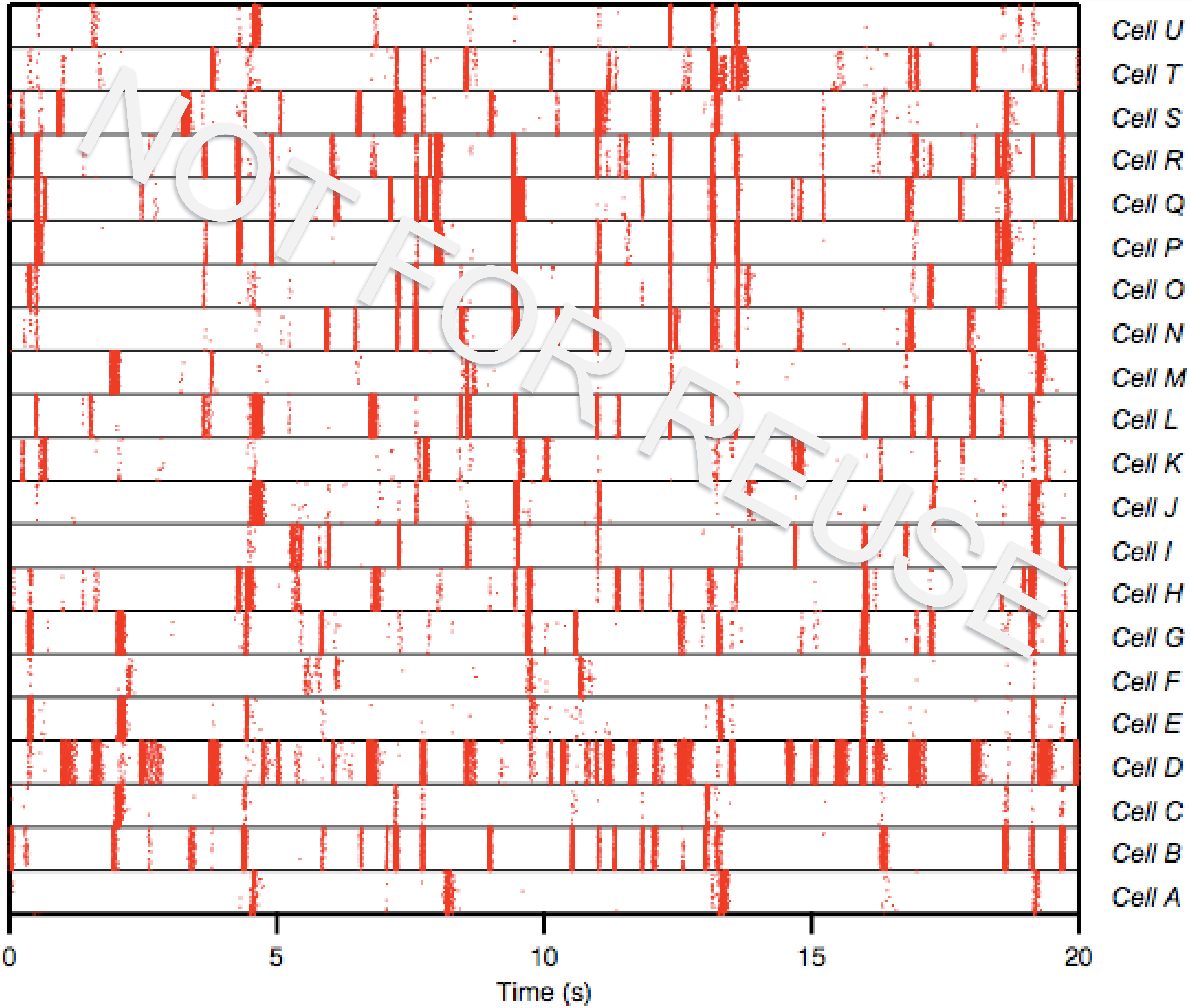
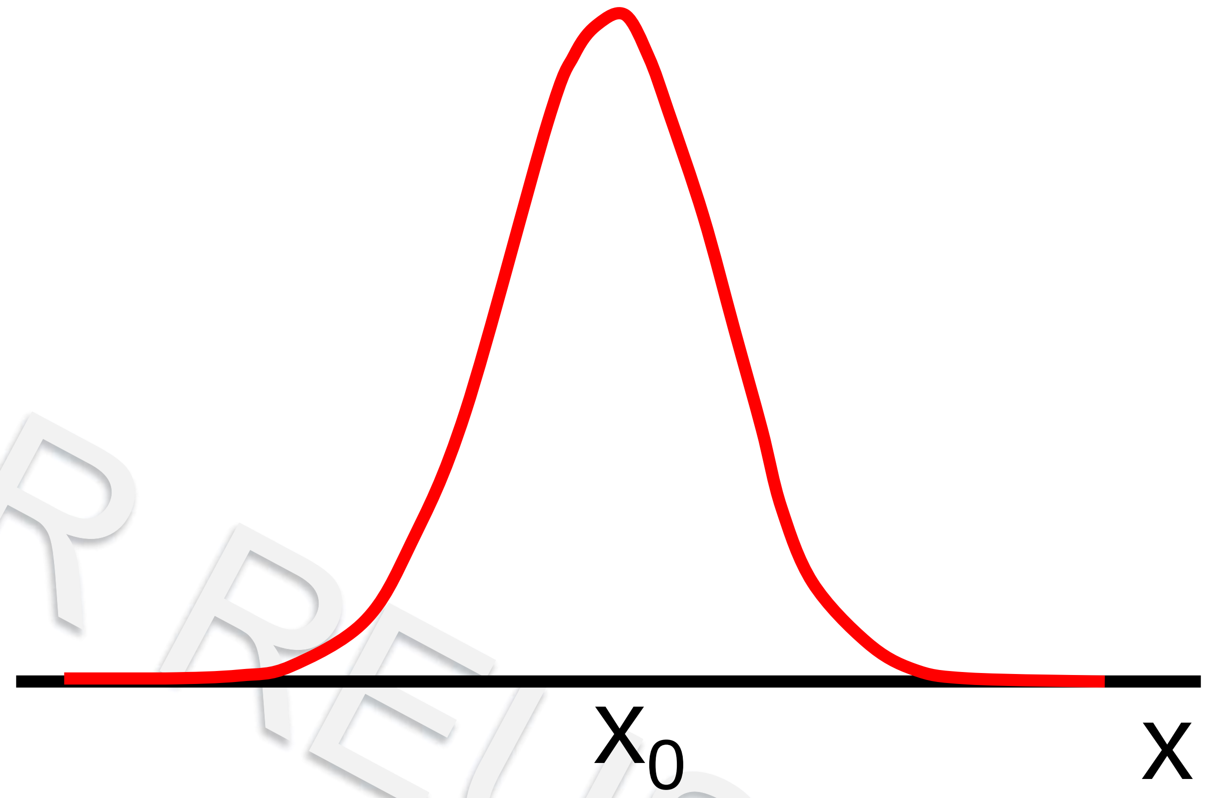
Better models



