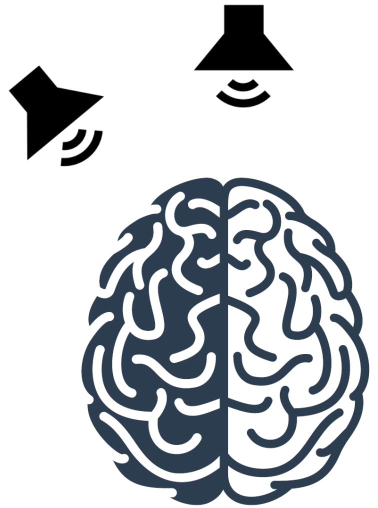
Detection of Code Juggling with Spike Counts Data Analysis

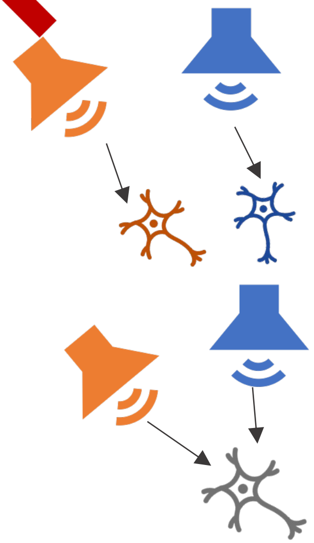
Yunran Chen and Surya T. Tokdar
Department of Statistical Science, Duke University, Durham, NC 27708, USA

Motivation

How does brain preserve information about multiple simultaneous items ?



- Designated neurons for each signal



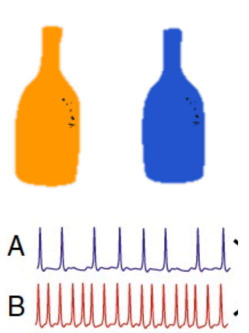
- Small receptive field (perceptual sensitivity)

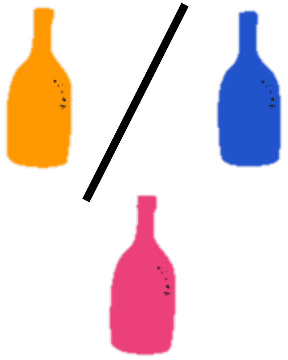
- Receptive fields too large
  - Visual (Alonso and Chen, 2009; Keliris et al., 2019)
  - Auditory (Groh et al., 2003; Werner-Reiss and Groh, 2008; Bulkin and Groh, 2011)

A single neuron will be exposed to multiple simultaneous stimuli !

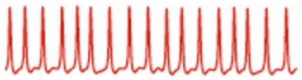
- Dynamics in presentation in a neural level
- Can a single neuron preserve info from both  and 

Motivation: potential dynamics





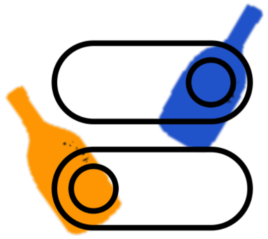
- Always encode A (or B)



- 1st order stochasticity
- Poisson distribution

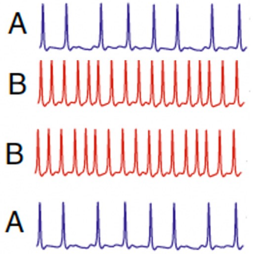
(Ventura et al., 2002; Kass et al., 2005)

- Treat it as a new stimulus

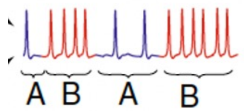


- Switch between A and B
- **Code Juggling**
- **2nd order stochasticity: constituent stimuli**
- Greater efficiency on information preserving
- Overdispersed + bounded by A & B

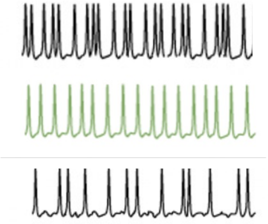
- Across trials
(Caruso et al., 2018)



- Within a trial
(Glynn et al., 2021)



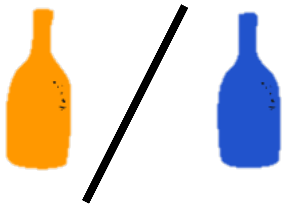
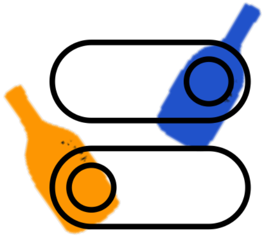
- Stochasticity not related A nor B
- Overdispersed
- Not related to A nor B



Higher trial-to-trial variability
(Semedo et al., 2019)

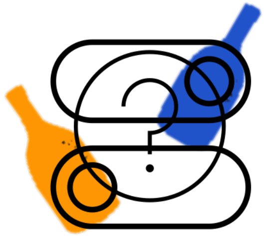
Research question:

How to model these biologically meaningful encoding schemes?

Poisson Mixtures

How to select the most likely scheme and quantify the uncertainty?

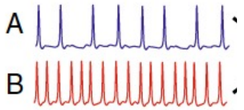


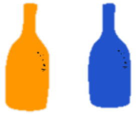
Hypothesis Testing
(Bayes Factor)

Statistical Analysis Framework

Assumptions:

- Two-stimuli cases





- Single stimulus spike counts follow Poisson distribution

$$P^{\mathrm{A}} = \mathrm{Poi}(\mu^{\mathrm{A}}),\ \ P^{\mathrm{B}} = \mathrm{Poi}(\mu^{\mathrm{B}}), \qquad \mu^{\mathrm{A}}, \mu^{\mathrm{B}} > 0.$$

Data: $(Y_j^e : j \in \{1, \ldots, n_e\}, e \in \{\mathrm{A}, \mathrm{B}, \mathrm{AB}\})$

Hypotheses:

$$\text{FIXED: } P^{\mathrm{AB}} = \mathrm{Poi}(\mu), \qquad \mu \in (0, \infty),$$

(whole-trial)

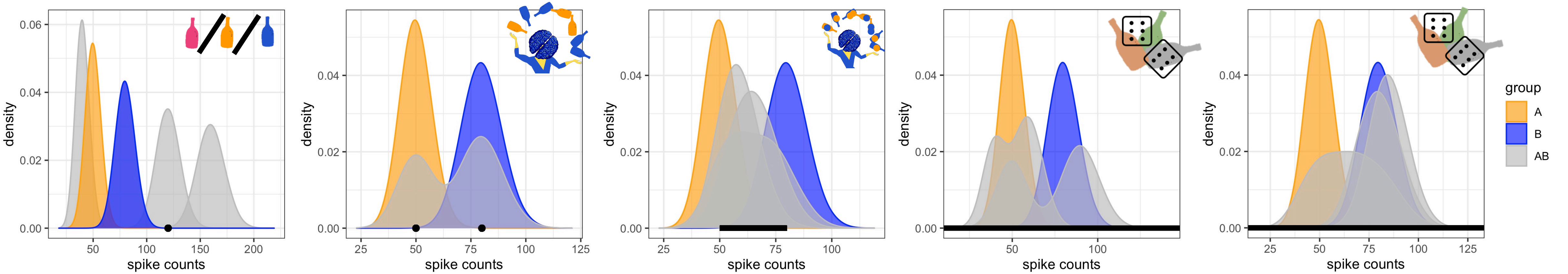
$$\text{MIXTURE: } P^{\mathrm{AB}} = \rho \mathrm{Poi}(\mu^{\mathrm{A}}) + (1-\rho)\mathrm{Poi}(\mu^{\mathrm{B}}), \qquad \rho \in (0,1),$$

(sub-trial)

$$\text{INTERMEDIATE: } P^{\mathrm{AB}} = \int \mathrm{Poi}(r\mu^{\mathrm{A}} + (1-r)\mu^{\mathrm{B}}) f(r) dr, \qquad f \text{ is a density on } (0,1),$$

