Modeling decisions

Note: many slides based on talk by Phil Holmes, Princeton

Moving dots decision task: Left or right?

Example of two-alternative decision task

Newsome, Movshon, Zohary, Shadlen, Gold, Britten … '90s and '00s

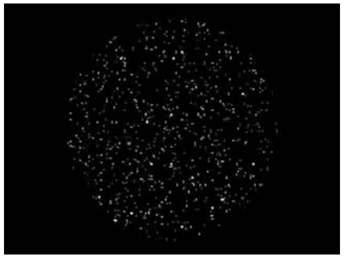
30 % coherent

5 % coherent

Bill Newsome

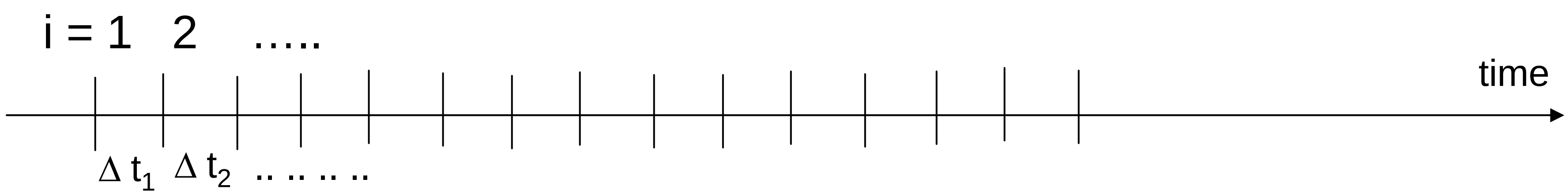
- Behavioral measures: reaction time (RT) distribution, error rate (ER).
- Neural measures: fMRI (humans), direct recordings in visual processing and motor areas (monkeys: MT, LIP, FEF).

Statistical hypothesis testing: discrimination among alternatives



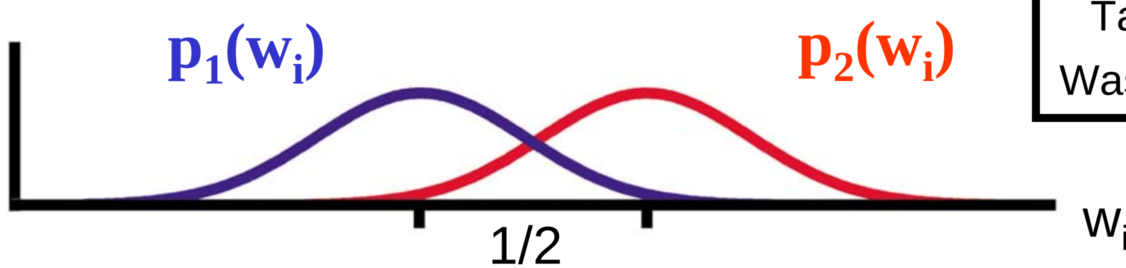
Hypothesis 1 – leftward dots
Hypothesis 2 – rightward dots

Consider increments of time Δt_i



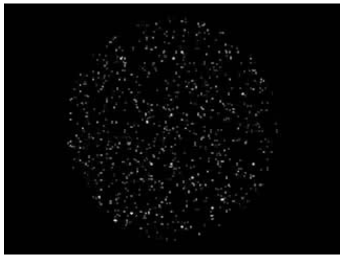
w_i = random string drawn from 2 distributions $p_1(w), p_2(w)$ representing i'th increment of evidence for hypotheses 1, 2 resp.

e.g. w_i = ‘fraction of right-going dots observed” over Δt_i



Task: given {w_i},
Was hyp. 1or 2 true?

Statistical hypothesis testing: discrimination among alternatives

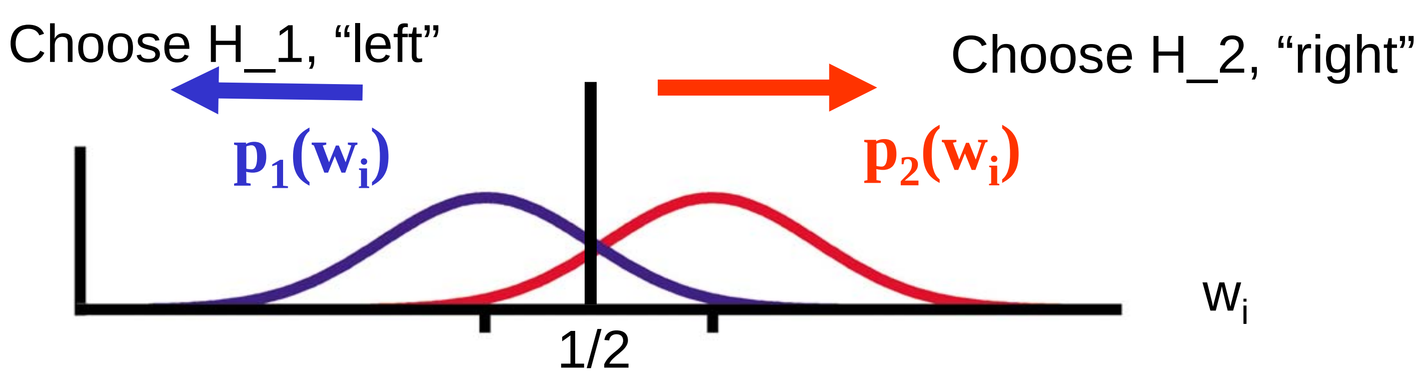


Hypothesis 1 – leftward dots
Hypothesis 2 – rightward dots

Partial (optimal) solution to "static" problem:

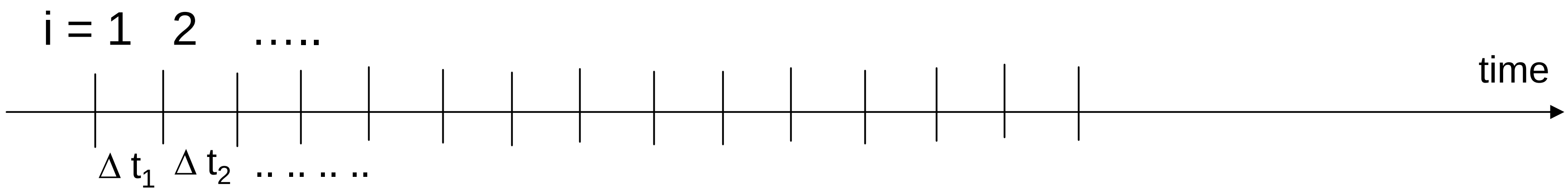
Max. likelihood based on a $\boxed{\text{‘}\underline{\text{single}}\text{’ time interval}}$

e.g. w_i = ‘fraction of right-going dots observed” over Delta t_i

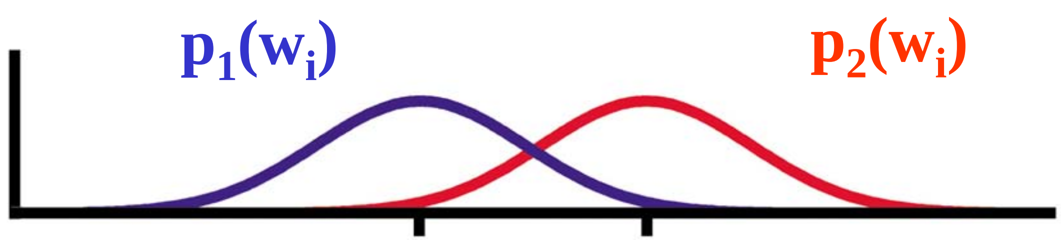


What about the DYNAMIC problem: decision?