The magical Gaussian

$$P(x) = Ae^{-((x-x_0)^2/2\sigma^2)}$$



When have you found a good feature or features?

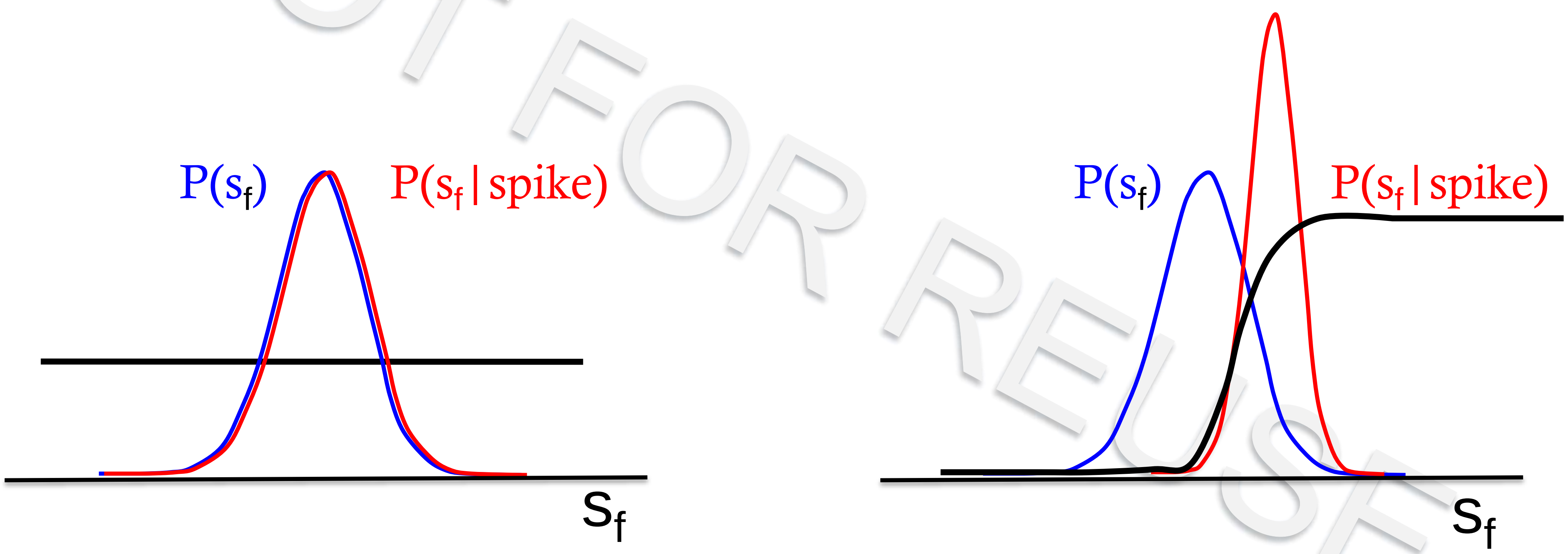
- When the input/output curve over your variable is *interesting*.
- How to quantify interesting?

When have you done a good job?

Tuning curve: $P(\text{spike}|s_f) = P(s_f|\text{spike})\, P(\text{spike}) / P(s_f)$