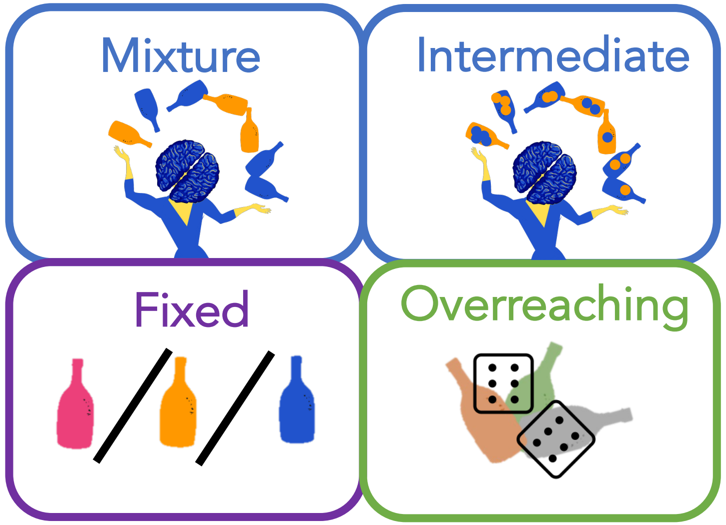
$$\text{OVERREACHING: } P^{\mathrm{AB}} = \int \mathrm{Poi}(m) g(m) dm, \qquad g \text{ is a density on } (0, \infty),$$

Probability distributions (support + continuity):



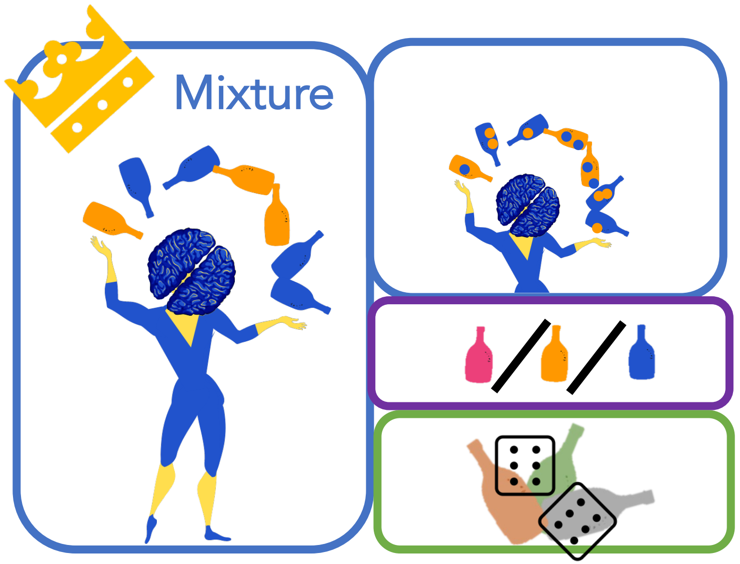
Hypothesis Testing in a Bayesian framework

Prior + Data → Posterior



Spike counts

$$(Y_j^e : j \in \{1, \ldots, n_e\}, e \in \{A, B, AB\})$$



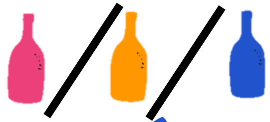
Challenges: nonparametric density estimation and fair competition

Predictive recursion marginal likelihood method (PRML)

(Newton et al., 1998; Newton, 2002; Tokdar et al., 2009; Martin and Tokdar, 2011)

Statistical Analysis Framework

Hypotheses:



FIXED: $P^{AB} = \mathrm{Poi}(\mu), \quad \mu \in (0, \infty),$

MIXTURE: $P^{AB} = \rho \mathrm{Poi}(\mu^{A}) + (1-\rho)\mathrm{Poi}(\mu^{B}), \quad \rho \in (0,1),$

INTERMEDIATE: $P^{AB} = \int \mathrm{Poi}(r\mu^{A} + (1-r)\mu^{B}) f(r)dr, \quad f$ is a density on $(0,1)$,

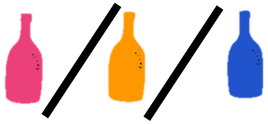
OVERREACHING: $P^{AB} = \int \mathrm{Poi}(m)g(m)dm, \quad g$ is a density on $(0,\infty)$,

Straightforward

unknown

unknown

Statistical Analysis Framework

Hypotheses:  FIXED: $P^{\mathrm{AB}} = \mathrm{Poi}(\mu),$ $\mu \in (0, \infty),$

Marginal likelihood: $p(Y^{\mathrm{A}}, Y^{\mathrm{B}}, Y^{\mathrm{AB}}) = \cancel{p(Y^{\mathrm{A}})p(Y^{\mathrm{B}})}\,p(Y^{\mathrm{AB}} | Y^{\mathrm{A}}, Y^{\mathrm{B}})$

$$p(Y^{\mathrm{AB}} | Y^{\mathrm{A}}, Y^{\mathrm{B}}) = p(Y^{\mathrm{AB}}) = \int p(Y^{\mathrm{AB}} | \mu) \pi(\mu) d\mu$$

Set a prior and can obtain closed form

Statistical Analysis Framework

OVERREACHING: $P^{AB} = \int \text{Poi}(m) g(m) dm,$ g is a density on $(0, \infty)$,

Marginal likelihood: $p(Y^{AB} | Y^{A}, Y^{B}) = p(Y^{AB})$

Predictive Recursion (Smith and Markov, 1978; Newton et al., 1998; Newton, 2002)

Data $Y_1, \ldots, Y_n$ from mixture distribution

known unknown

$$m_f(y) = \int \kappa(y \mid u) f(u) d\nu(u)$$

For $i \in \{1, \ldots, n\}$

$$f_i(u) = (1 - w_i) f_{i-1}(u) + w_i \frac{\kappa(Y_i \mid u) f_{i-1}(u)}{m_{i-1}(Y_i)}, \; u \in \mathcal{U},$$

$$m_{i-1}(y) = \int \kappa(y \mid u) f_{i-1}(u) d\nu(u), \quad y \in \mathcal{Y}$$

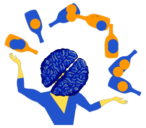
Theoretical guarantee (Tokdar et al., 2009) :

f_n m_n converge to the truth asymptotically (provided $\sum_{i=1}^{\infty} w_i = \infty$ $\sum_{i=1}^{\infty} w_i^2 < \infty$)

Bayesian inferential paradigm interpretation (Tokdar et al., 2009)

Statistical Analysis Framework

$$\text{MIXTURE: } P^{\mathrm{AB}} = \rho \mathrm{Poi}(\mu^{\mathrm{A}}) + (1-\rho)\mathrm{Poi}(\mu^{\mathrm{B}}), \qquad \rho \in (0,1),$$



$$\text{INTERMEDIATE: } P^{\mathrm{AB}} = \int \mathrm{Poi}(r\mu^{\mathrm{A}} + (1-r)\mu^{\mathrm{B}}) f(r) dr, \qquad f \text{ is a density on } (0,1),$$

Marginal likelihood:

$$p(Y^{\mathrm{AB}} | Y^{\mathrm{A}}, Y^{\mathrm{B}}) = \int p(Y^{\mathrm{AB}} | \theta) p(\theta | Y^{\mathrm{A}}, Y^{\mathrm{B}}) d\theta \qquad \theta = (\mu^{\mathrm{A}}, \mu^{\mathrm{B}})$$

First-stage prior:

$$\pi_0(\mu^{\mathrm{A}}, \mu^{\mathrm{B}}) = 1/\sqrt{\mu^A \mu^{\mathrm{B}}},$$

Second-stage prior:

$$\pi(\mu^{\mathrm{A}}, \mu^{\mathrm{B}}) = \mathrm{Gam}(\mu^{\mathrm{A}} \mid 0.5 + \textstyle\sum_{j=1}^{n_{\mathrm{A}}} Y_j^{\mathrm{A}}, n_{\mathrm{A}}) \times \mathrm{Gam}(\mu^{\mathrm{B}} \mid 0.5 + \textstyle\sum_{j=1}^{n_{\mathrm{B}}} Y_j^{\mathrm{B}}, n_{\mathrm{B}})$$

$$m_f(y) = \int \kappa_\theta(y \mid u) f(u) d\nu(u)$$

PRML score

(Martin and Tokdar, 2011)