Consider increments of time Δt_i



w_i = random string drawn from 2 distributions $p_1(w), p_2(w)$ representing i'th increment of evidence for hypotheses 1, 2 resp.



<u>SEQUENTIAL PROBABILITY RATIO TEST (SPRT): Wald, 1947</u>

See review, Gold and Shadlen 2001

Consider the likelihood ratio

$$R_n = \prod_{i=1}^{n} \frac{p_1(w_i)}{p_2(w_i)}$$

SPRT: Fix thresholds $Z_1 < 1 < Z_2$ and continue observing as long as

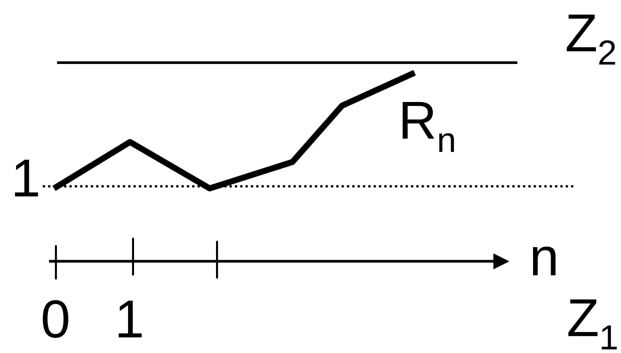
$$Z_1 < R_n < Z_2 \quad .$$

As soon as

$$R_n \leq Z_1 \;\Rightarrow\; \text{declare hyp 1 true,}$$

or

$$R_n \geq Z_2 \;\Rightarrow\; \text{declare hyp 2 true.}$$



"Random walk"

$$R_n = \prod_{i=1}^{n} \frac{p_1(w_i)}{p_2(w_i)}$$

SPRT: Fix thresholds $Z_1 < 1 < Z_2$ and continue observing as long as

$$Z_1 < R_n < Z_2 \; .$$

As soon as

$$R_n \le Z_1 \;\Rightarrow\; \text{declare hyp 1 true,}$$

or

$$R_n \ge Z_2 \;\Rightarrow\; \text{declare hyp 2 true.}$$

Let $E_j(N)$ = expected no. obs needed to declare hyp j true with specified error probabilities a_j, $j = 1, 2$.

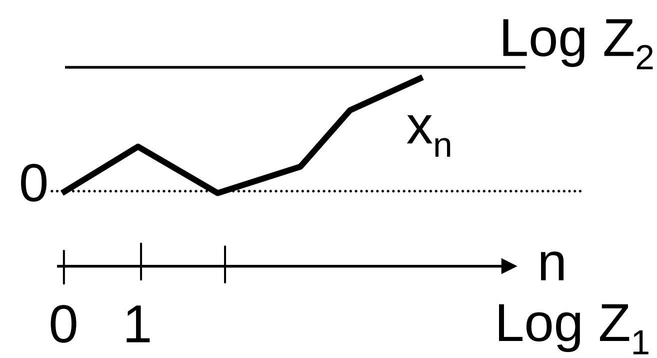
Theorem (Wald, Barnard) Among all fixed sample or sequential tests, SPRT with error probabilities a_j minimises $E_j(N)$.

There are formulae for Z_j in terms of a_j.

Take logarithm to make SPRT an ‘additive’ test in time

$$x_n = \log(R_n) = \log \prod_{i=1}^{n} \frac{p_1(w_i)}{p_2(w_i)}$$

$$= \sum \log \frac{p_1(w_i)}{p_2(w_i)} = \sum \delta I_i$$



E.g. If $p_j($ are normal distributions,

$$p_j(w) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[\frac{-(w - \mu_j)^2}{2\sigma^2}\right] ,$$

a short calculation shows that:

$$E(\delta I_i) = :(\mu_1 - \mu_2)^2/\sigma^2 \stackrel{\text{def}}{=} A\,, \text{ if } w_i \text{ drawn from } p_1;$$ (opposite sign if from p_2)

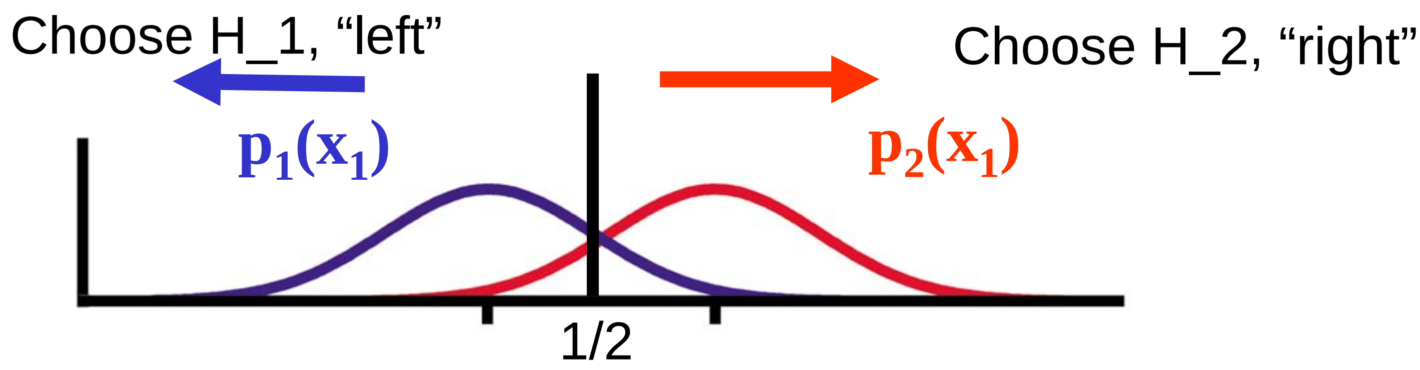
$$Var(\delta I_i) = |\mu_2 - \mu_1| \stackrel{\text{def}}{=} D\,.$$