Boring: spikes unrelated to stimulus

Interesting: spikes are selective



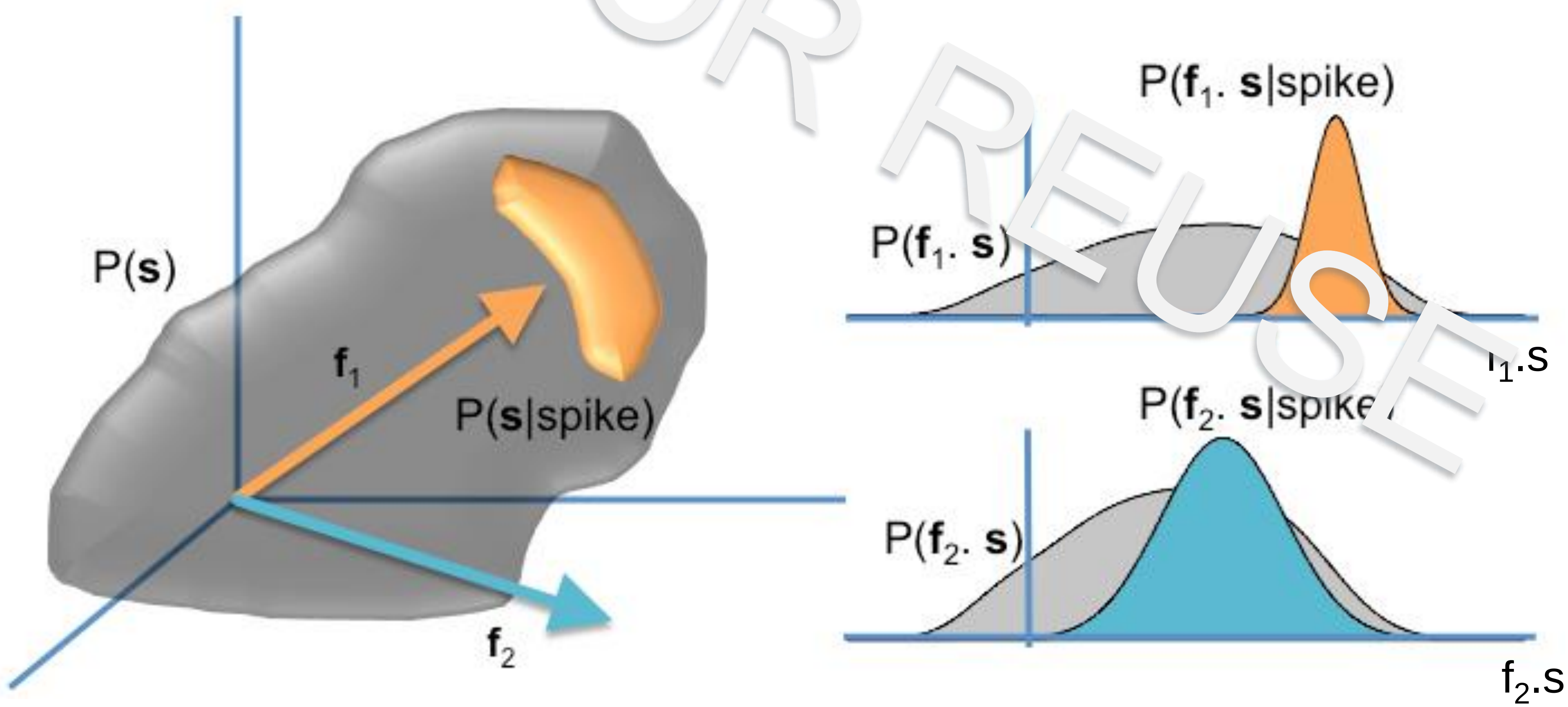
Introducing the Kullback-Leibler divergence

$$D_{KL}(P(s),Q(s)) = \int ds\, P(s) \log_2 P(s)/Q(s)$$

Goodness measure: $D_{KL}(P(s_f|\text{spike}), P(s_f))$

Maximally informative dimensions

Choose filter in order to maximize D_{KL} between spike-conditional and prior distributions



Sharpee, Rust and Bialek, Neural Computation (2004)

Image from Fairhall, Barreiro, Shea-Brown (2012)

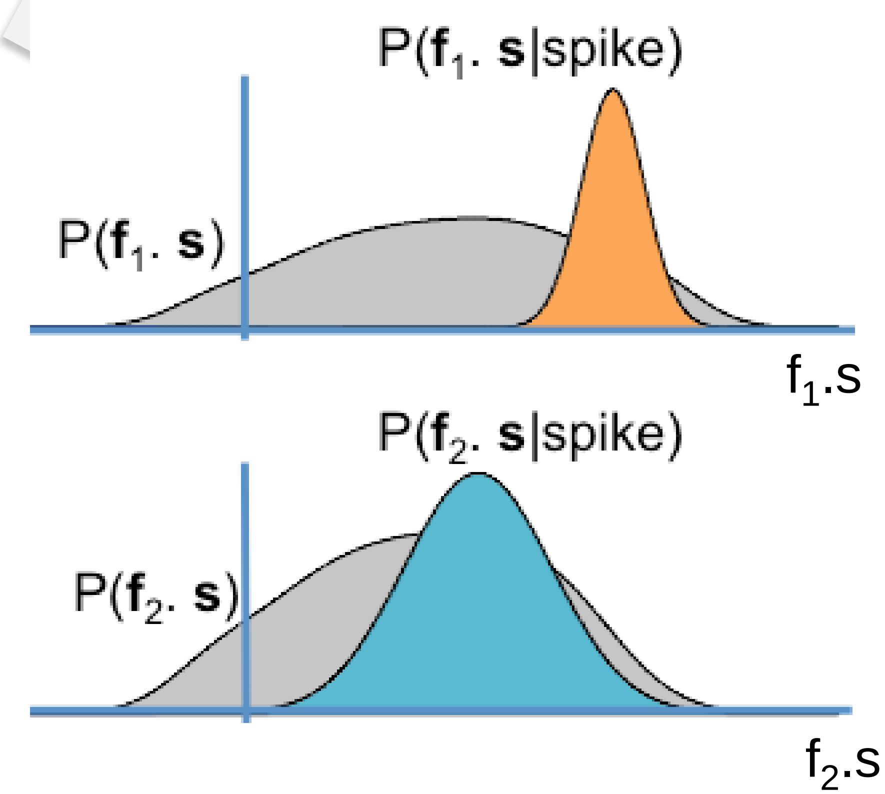
Maximally informative dimensions

Choose filter in order to maximize D_{KL} between spike-conditional and prior distributions