Predictive Recursion Marginal Likelihood

Data $Y_1, \ldots, Y_n$ from mixture distribution

$i \in \{1, \ldots, n\}$

Predictive Recursion

(Newton et al., 1998; Newton, 2002)

$$m_f(y) = \int \kappa_\theta(y \mid u) f(u) d\nu(u)$$

$$f_i(u) = (1 - w_i) f_{i-1}(u) + w_i \frac{\kappa_\theta(Y_i \mid u) f_{i-1}(u)}{m_{i-1}(Y_i)}, \; u \in \mathcal{U},$$

$$m_{i-1,\theta}(y) = \int \kappa_\theta(y \mid u) f_{i-1}(u) d\nu(u), \quad y \in \mathcal{Y}$$

$\kappa = \kappa_\theta$

PRML score

(Martin and Tokdar, 2011)

$$L_n(\theta) := \prod_{i=1}^{n} m_{i-1,\theta}(Y_i)$$

- PRML score estimate true marginal density: $p(Y_1, \ldots, Y_n \mid \theta) = \prod_{i=1}^{n} p(Y_i \mid Y_1, \ldots, Y_{i-1}, \theta)$
- Asymptotic consistency: $n^{-1} \log\{L_n(\theta)/L_n(\theta^\star)\} \to -\inf_{f \in \mathbb{F}} d_{\mathrm{KL}}(m^\star, m_{f,\theta})$.
- Interpretable: expectation filtration approximation to a fully Bayesian estimation $w_i = (1 + a)^{-1}, \; a > 0$

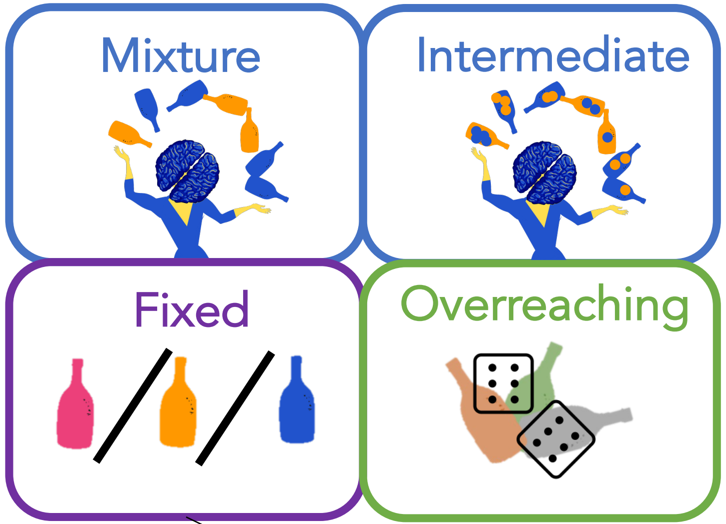
Laplace Approximation:

$$I(M_h) := \int_{\Theta_h} L_{n,h}(\theta_h) \pi_h(\theta_h) d\theta_h, \quad h \in H,$$

$$\hat{I}(M_h) = L_{n,h}(\hat{\theta}_h)(2\pi)^{-d_h/2} |\Sigma_h|^{1/2},$$

Fair competition: choice of prior

Prior + Data → Posterior



Spike counts

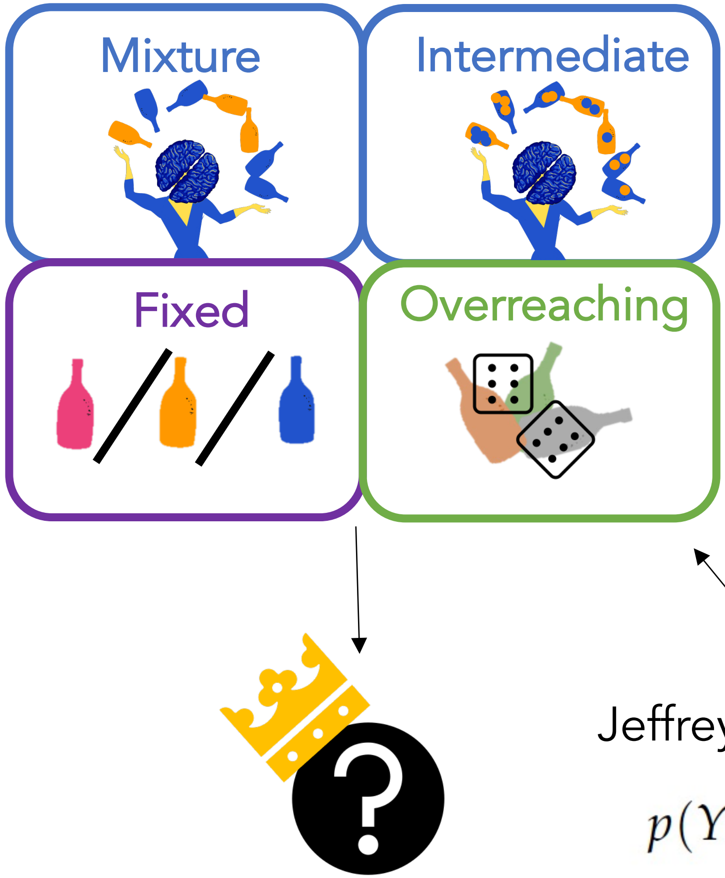
$$(Y_j^e : j \in \{1, \ldots, n_e\}, e \in \{\mathrm{A}, \mathrm{B}, \mathrm{AB}\})$$

$$p_h(Y) = \frac{I(M_h)\, p_{0,h}}{\sum_{h' \in H} I(M_{h'})\, p_{0,h'}}$$

- Sensitive to the choice of prior
  - Bounded uniform (same as overreaching)
  - Jeffreys' prior (improper) - - > tune the constant
  - Jeffreys' prior (intrinsic Bayes factor) (Berger & Pericchi, 1996)

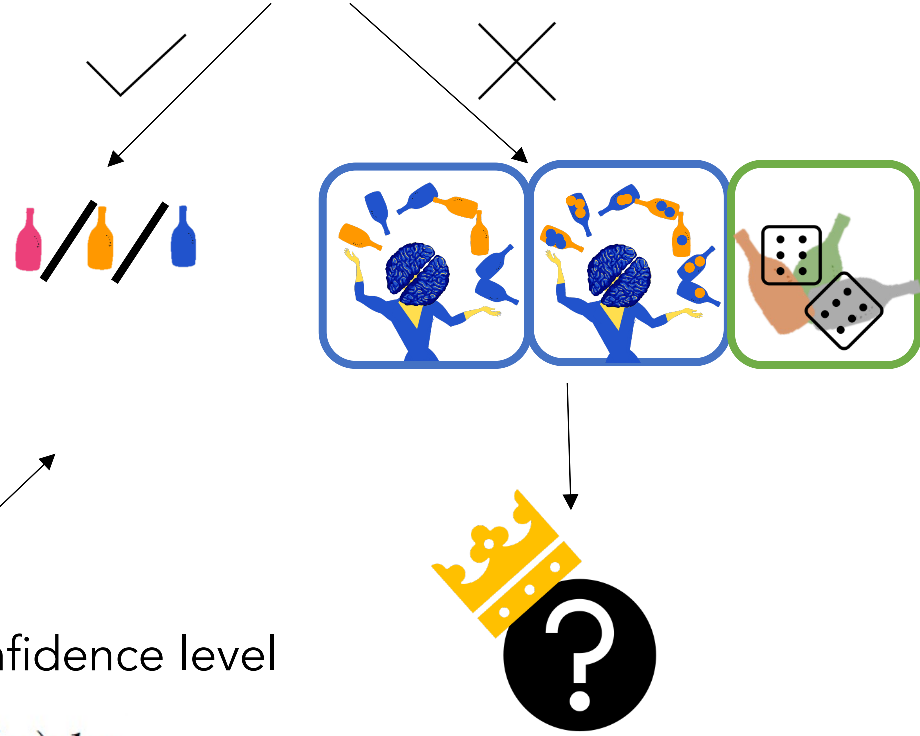
Two potential frameworks

Full Bayesian



Two-stage procedure

Poisson variance test → Is it Fixed?
(Brown and Zhao, 2002)