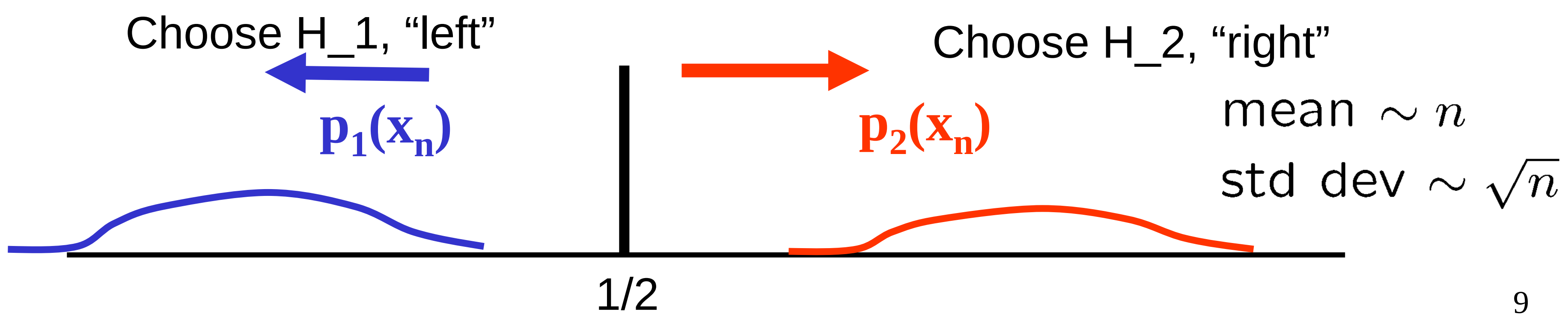
"Random walk"

Take logarithm to make SPRT an ‘additive’ test in time

$$x_n = \log(R_n) = \log \prod_{i=1}^{n} \frac{p_1(w_i)}{p_2(w_i)} \qquad (1)$$

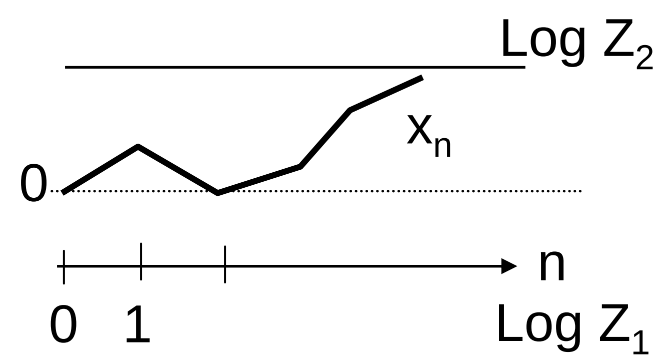
$$= \sum \log \frac{p_1(w_i)}{p_2(w_i)} = \sum \delta I_i \qquad (2)$$





Take logarithm to make SPRT an 'additive' test in time

$$
\begin{aligned}
x_n &= \log(R_n) = \log \prod_{i=1}^{n} \frac{p_1(w_i)}{p_2(w_i)} \qquad (1)\\
&= \sum \log \frac{p_1(w_i)}{p_2(w_i)} = \sum \delta I_i \qquad (2)
\end{aligned}
$$



E.g. If $p_j($ are normal distributions,

$$p_j(w) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[\frac{-(w - \mu_j)^2}{2\sigma^2}\right] ,$$

a short calculation shows that:

$$E(\delta I_i) = (\mu_1 - \mu_2)^2/\sigma^2 \stackrel{\text{def}}{=} A\,,\ \text{if } w_i \text{ drawn from } p_1, \quad \text{(opposite sign if from } p_2)$$

$$Var(\delta I_i) = |\mu_2 - \mu_1| \stackrel{\text{def}}{=} D\,.$$

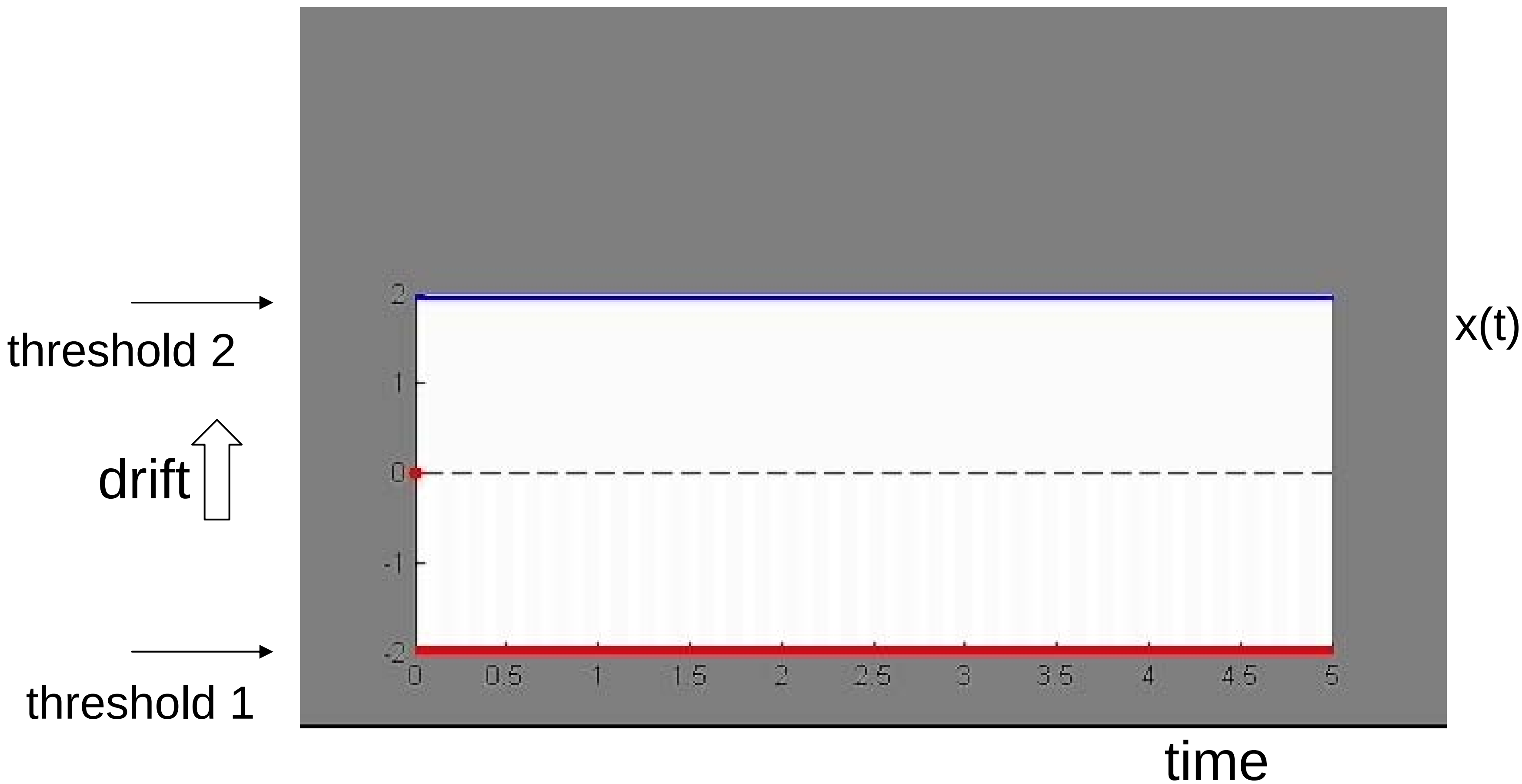
Continuous time limit of random walk is **DRIFT-DIFFUSION MODEL**

$$\frac{dx}{dt} = \pm A + \sqrt{D}\eta(t) \quad \leftarrow \quad \text{noise term}$$