Equivalent to maximizing mutual information between stimulus and spike

Does not depend on white noise inputs

Can be used for deriving models from natural stimuli



Finding relevant features

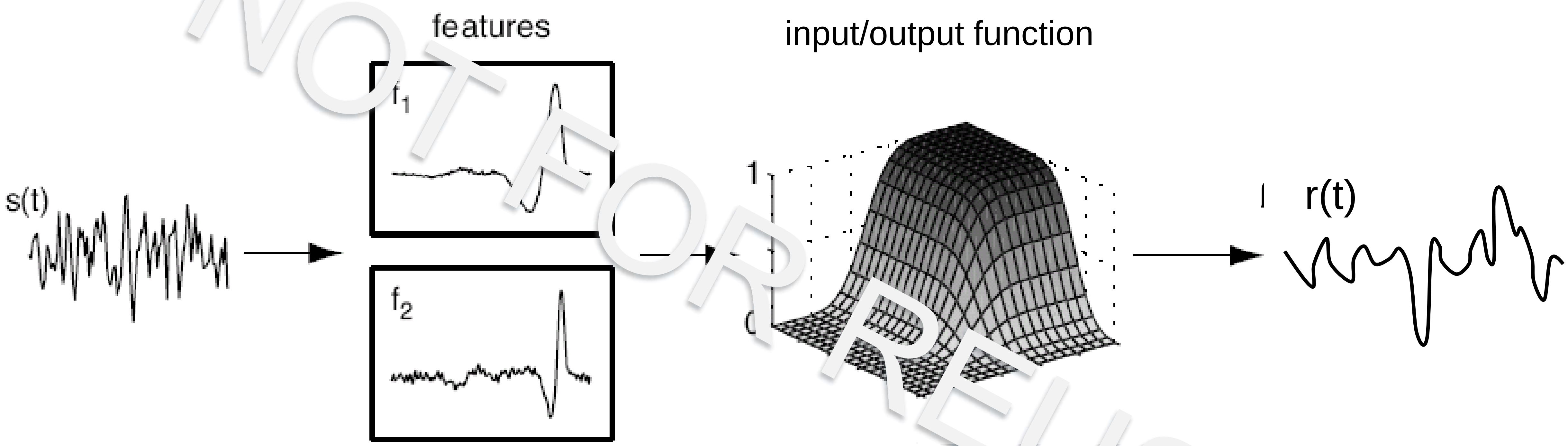
1. Single filter determined by the conditional average