Jeffreys’ prior: tune constant to match the confidence level

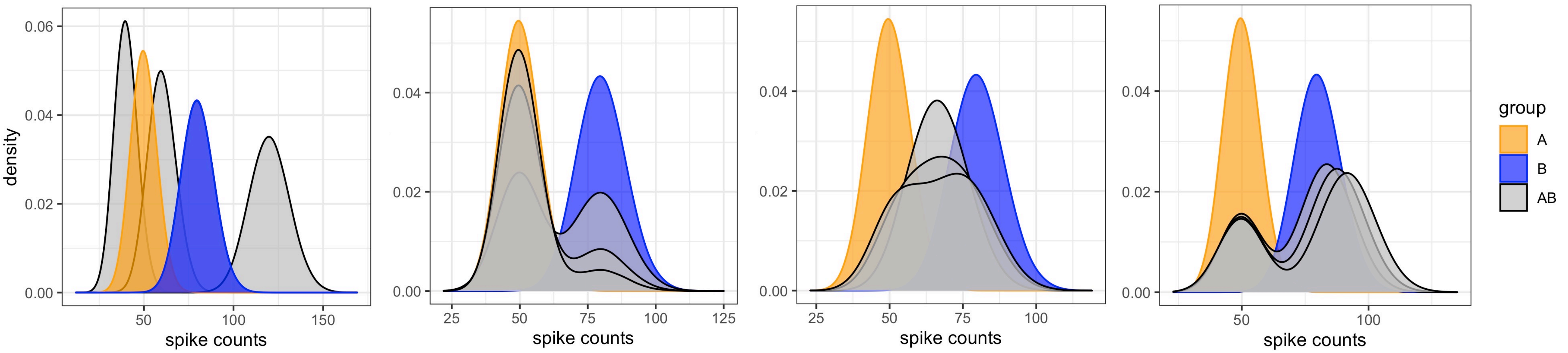
$$p(Y^{\mathrm{AB}}|Y^{\mathrm{A}}, Y^{\mathrm{B}}) = p(Y^{\mathrm{AB}}) = \int p(Y^{\mathrm{AB}}|\mu)\pi(\mu)d\mu$$

$$\pi(\mu) \propto \mu^{-1/2}$$

Performance Assessment

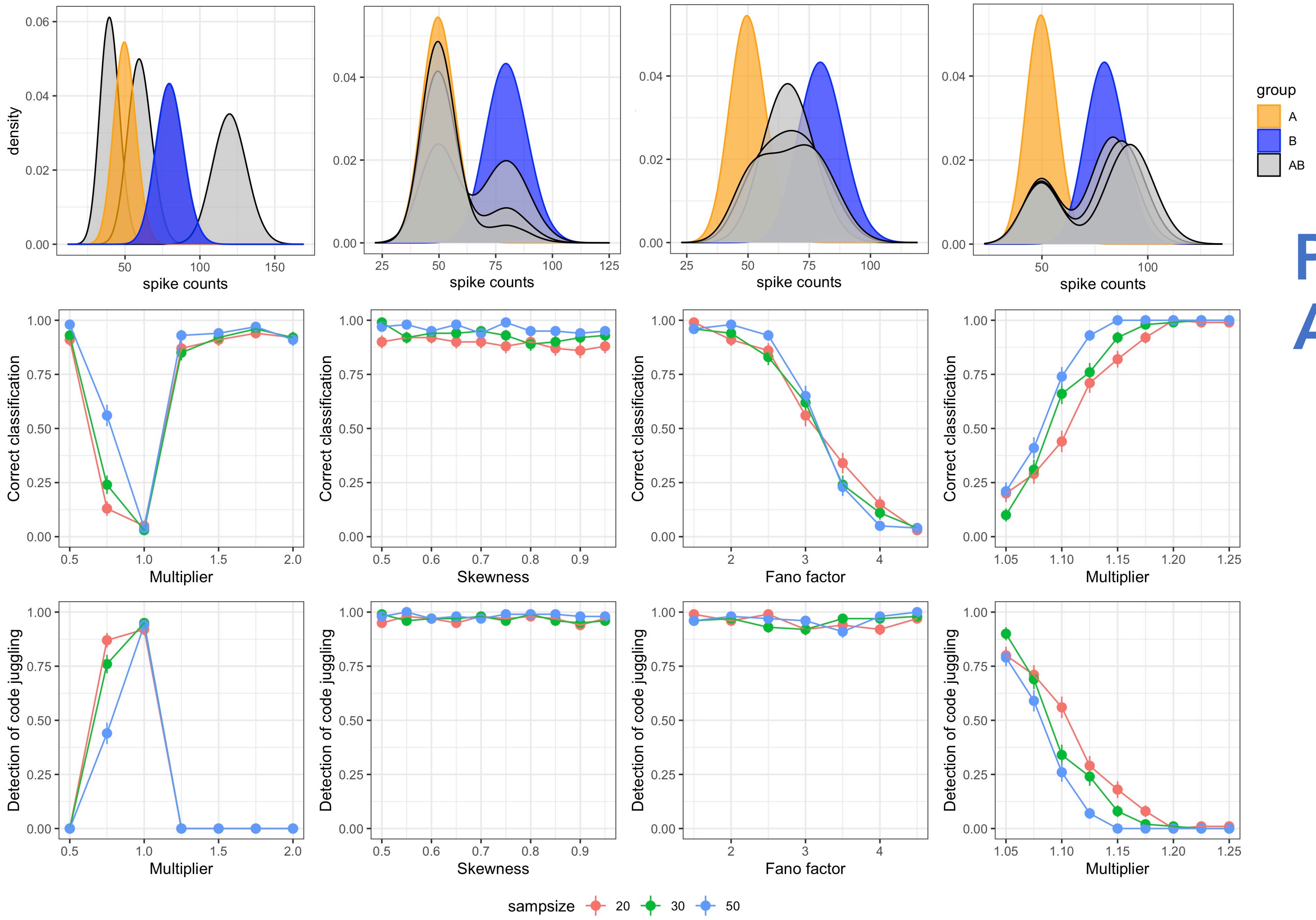
100 experiment sets: $P^{A} = \text{Poi}(50)$ $P^{B} = \text{Poi}(80)$ $n_A = n_B = n_{AB}$