(e.g. "-A" if draw from p_2, i.e. hyp. 2)

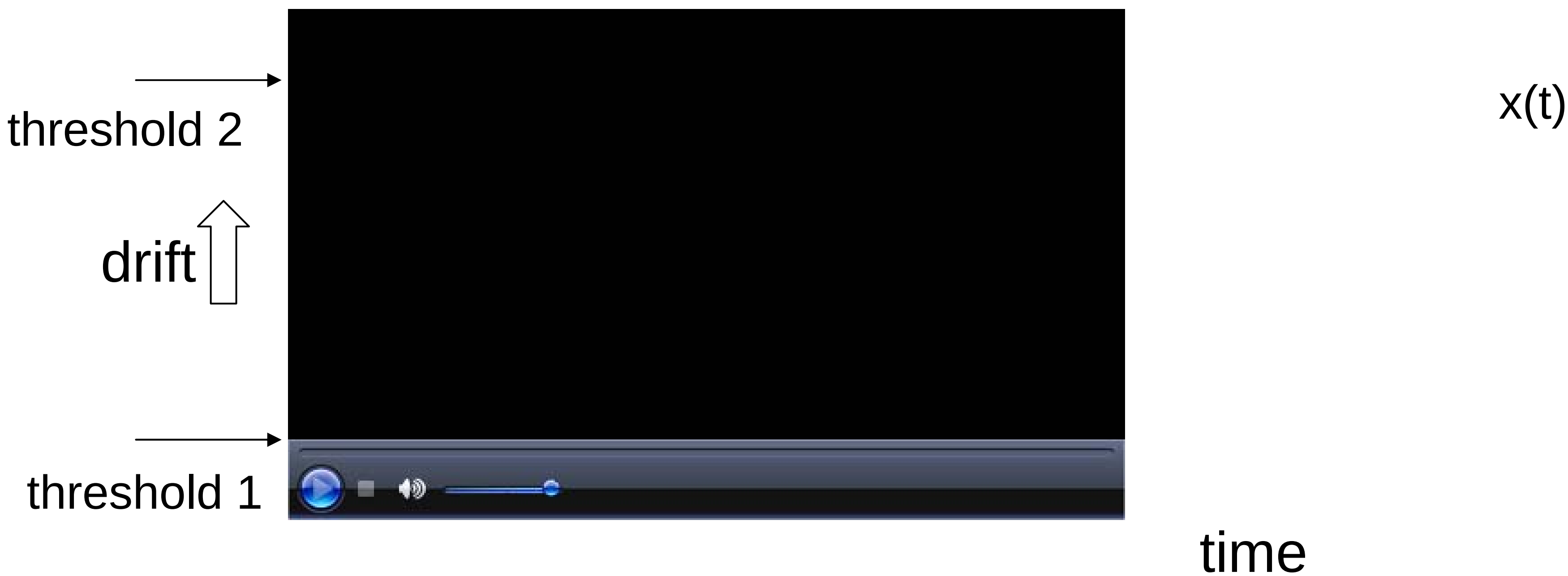
4.1. Behavioral Evidence for DD process

[Ratcliff, 1978; Ratcliff et al 1999], SPRT Used by Laming 1968. With 5-7 adjustable parameters, can match individual subject RTs well.