2. A family of filters derived using PCA

3. Information theoretic methods use the whole distribution

Removes requirement for Gaussian stimuli

Modeling the noise



Bernoulli trials

Binomial spiking

Distribution: $P_n[k] = ?$

Mean: $\langle k \rangle = ?$

Variance: $\mathrm{Var}(k) = ?$

Poisson spiking

Distribution: $P_T[k] = (rT)^k \exp(-rT)/k!$

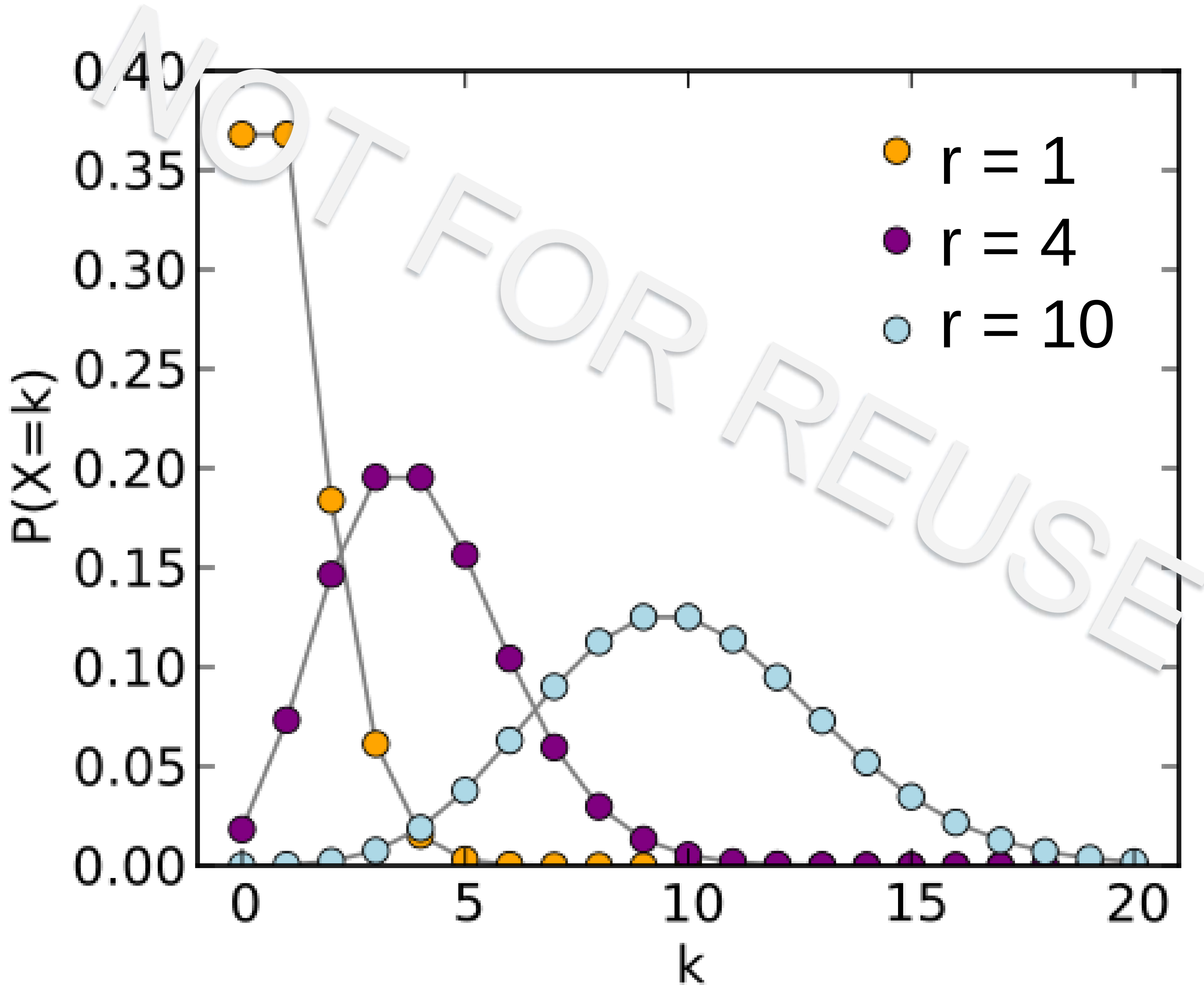
Mean: $<k> = rT$

Variance: $Var(k) = rT$

Fano factor: $F = 1$

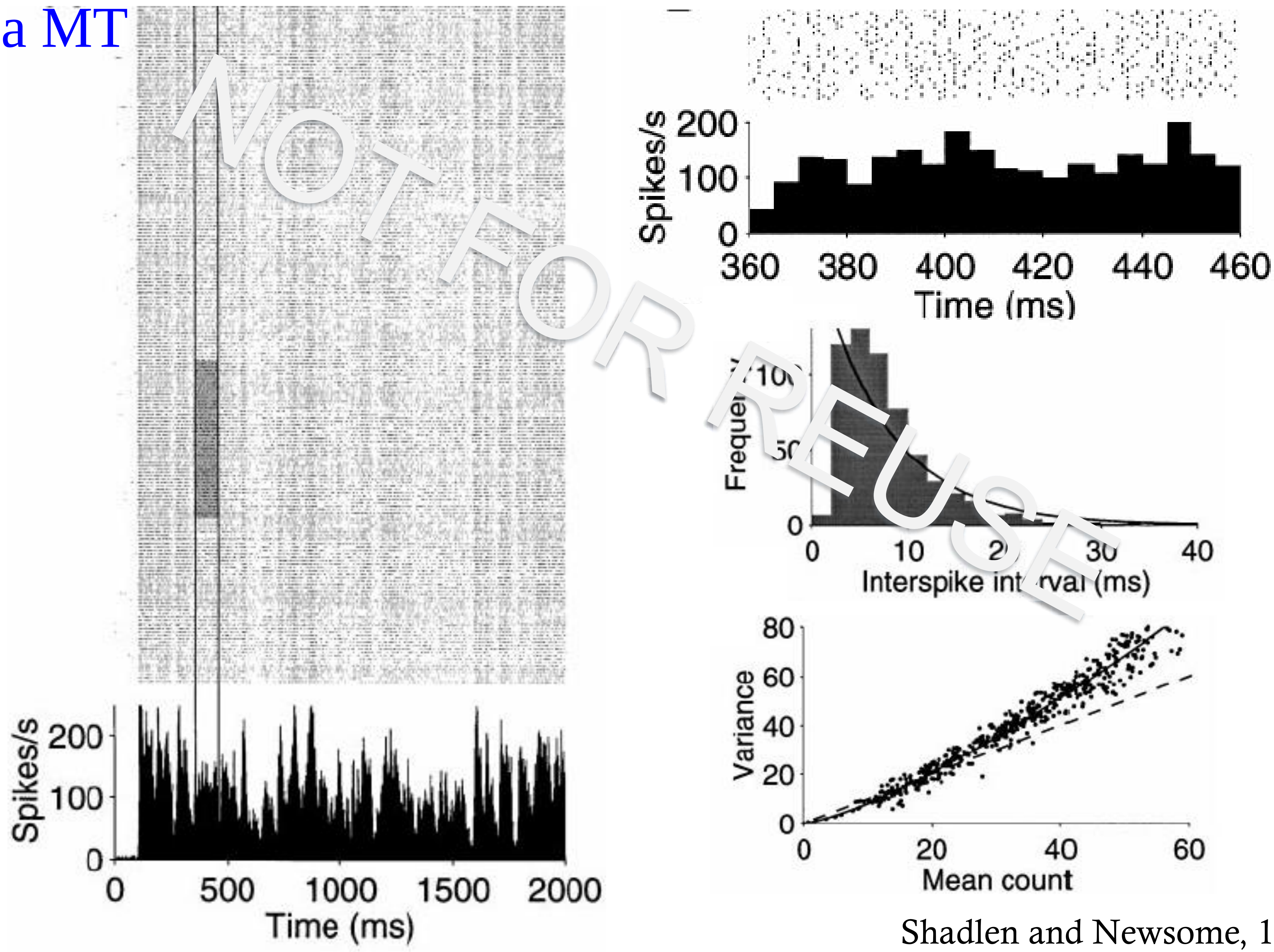
Interval distribution: $P(T) = r \exp(-rT)$