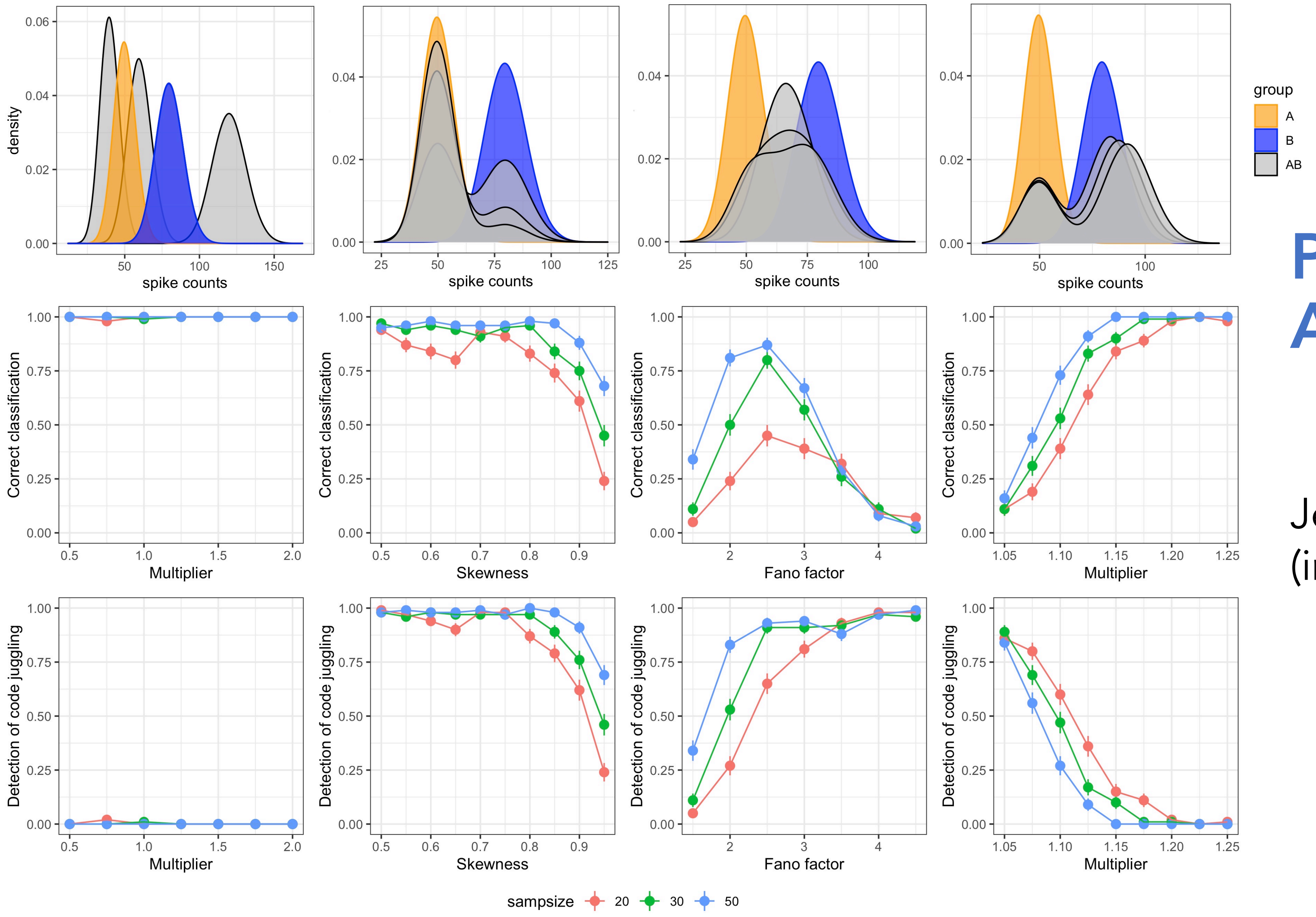
FIXED:	$P^{AB} = \text{Poi}(80m)$	Multiplier
MIXTURE:	$P^{AB} = m\text{Poi}(50) + (1-m)\text{Poi}(80)$	Skewness
INTERMEDIATE:	$P^{AB} = \int \text{Poi}(r50 + (1-r)80)\text{Beta}(r; a, b)dr$	Fano factor
OVERREACHING:	$P^{AB} = 1/3\text{Poi}(50) + 2/3\text{Poi}(80m)$	Multiplier



Performance Assessment

Bounded Uniform



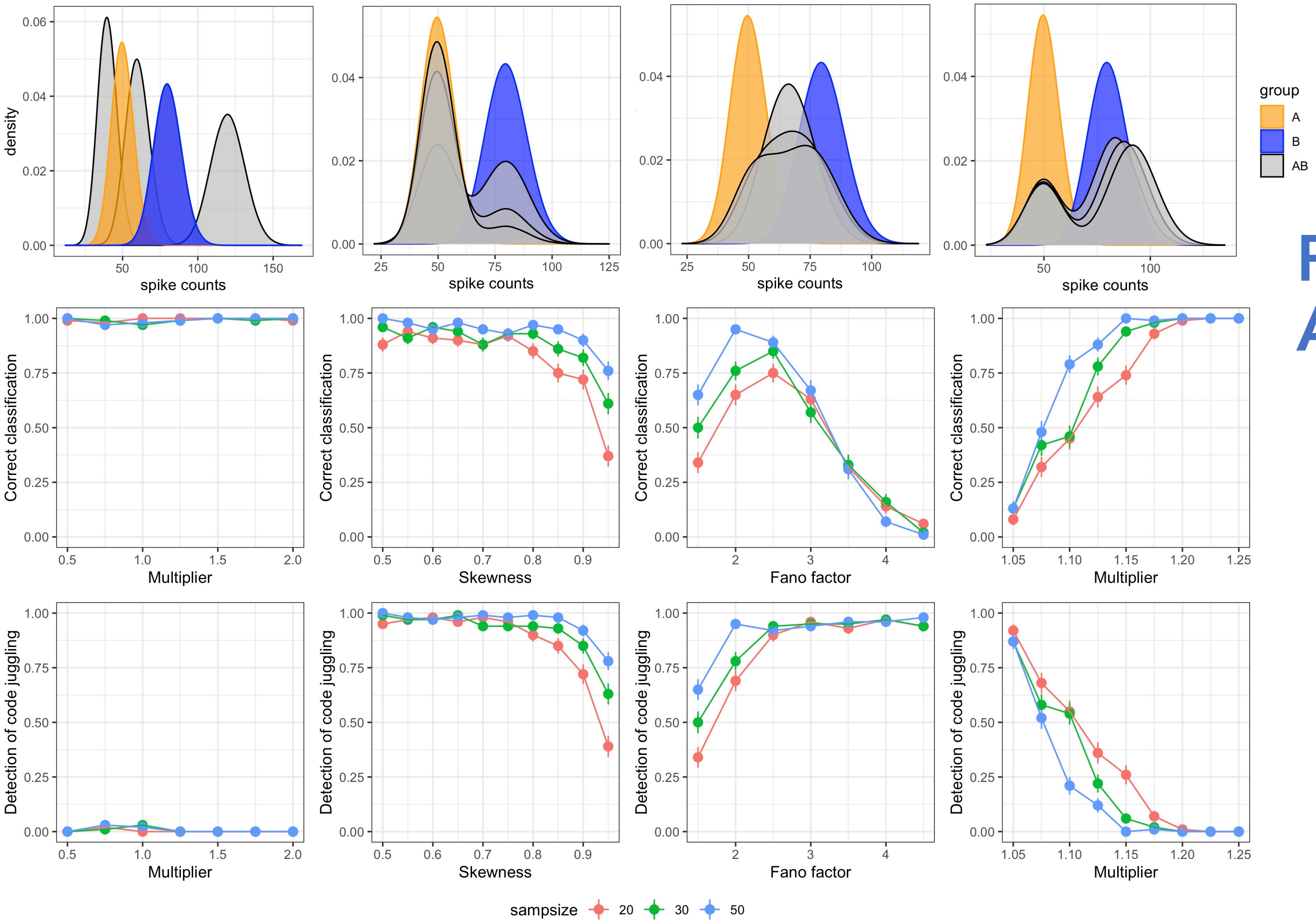


Performance Assessment

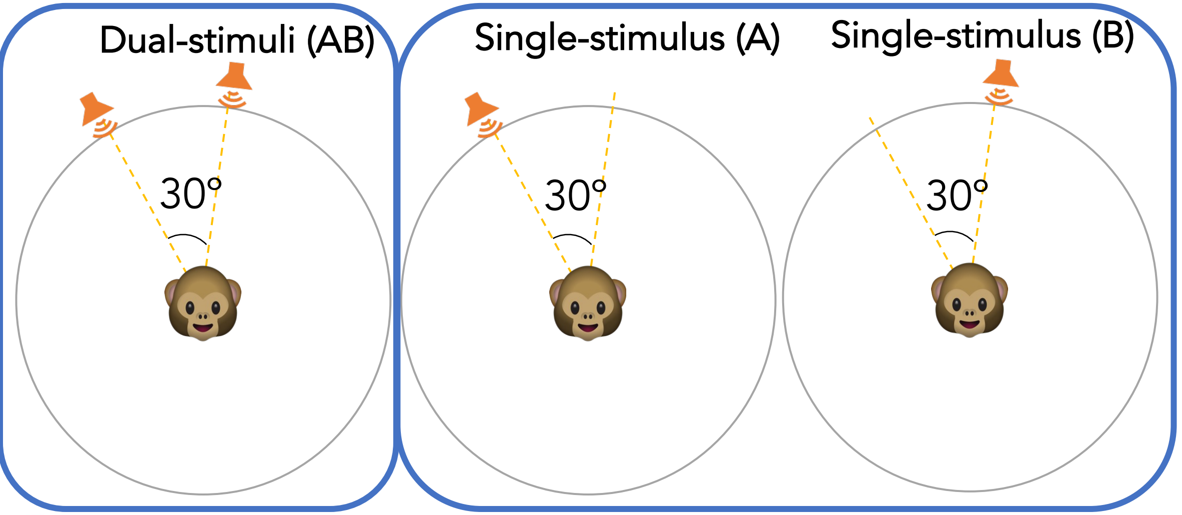
Jeffreys’ prior
(intrinsic Bayes factor)

