Mathematical Neural Models for Visual Attention

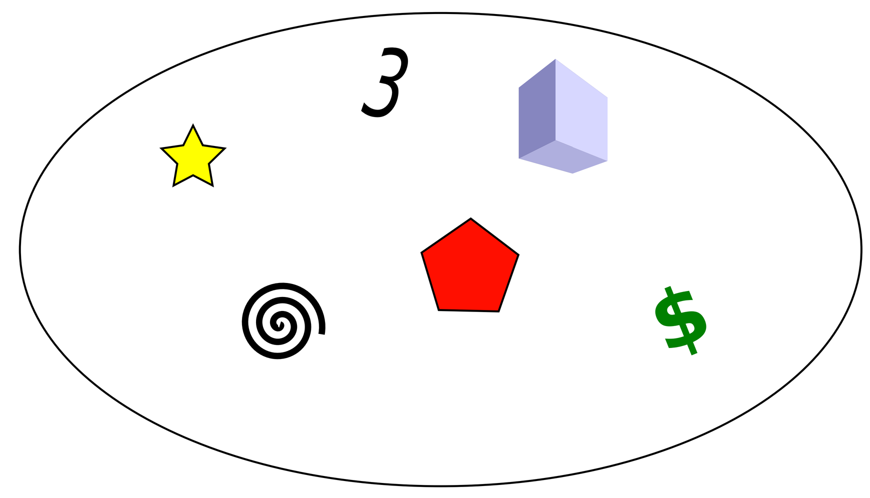
Kang Li[1,2,*], Søren Kyllingsbæk[2], Claus Bundesen[2] and Susanne Ditlevsen[1]

[1]Department of Mathematical Sciences [2]Department of Psychology, University of Copenhagen *kang@math.ku.dk

KØBENHAVNS UNIVERSITET

Motivation

Visual attention refers to the selection of important visual information from a complicated visual field. In psychology, visual attention is usually studied by behavioral tasks using the recorded **response times** or **accuracies**. However, the biological neural mechanisms of the brain is not directly touched.



Motivation: Explain visual attention from a biological level of neurons, the basic processing units in the nervous system. We construct statistical models that combine neural observation (e.g. spike trains) and visual attention theories.

Objective

- Explore, develop and verify neural models for visual attention.
- Explain the neural mechanism during visual attention.
- Investigate the neural code relating external signal to internal spikes, under visual attention theories.

Neural explanation for visual attention

Our neural explanation relies on the **Neural Theory for Visual Attention (NTVA)** proposed by Bundesen et al (2005), which states that a neuron, when presented to multiple objects, can only respond to a single stimulus object at one time. On the other hand, empirical studies by Reynolds et al (1999) show that the neuronal response to multiple stimuli is a weighted average of responses to single objects. Following the two opposing hypotheses, we formulate two models on a single neuron level:

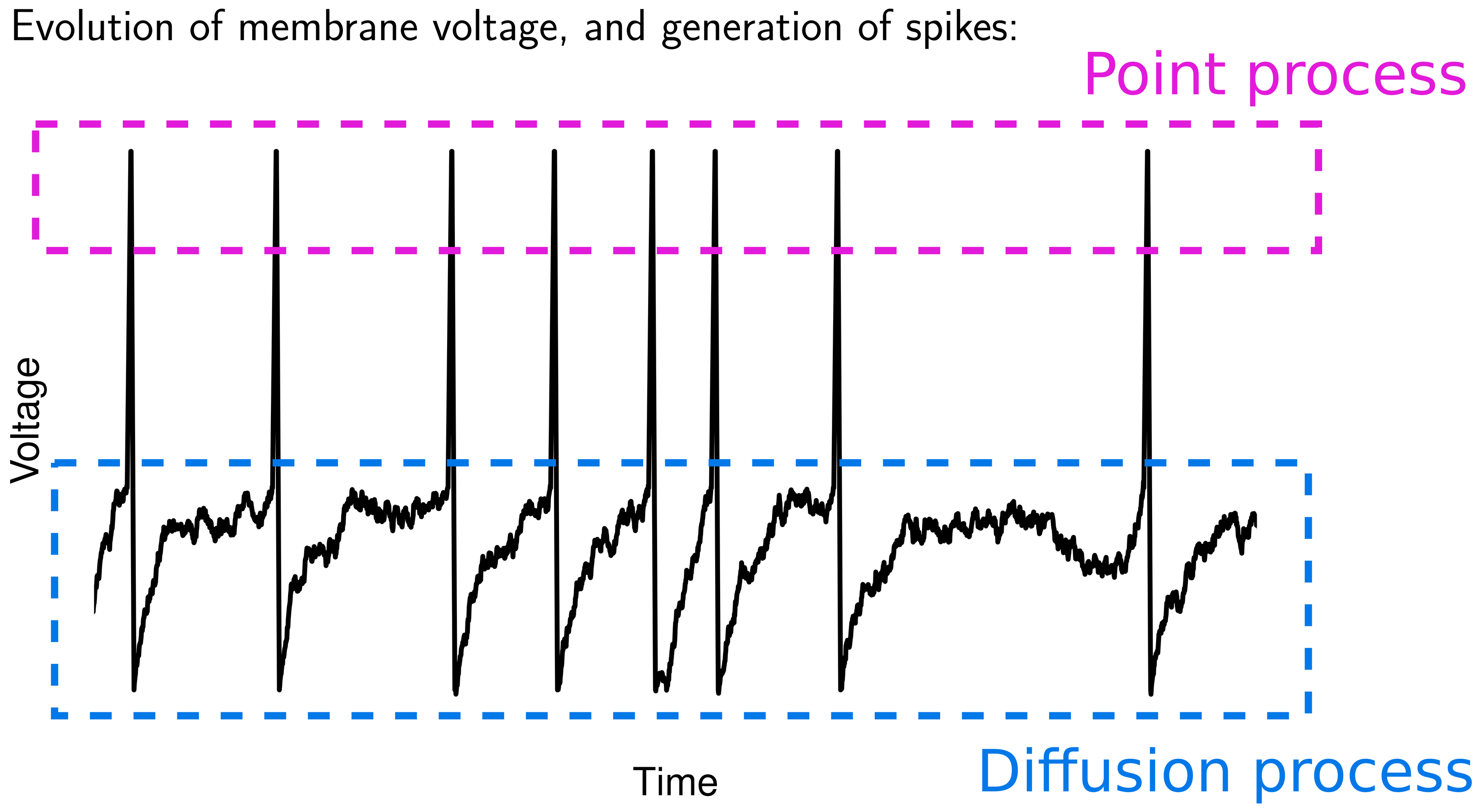
- **Probability mixing**: the neuron follows a probability mixture, responding to each single object with probabilities;
- **Response averaging**: the neuron's response is a weighted average of responses to each single object.