The Poisson distribution



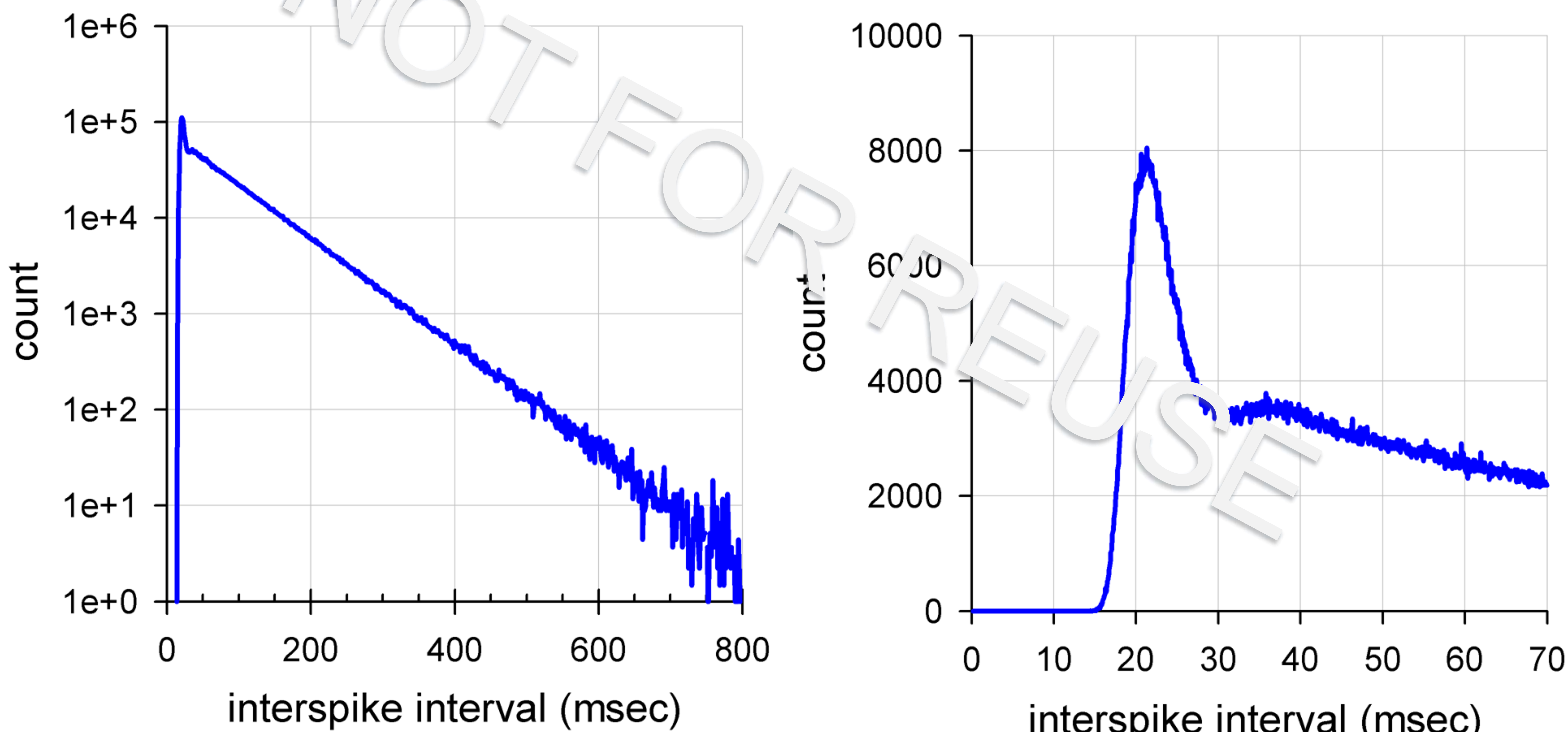
Poisson or not?