Performance Assessment

Jeffreys' prior (w/ constant 1)



Application in IC data

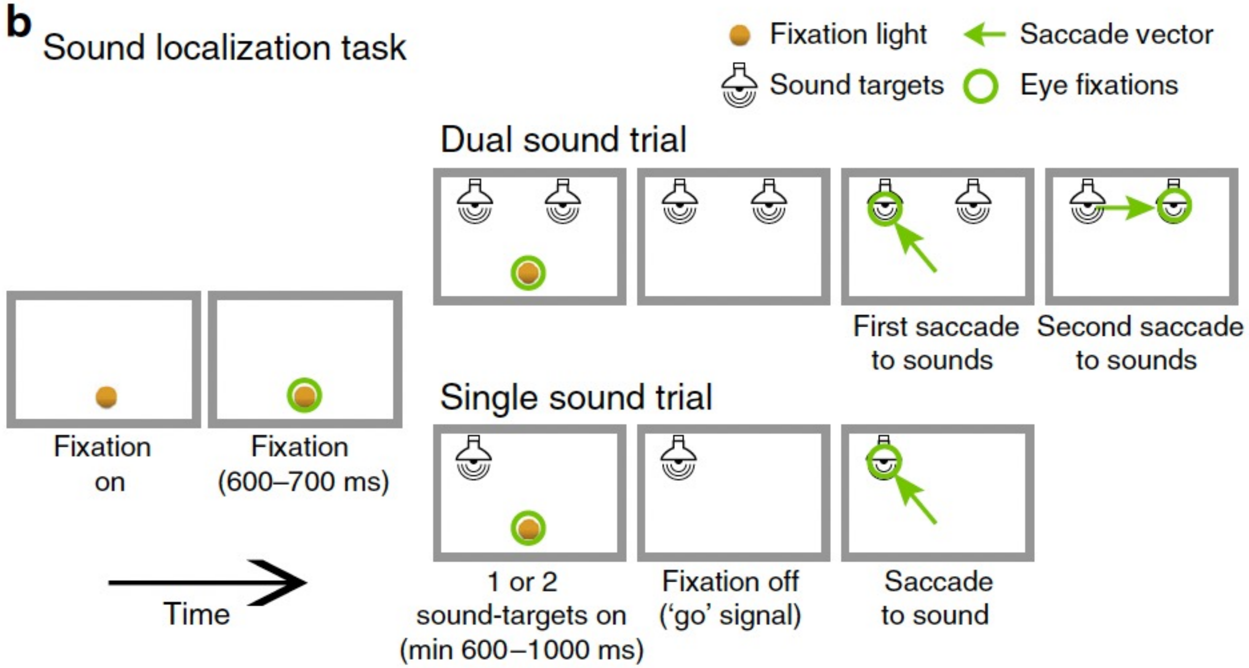


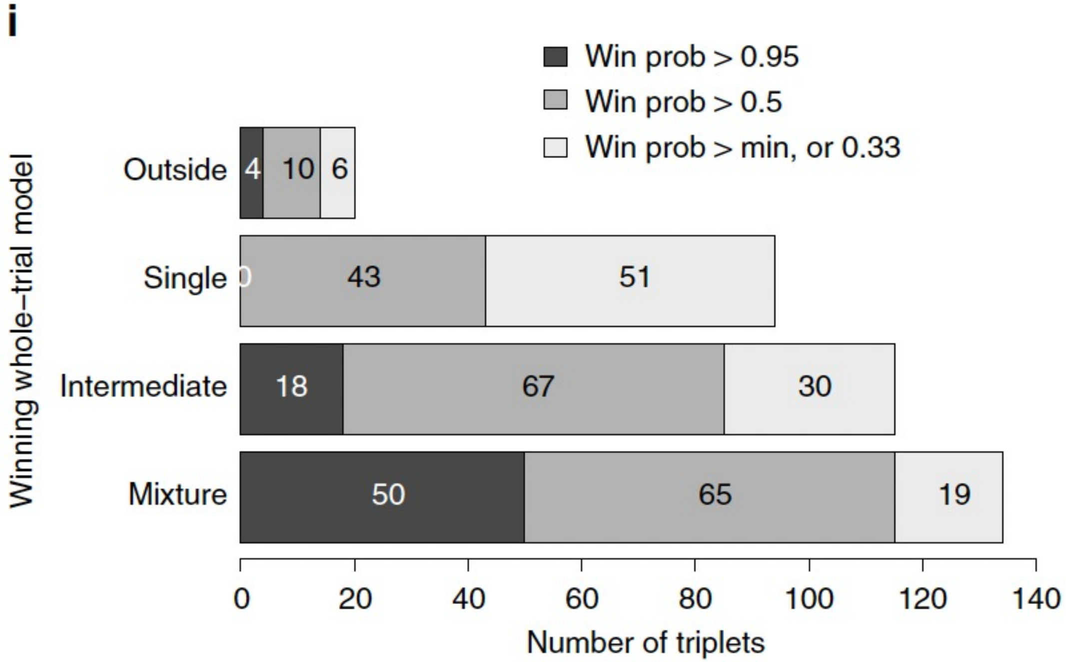
Experiment design (Caruso et al., 2018)

- Localization task: eye movements to sound (saccades)
- Single cell recording in Inferior colliculus (IC): accurate sound localization

Results from Caruso et al. (2018)

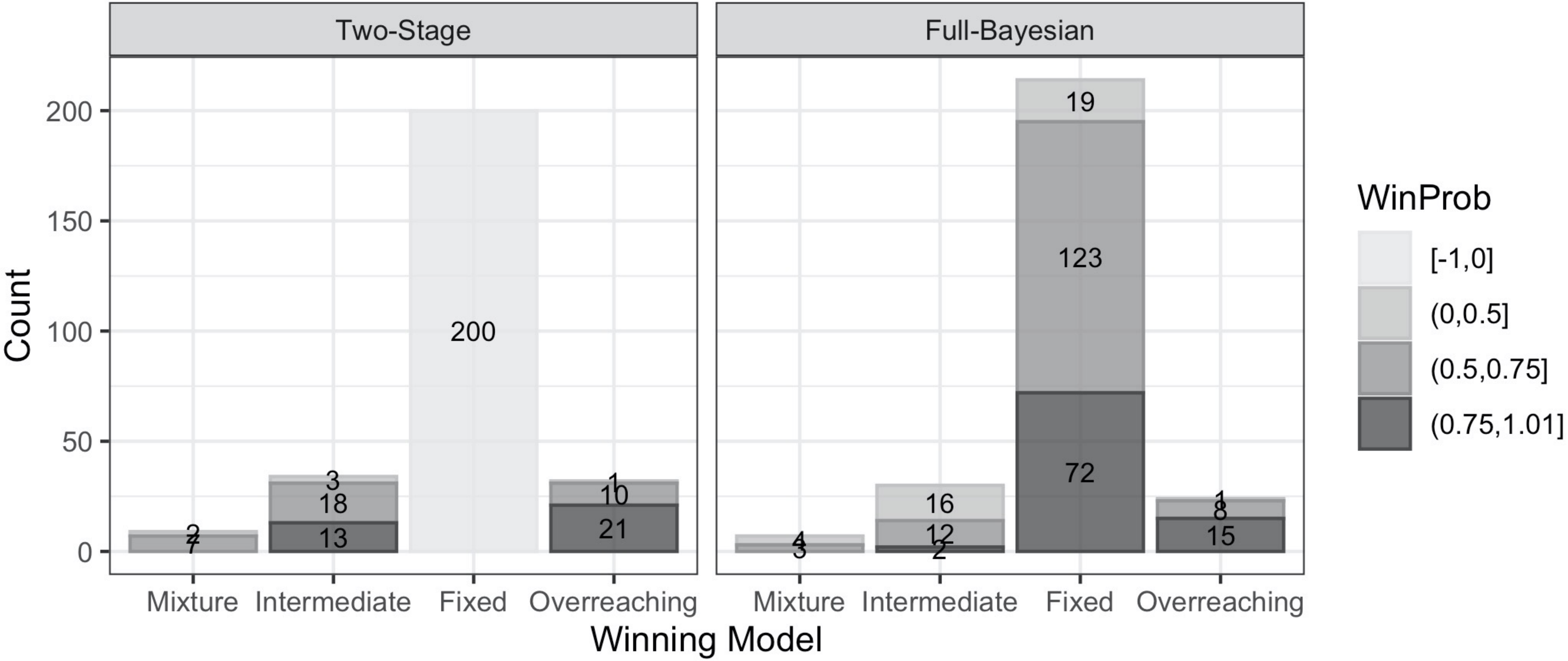
- Chi-square goodness of fit test: 363 triplets