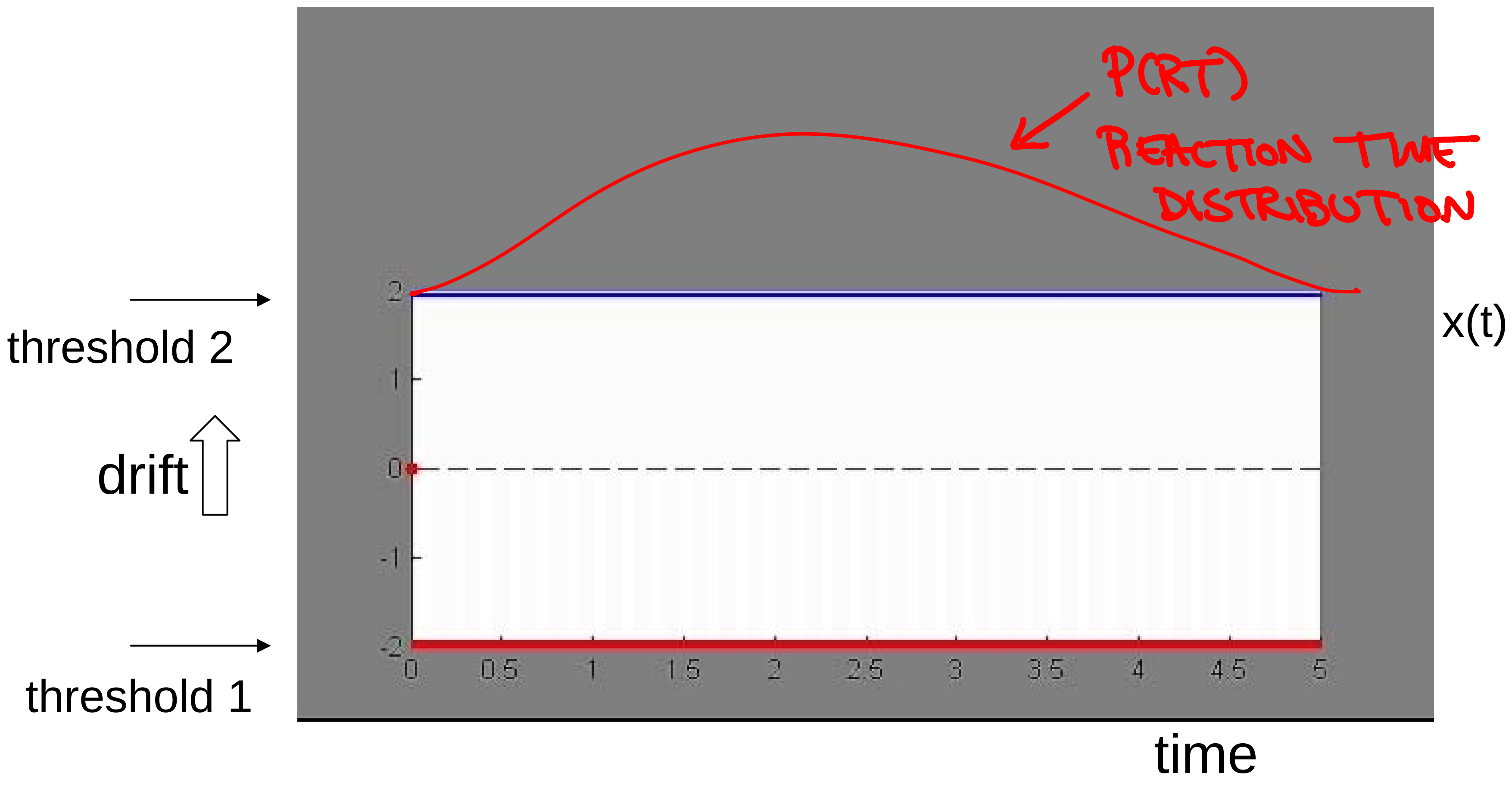
4.1. Behavioral Evidence for DD process

[Ratcliff, 1978; Ratcliff et al 1999], SPRT Used by Laming 1968. With 5-7 adjustable parameters, can match individual subject RTs well.



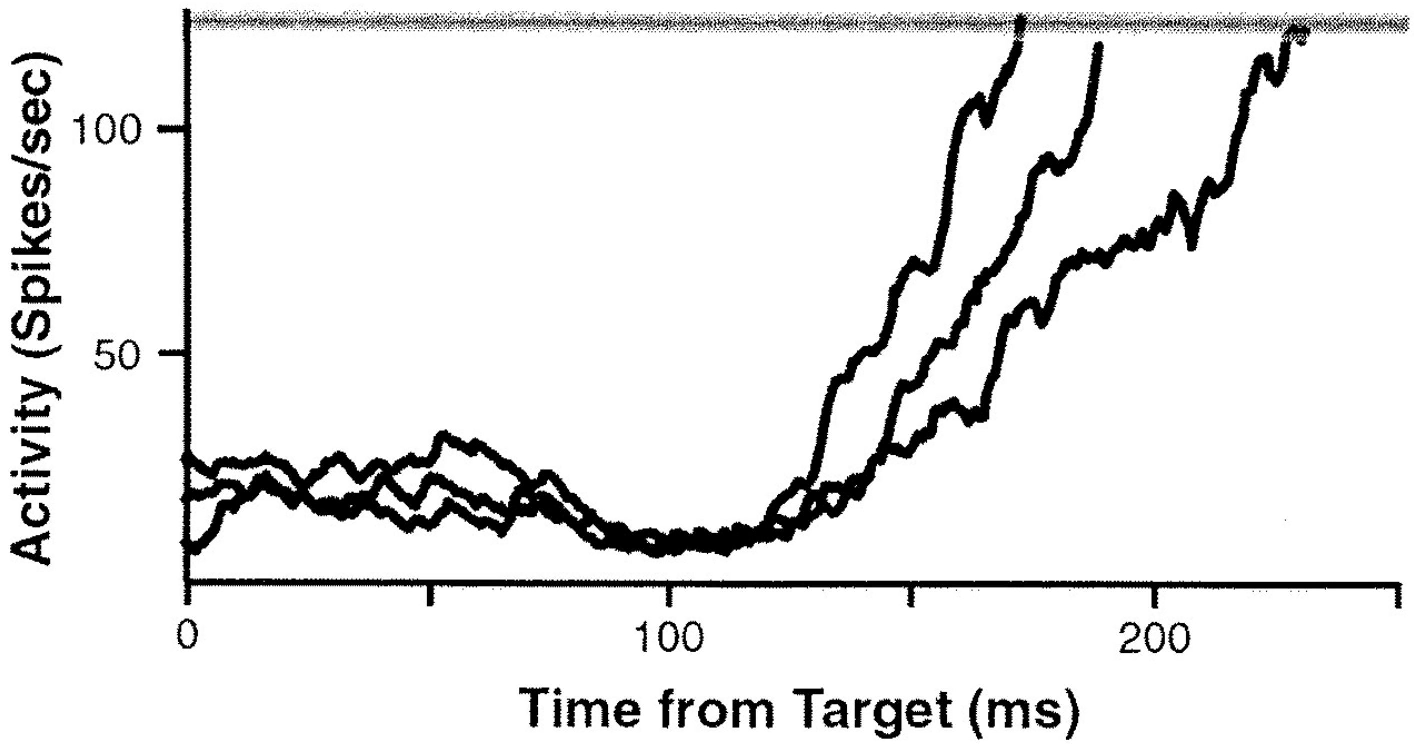
4.1. Behavioral Evidence for DD process

[Ratcliff, 1978; Ratcliff et al 1999], SPRT Used by Laming 1968. With 5-7 adjustable parameters, can match individual subject RTs well.

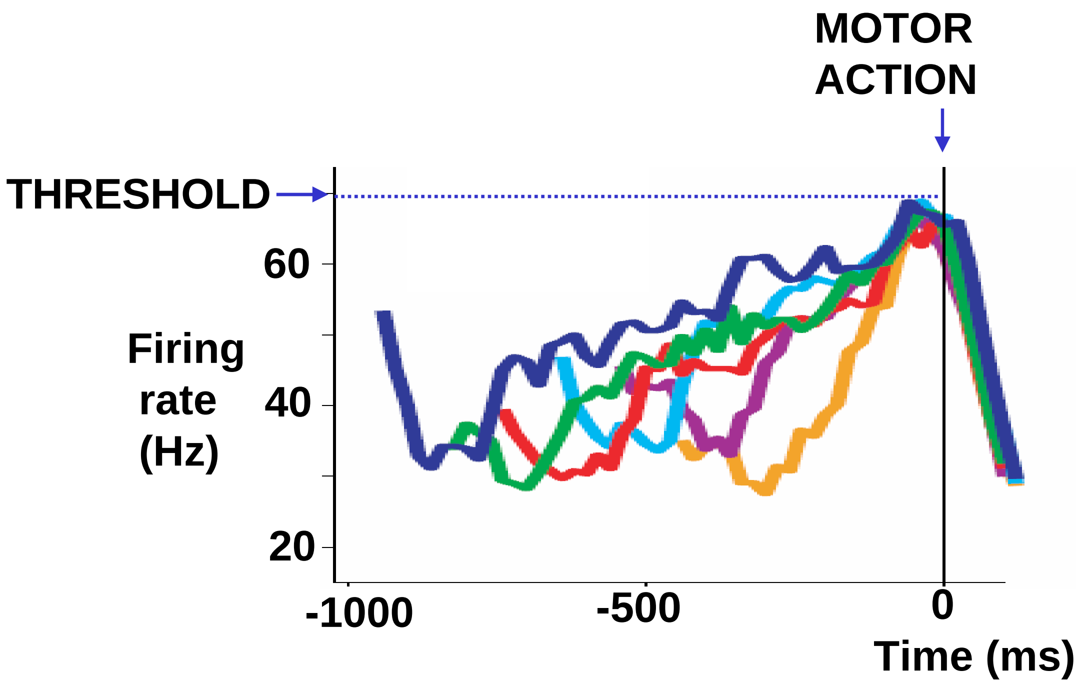


4.2. Neural Evidence for DD process

LIP and FEF neural spike rates vs. time - evidence for crossing fixed threshold prior to response (saccade).