Area MT

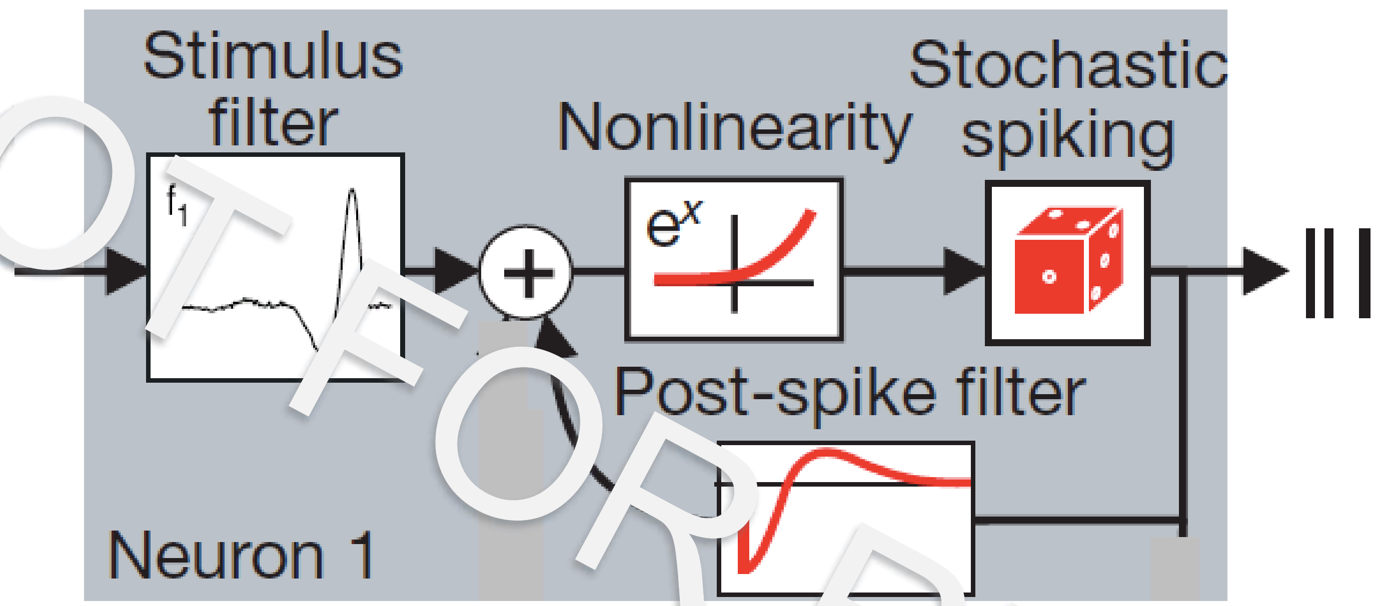


Shadlen and Newsome, 1998

Interspike interval distributions



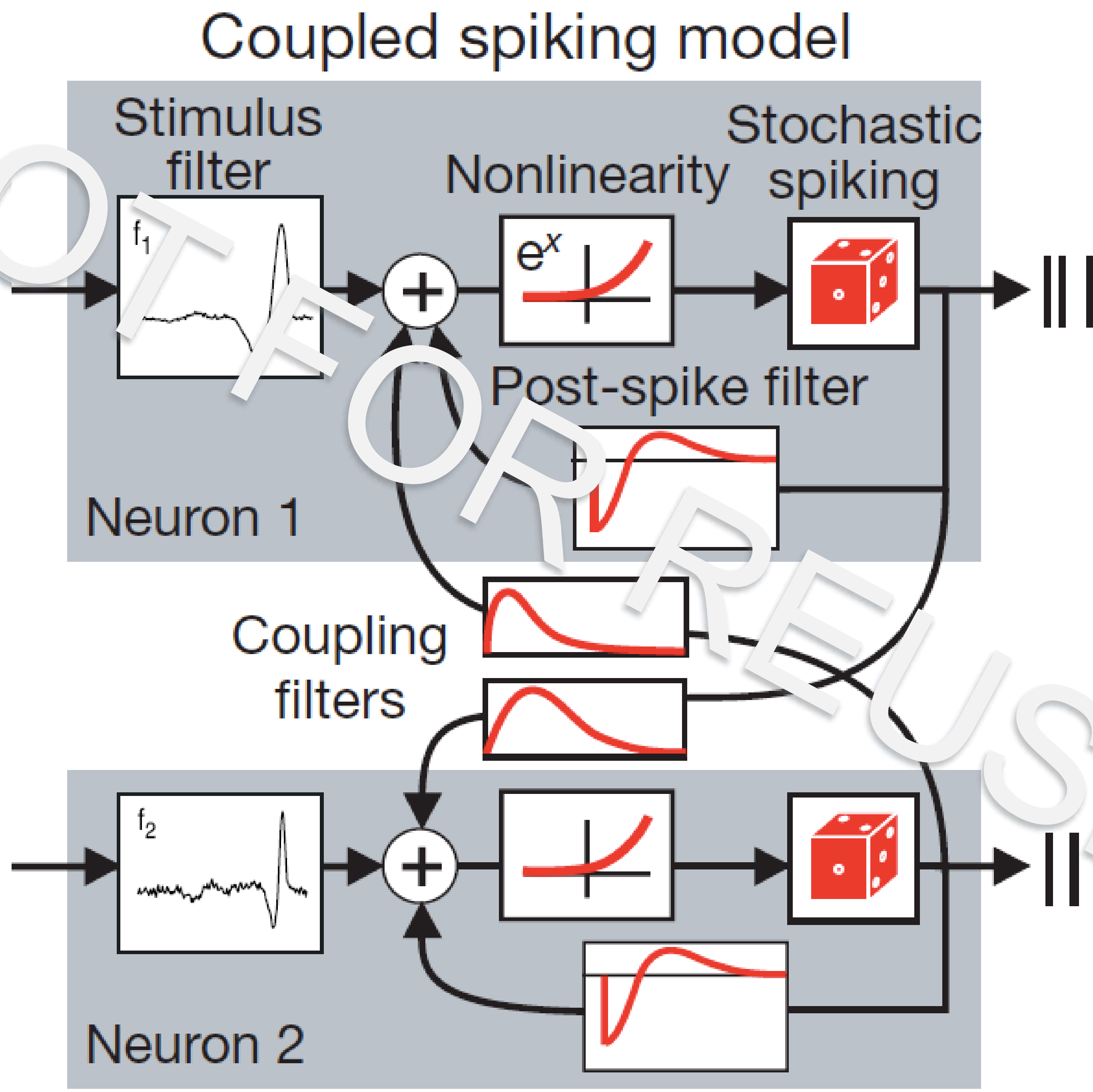
The generalized linear model



GLM: P(spike at t) ~ $\exp(f_1 * s + h_1 * r)$

Pillow et al., *Nature* 2008

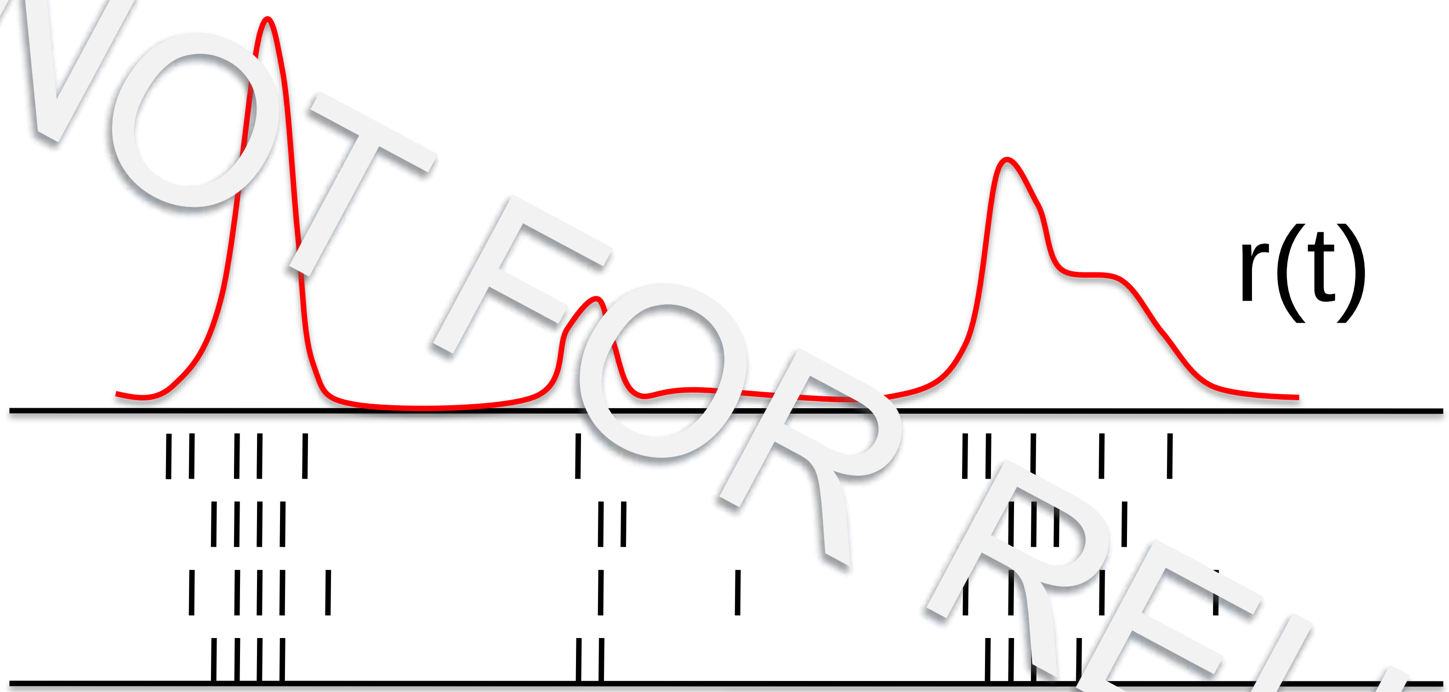
But wait, there's more!