Furthermore, based on NTVA and probability mixing, we formulate two opposing models for neural ensembles:

- **Parallel processing**: Neurons split the attention between different objects, with some neurons attending one object while some others attending another one. All objects are processed simultaneously (in parallel).
- **Serial processing**: At any given time, all neurons attend the same stimulus, and they change the attention together. Objects are processed sequentially.

Spiking neuron models

Evolution of membrane voltage, and generation of spikes:

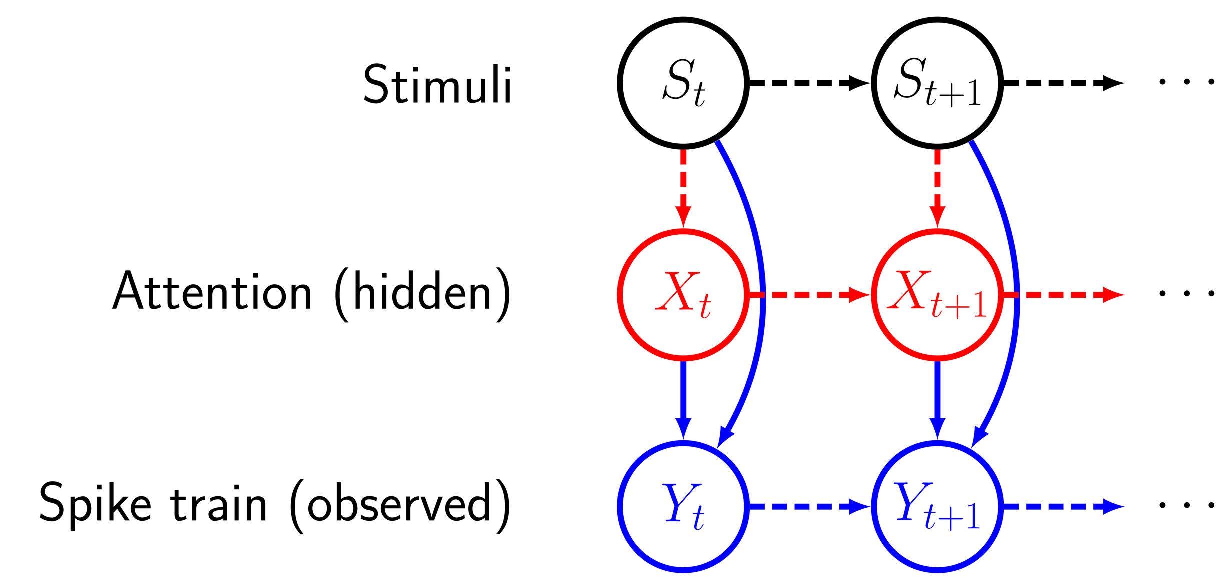


We employ two types of models:

- **Point process** models, where we only look at the discrete sequence of spikes, and model the **conditional intensity function** for the point process.
- **Leaky integrate-and-fire** models, where the voltage is modeled as a **diffusion process** incorporating spiking history effects. A spike is formed whenever the voltage passes a threshold value. The likelihood function can be evaluated via the **first-passage time** problem by solving the **Fokker-Planck** equations.

State-space representation

We construct a unified state-space model to describe visual attention through neurons using spike train data, combining the above two components.



