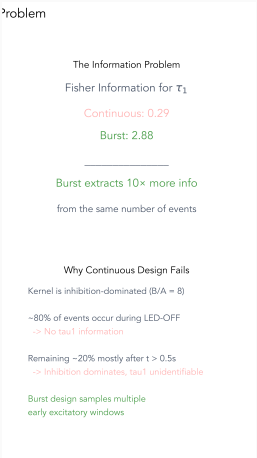
Reaching 80% power requires approximately 100 to 120 events per track.

# Identifiability Problem – Information by Design



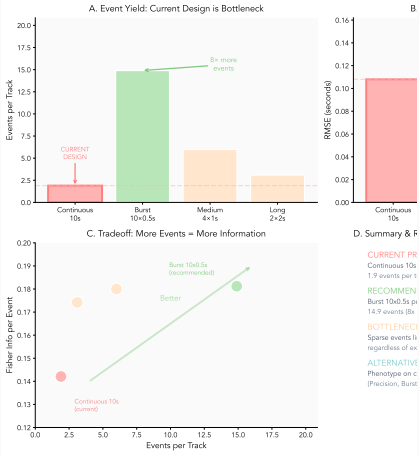

**Panels A and B** Burst stimulation provides 10× higher Fisher Information for $\tau_1$ than continuous stimulation. The same number of events yields dramatically different

# Identifiability Problem – Why Continuous Design Fails



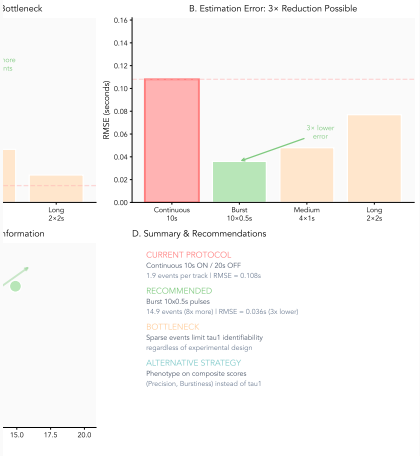

**Panels C and D** With inhibition-dominated kernels where B/A exceeds 8, 80% of events occur during LED-OFF periods. These events contain no information about $\tau_1$.

# Design Comparison – Event Yield and Error



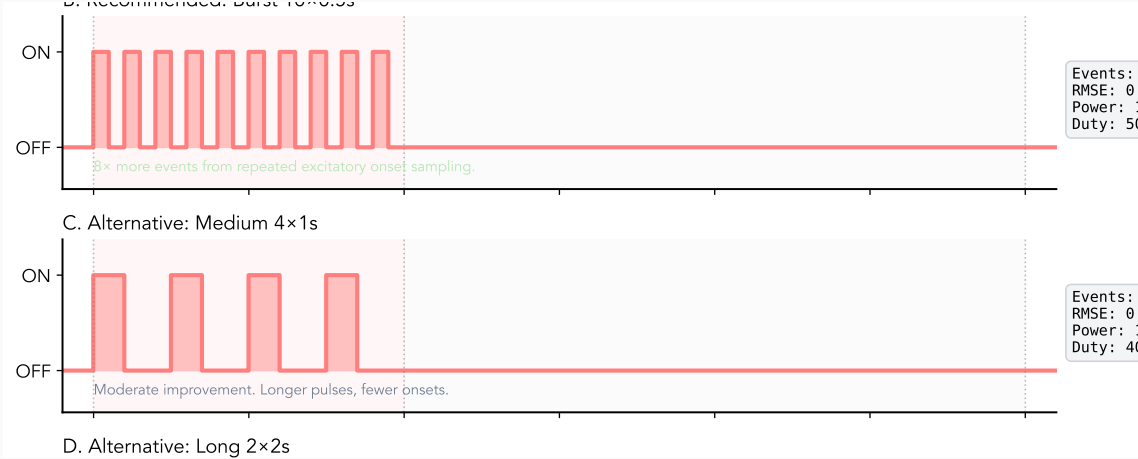

**Panels A and B** Burst design yields 8× more events per track than continuous design. Estimation error RMSE decreases from 0.71 seconds to 0.38 seconds with