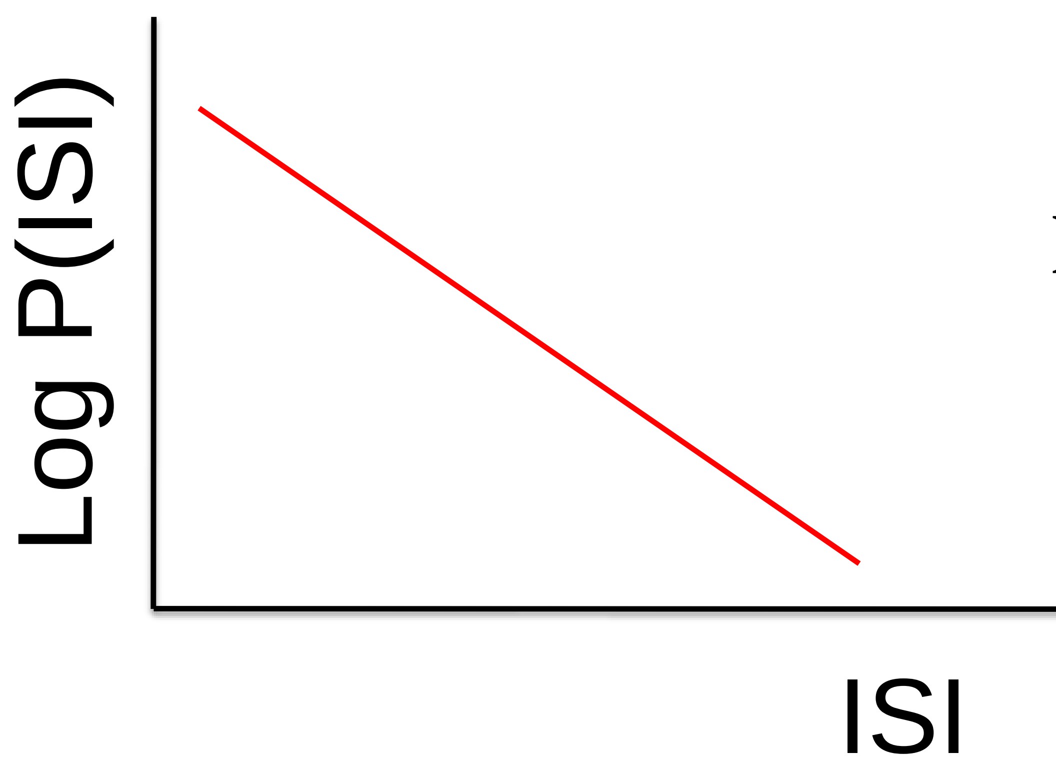


$$\text{GLM: } r(t) = g(f_1{*}s + h_1{*}r_1 + h_2{*}r_2 + \ldots)$$

Pillow et al., *Nature* 2008

Time-rescaling theorem





Exponential distribution of *scaled* ISIs: $(t_{i-1-}\, t_i)\ r(t)$

Brown et al. (2001)

That's it for encoding!

That's it for encoding!

Next week…

- Reading minds!
- Decoding methods