J. Schall, V. Stuphorn, J. Brown, *Neuron*, 2002



[Roitman + Shadlen ’02]

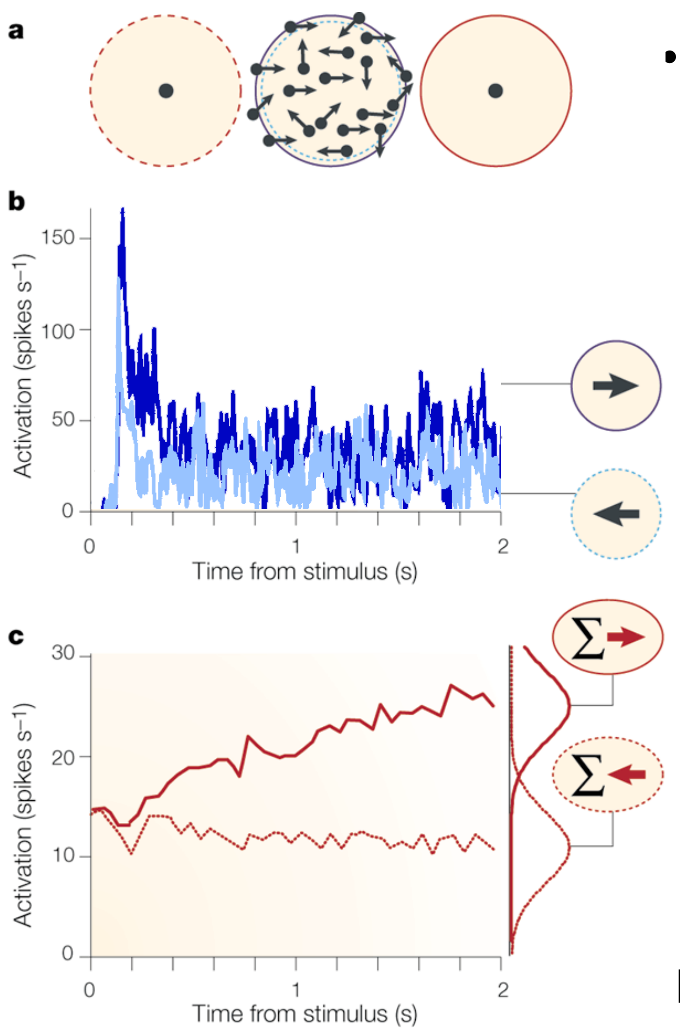
Neural basis of decisions: Shadlen, Schall, Newsome, Movshon, Gold, et al

- **Task: saccade in the direction of movement**
- **MT: encode direction of movement**