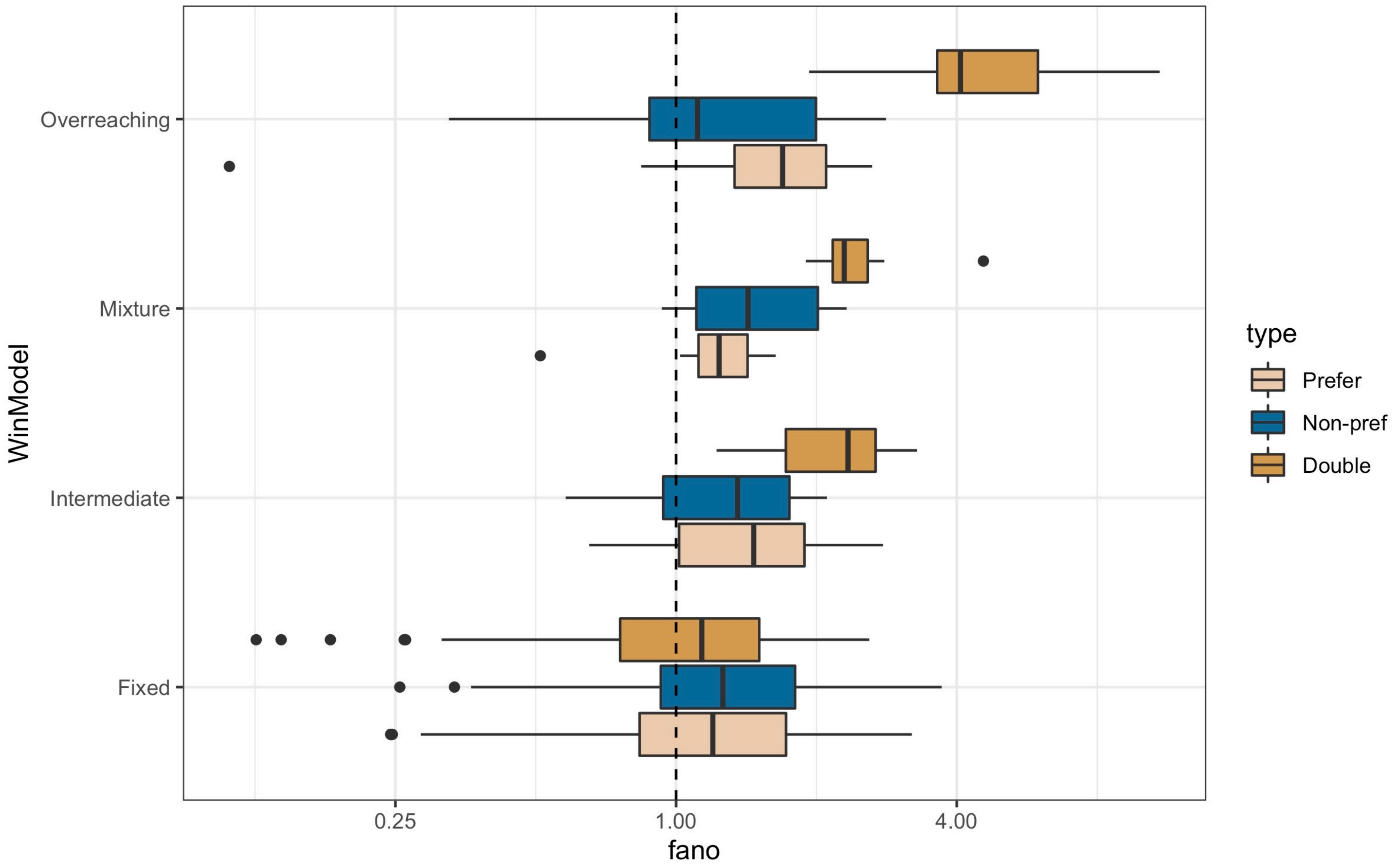


Results

Preprocessing: Poisson variance test (Brown and Zhao, 2002) ensure Poisson-like distribution for single stimulus