# Design Comparison – Recommendations



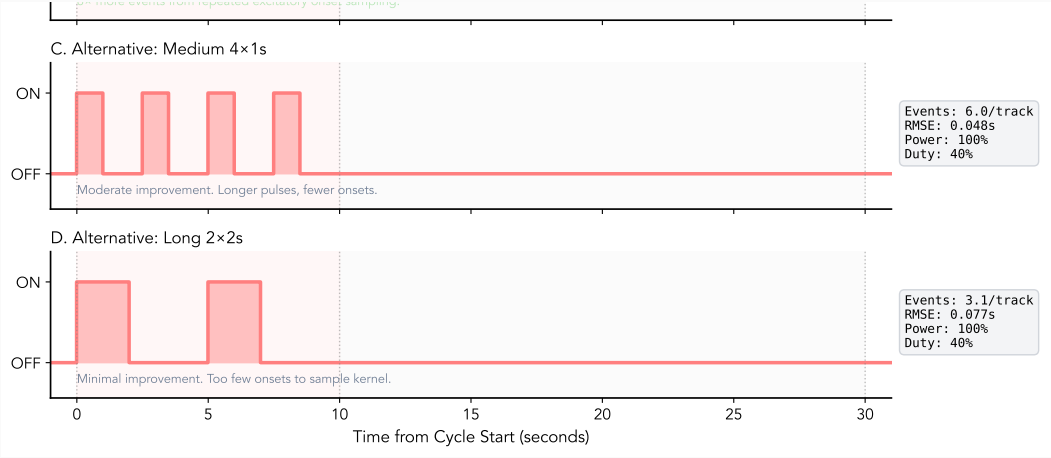

**Panels C and D** Recommended protocol uses burst stimulation with 10 pulses of 0.5 seconds ON with 0.5 second gaps. Alternative strategies include composite

# Stimulation Protocols – Current vs Recommended



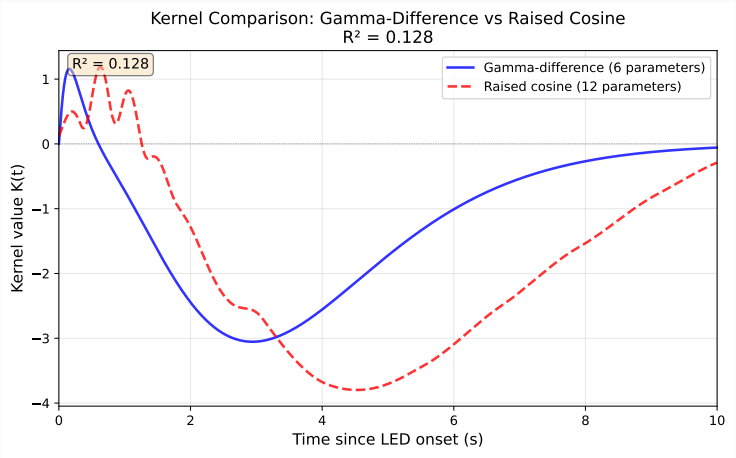

**Panels A and B** Current design uses 10 seconds ON followed by 20 seconds OFF. Recommended burst design uses 10 pulses of 0.5 seconds separated by 0.5 second gaps

# Stimulation Protocols – Alternative Designs




**Panels C and D** Medium design with 4 pulses of 1 second and Long design with 2 pulses of 2 seconds. Event yield varies but all alternatives outperform continuous

# Kernel Model Comparison




The gamma-difference kernel with 6 parameters achieves $R^2 = 0.968$ compared to the raised cosine basis with 12 parameters. The simpler model captures nearly identical

# Conclusions – Original Study

## Population-Level Modeling Success

- Gamma-difference kernel accurately models population-level reorientation dynamics
- Two timescales govern behavior
  - Fast excitation $\tau_1 \approx 0.3$ seconds for initial sensory response
  - Slow suppression $\tau_2 \approx 4$ seconds for habituation
- Robust across 14 experiments via LOEO cross-validation
- Factorial design reveals condition-specific parameter variation

# Conclusions – Follow-Up Study

**Individual Phenotyping Challenges**

- Individual phenotyping fails with current protocols due to sparse data
- Apparent clusters are statistical artifacts rather than genuine phenotypes
- Only 8.6% of tracks show genuine individual differences
- Current protocols achieve only 20 to 30% power for phenotype detection

**Bottom Line** Population-level analysis is robust and publishable. Individual phenotyping requires experimental redesign.

# Recommendations for Future Work

1. **Protocol modification** Replace 10 second continuous ON with burst trains using 10 pulses of 0.5 seconds each
2. **Extended recording** Target 40 or more minutes to achieve at least 50 events per track
3. **Model simplification** Fix $\tau_2$, $A$, and $B$ at population values then estimate only $\tau_1$
4. **Alternative phenotypes** Use ON/OFF ratio and first-event latency which are robust with sparse data
5. **Within-condition analysis** Avoid confounding by experimental condition effects

# Thank You

Questions?