- **LIP: Integrate noisy evidence**

MT

LIP



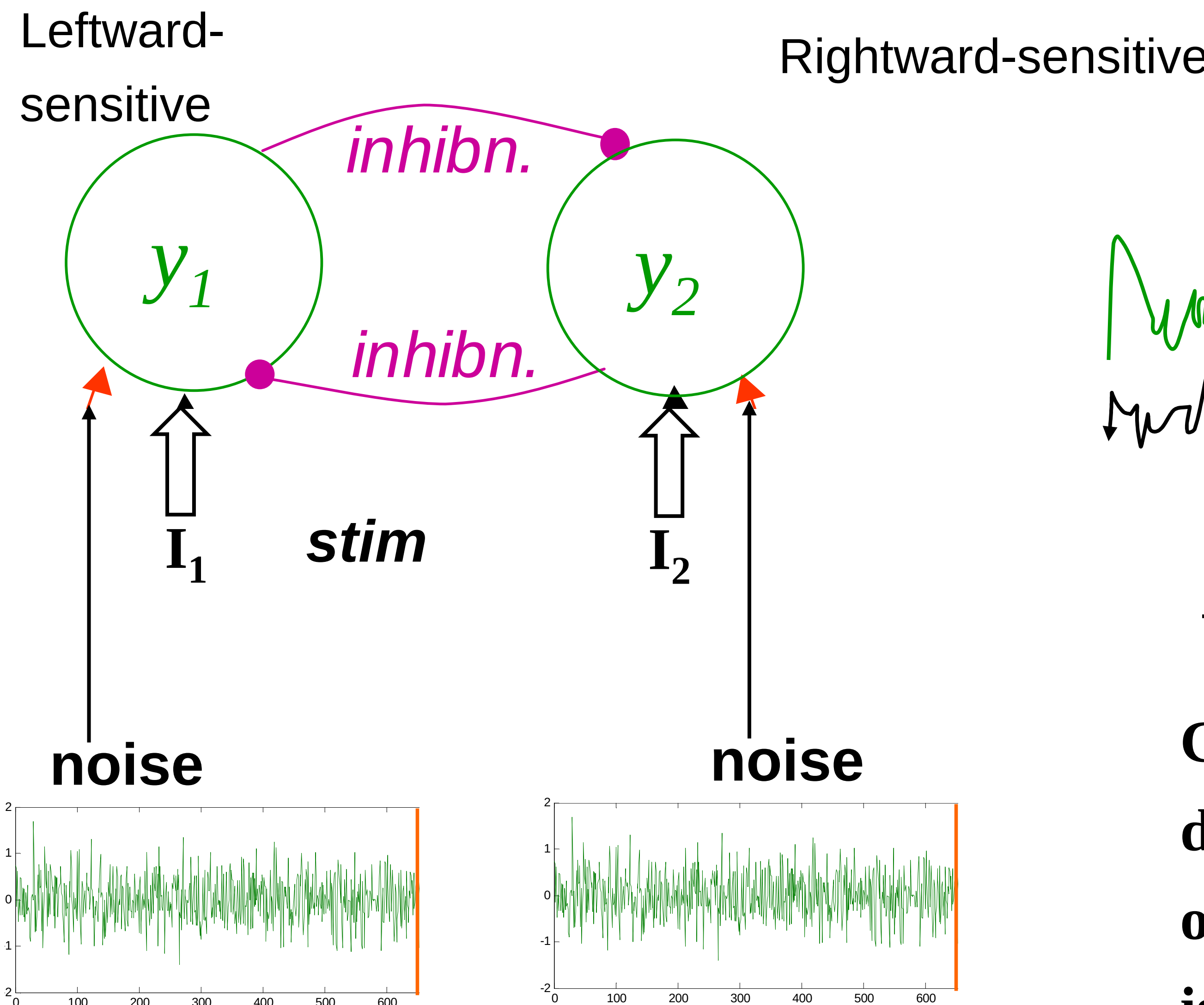
'rightward' sensitive cell

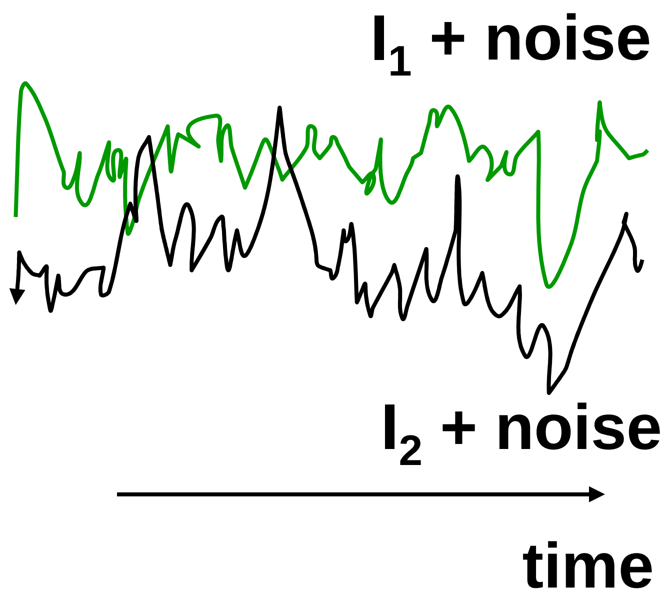
'leftward' sensitive cell

'rightward' sensitive cell

'leftward' sensitive cell

From: Schall, 2001; Shadlen & Newsome, 1996





Goal:

decide which

of I_1 or I_2

is larger, i.e. compute

$I_1 - I_2$

- **ASIDE – where does noise come from?**

Neural representation of incoming information fluctuates in time

mechanism 1: sensory scanning

ABABABBBAB

ABABBABABA

BBABABABAB

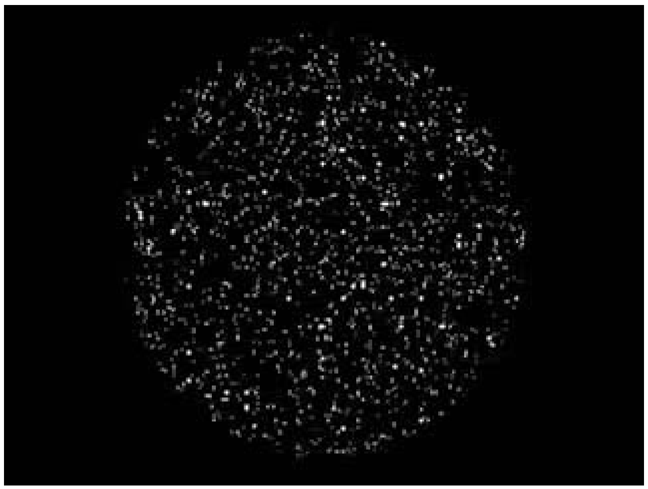
BABABABABA

BBABBABABA

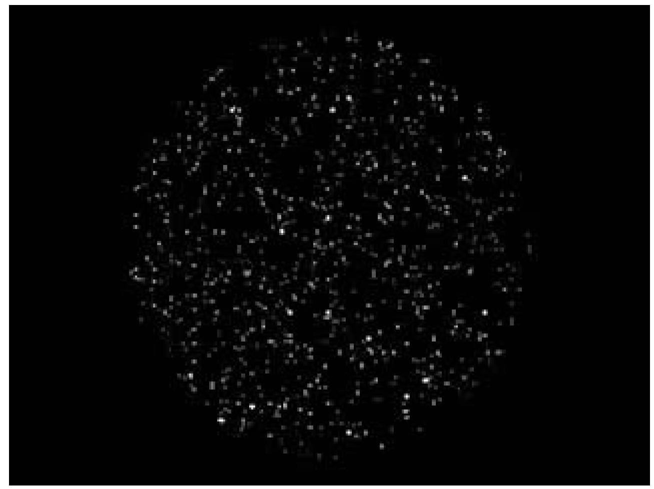
Neural representation of incoming information fluctuates in time

mechanism 2: stimulus itself fluctuates

30 % coherent



5 % coherent



Bill Newsome

Neural representation of incoming information fluctuates in time

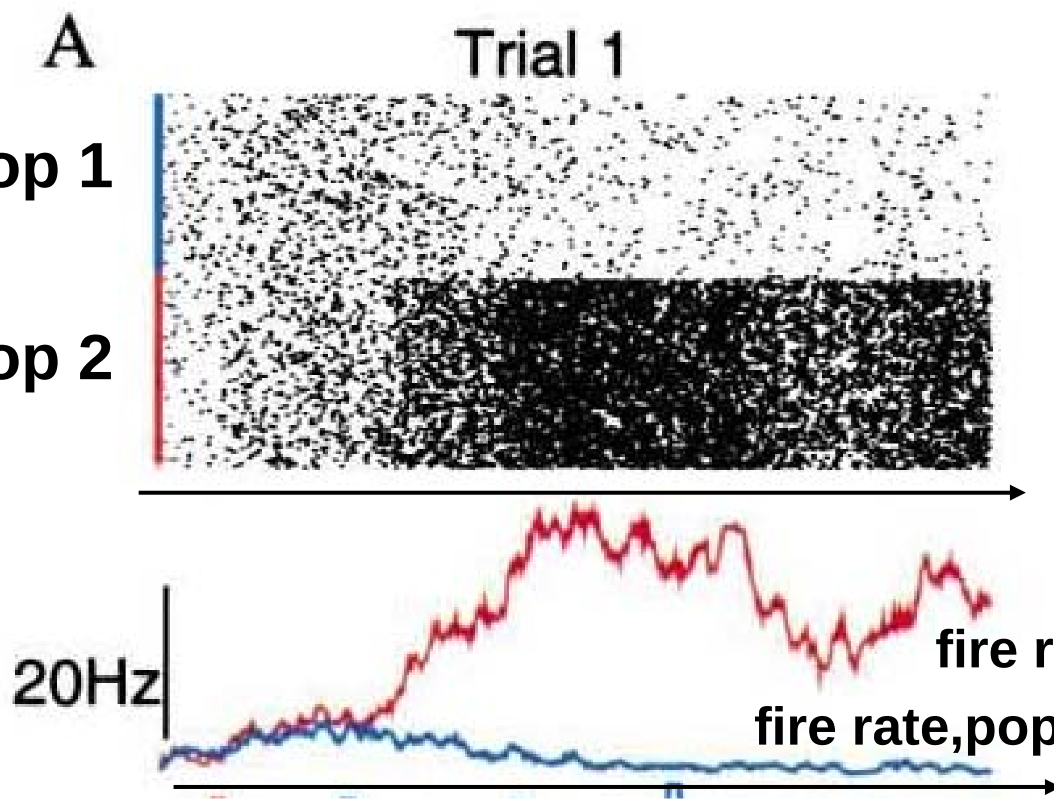
mechanism 3: ***intrinsic fluctuations due to finite-size effects***