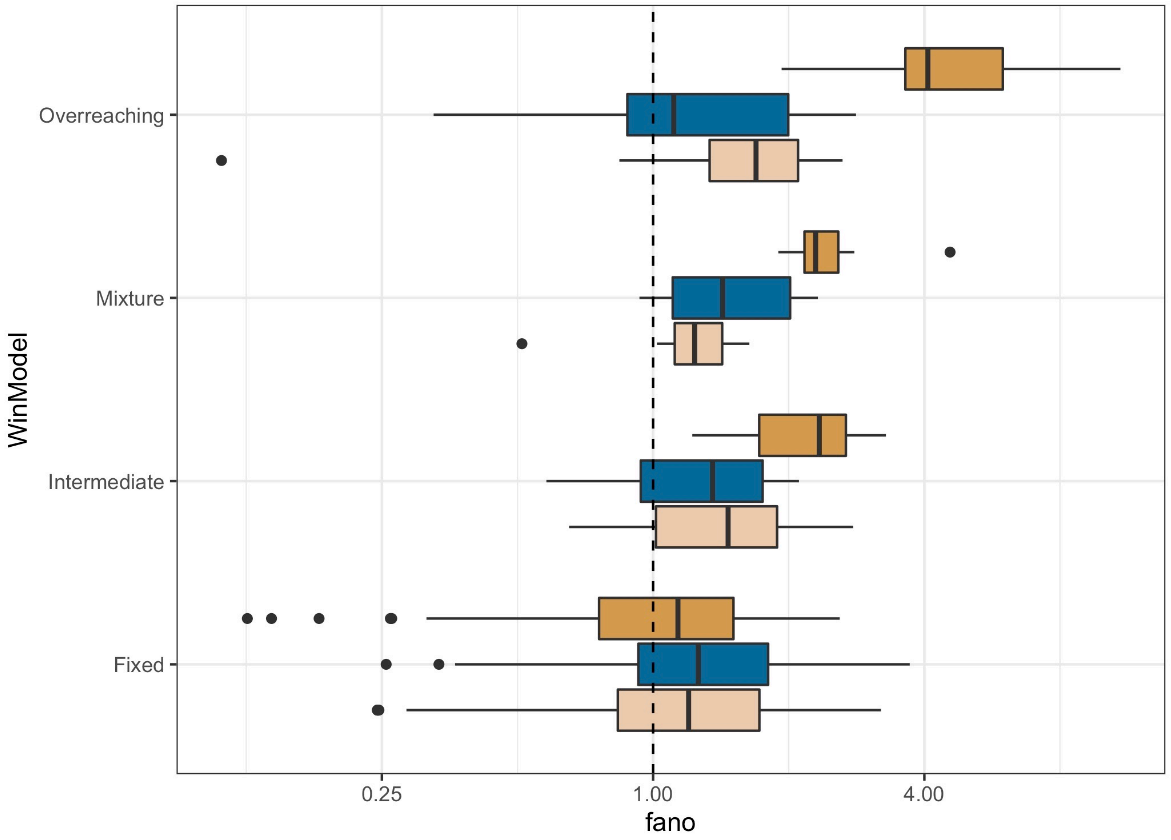


Results

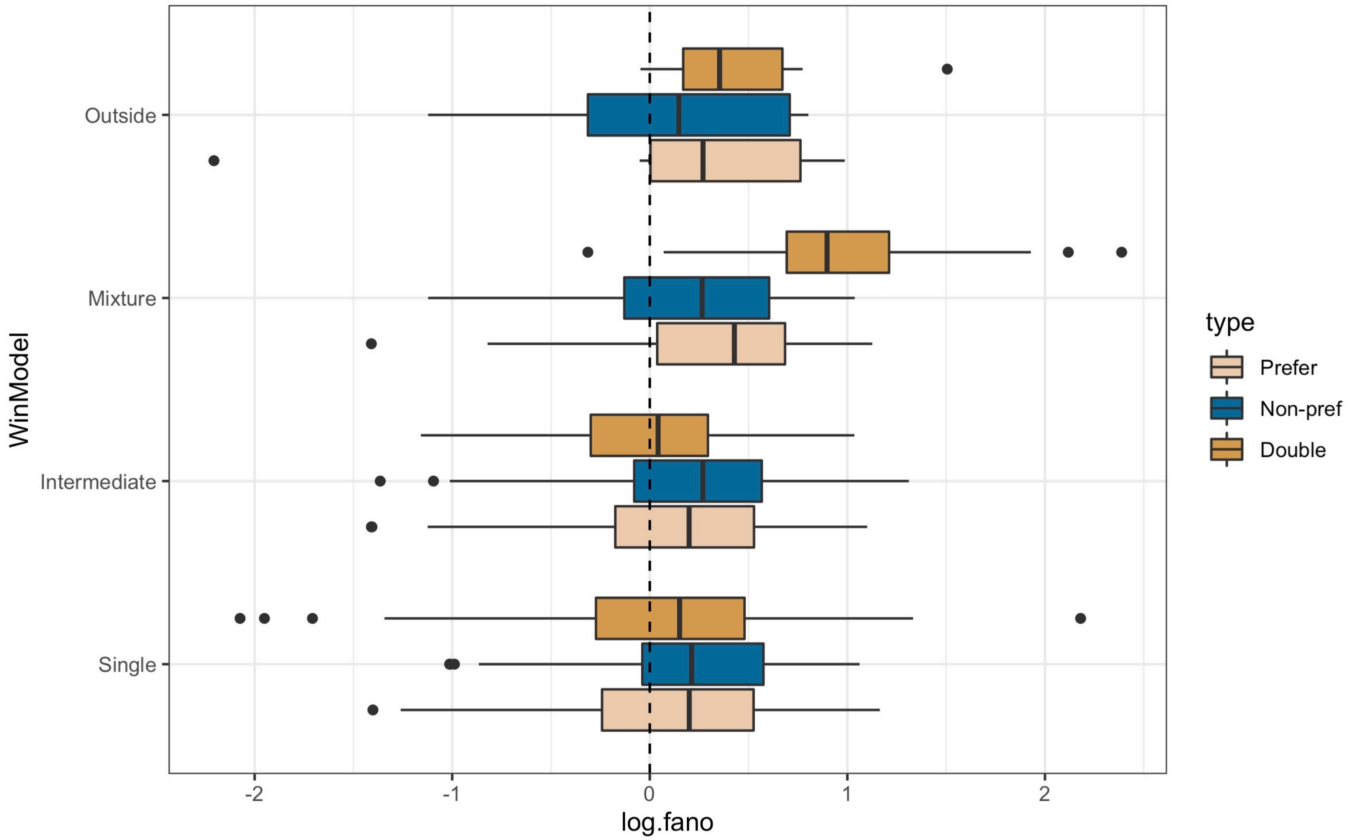

Overreaching
Mixture
Intermediate
Fixed
WinModel
0.25
1.00
4.00
fano
type
Prefer
Non-pref
Double

Results

New Method



Old Poisson



Thank you

Reference

Abramowitz, M. and I. A. Stegun (1965). *Handbook of mathematical functions: with formulas, graphs, and mathematical tables*, Volume 55. Courier Corporation.

Brown, L. D. and L. H. Zhao (2002). A test for the poisson distribution. *Sankhyā: The Indian Journal of Statistics, Series A*, 611–625.

Caruso, V. C., J. T. Mohl, C. Glynn, J. Lee, S. M. Willett, A. Zaman, A. F. Ebihara, R. Estrada, W. A. Freiwald, S. T. Tokdar, et al. (2018). Single neurons may encode simultaneous stimuli by switching between activity patterns. *Nature communications 9*(1), 2715.

Glynn, C. D., S. T. Tokdar, A. Zaman, V. C. Caruso, J. T. Mohl, S. M. Willett, and J. M. Groh (2019). Analyzing second order stochasticity of neural spiking under stimuli-bundle exposure. Submitted.

Martin, R. and S. T. Tokdar (2011). Semiparametric inference in mixture models with predictive recursion marginal likelihood. *Biometrika 98*(3), 567–582.

Newton, M. A. (2002). On a nonparametric recursive estimator of the mixing distribution. *Sankhyā: The Indian Journal of Statistics, Series A*, 306–322.

Newton, M. A., F. A. Quintana, and Y. Zhang (1998). Nonparametric bayes methods using predictive updating. In *Practical nonparametric and semiparametric Bayesian statistics*, pp. 45–61. Springer.

Semedo, J. D., A. Zandvakili, C. K. Machens, M. Y. Byron, and A. Kohn (2019). Cortical areas interact through a communication subspace. *Neuron 102*(1), 249–259.

Tokdar, S. T., R. Martin, J. K. Ghosh, et al. (2009). Consistency of a recursive estimate of mixing distributions. *The Annals of Statistics 37*(5A), 2502–2522.

PRML Algorithm

Input: i.i.d observations $Y_1, ..., Y_n$
Output: marginal likelihood $L_n(\theta) = \prod_{i=1}^n m_{i-1,\theta}(Y_i)$, gradient $\nabla \log L_n(\theta) = \sum_{i=1}^n \nabla \log m_{i-1,\theta}(Y_i)$ and mixing density $f_{n,\theta}$
Initialize: $f_{0,\theta} \in \mathbb{F}$ of f (usually uniform);compute $\nabla f_{0,\theta}(u)$;
weights $w_1, .., w_n \in (0,1)$ (usually $w_i = (i+1)^{-\gamma}$)
for $i = 1,...,n$ **do**

$$m_{i-1,\theta}(Y_i) = \int p(Y_i|\theta,u') f_{i-1,\theta}(u') d\mu(u') \tag{13}$$

$$f_{i,\theta}(u) = (1-w_i) f_{i-1,\theta}(u) + w_i \frac{p(Y_i|\theta,u) f_{i-1,\theta}(u)}{m_{i-1,\theta}(Y_i)} \tag{14}$$

For $\nabla f_{i,\theta}(u)$:

$$G(\theta,u) = p(Y_i|\theta,u) \nabla f_{i-1,\theta}(u) + \nabla p(Y_i|\theta,u) f_{i-1,\theta}(u) \tag{15}$$

$$\nabla \log m_{i-1,\theta}(Y_i) = \frac{\int G(\theta,u) d\mu(u)}{m_{i-1,\theta}(Y_i)} \tag{16}$$

$$\nabla f_{i,\theta}(u) = (1-w_i) \nabla f_{i-1,\theta}(u) + w_i \{ \frac{G(\theta,u) - p(Y_i|\theta,u) f_{i-1,\theta}(u) \nabla \log m_{i-1,\theta}(Y_i)}{m_{i-1,\theta}(Y_i)} \} \tag{17}$$

end