Neural explanation for visual attention governs the transitions of the attention states X. **Spiking neuron models** govern the formation of spike trains Y.

Neural coding in visual attention

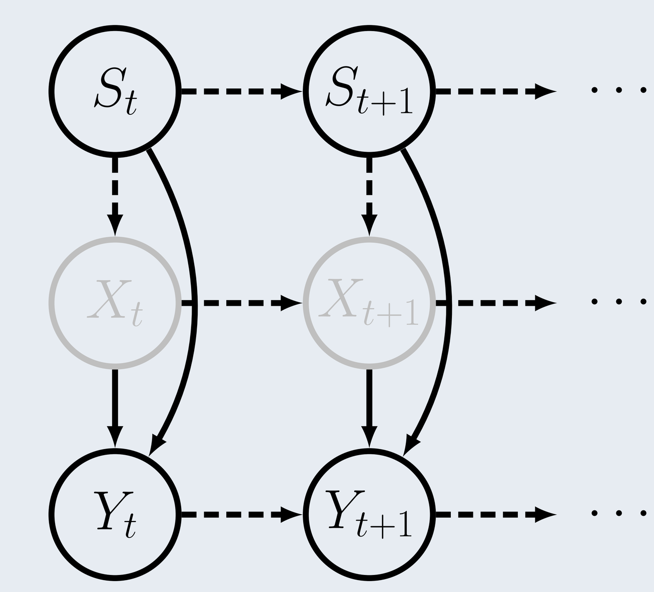
Encoding

Goal: S and Y known, X hidden. Given $S_{1:T} = s_{1:T}$ and $Y_{1:T} = y_{1:T}$, estimate the parameters θ for all underlying distributions.

Maximum likelihood:

$$\hat{\theta} = \arg\max_{\theta} \int p(y_{1:T}|s_{1:T}, x_{1:T}) p(x_{1:T}|s_{1:T}) \mathrm{d}x_{1:T}$$

For **discrete** X: evaluating marginal likelihood;
For **continuous** X: (Sequential) Monte Carlo for X, giving pseudo-marginals.



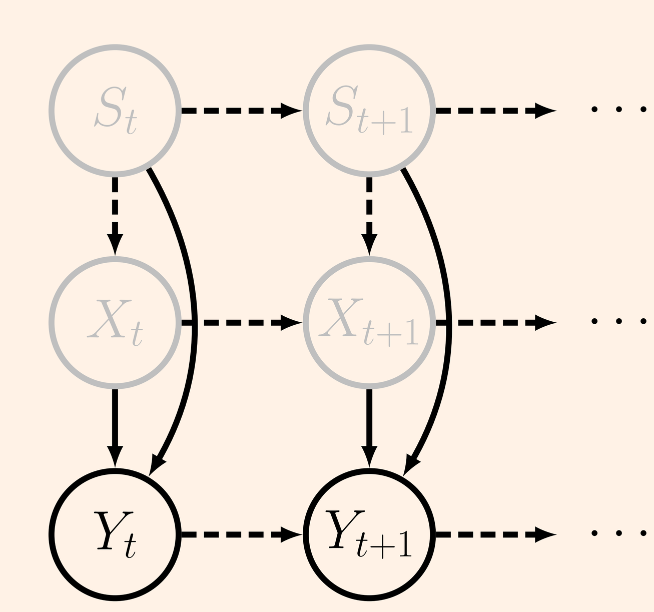
Decoding

Goal: Y known, X and S hidden. Given fitted parameters θ and $Y_{1:T} = y_{1:T}$, infer unknown $S_{1:T}$ and/or $X_{1:T}$.

Method: Obtain the conditional distribution through **sequential Monte Carlo**:

$$\begin{aligned} & p_\theta(s_{1:T}, x_{1:T}|y_{1:T}) \\ \propto\ & p_\theta(y_{1:T}|x_{1:T}, s_{1:T}) p_\theta(x_{1:T}|s_{1:T}) p_\theta(s_{1:T}) \end{aligned}$$

- Online filtering; offline smoothing
- (Auxiliary) Particle filter; parameter learning



Result overview

The state-space model has been applied in different situations:

1. Compare **probability mixing** and **response averaging** on experimental data, using **Hawkes point process** models for spike trains. NTVA and probability mixing were supported. (Li et al, Frontiers in Computational Neuroscience, 2016)
2. Distinguish between **parallel** and **serial processing** on experimental data, also using **Hawkes point process** models.
3. Distinguish between **probability mixing** and **response averaging** in more realistic biophysical settings using the **LIF** model with simulations. (Li et al, JMN, 2016)
4. Investigate **neural decoding** in biophysical settings using the **LIF** model with simulations. Various SMC methods were explored.

Summary

- Constructed and verified novel mathematical neural models for visual attention.
- Combined spiking neuron models with visual attention theories.
- Applied Hawkes point process models and leaky integrate-and-fire models.
- Explained NTVA considering a single neuron (probability mixing and response averaging) and neuron ensembles (parallel and serial processing).
- Formulated the models as a unified state-space framework for neural encoding and decoding in visual attention.
- The application of these models provides both biological and statistical insights.