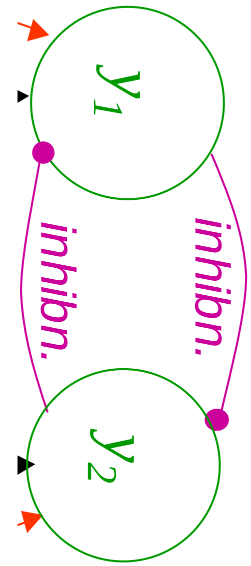
internal dynamics of neural populations are noisy …

(Wang, 2002)

spikes in pop 1

spikes in pop 2





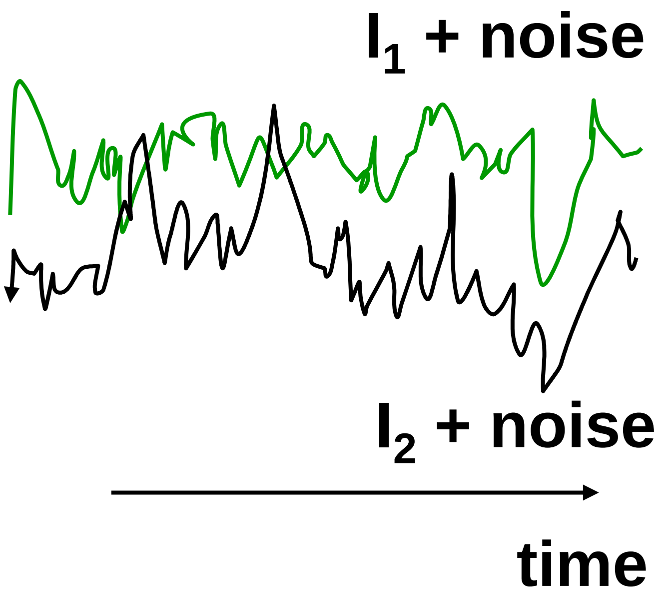
fire rate,pop 2 (winner)

fire rate,pop 1

time (approx 3 sec elapsed)

(AND may be correlated: Zohary et al, Science 1996)

Back to problem at hand …



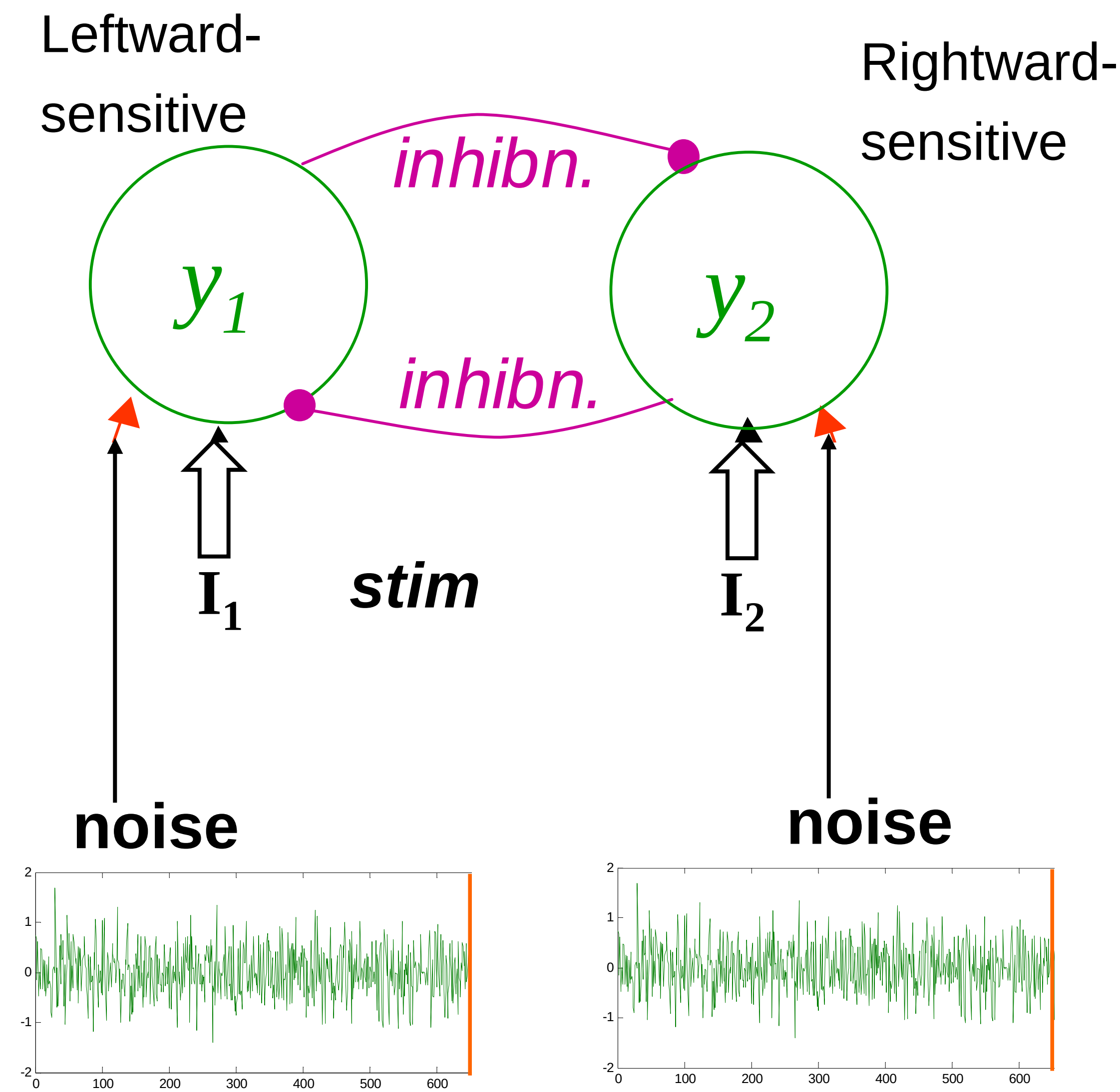
Goal:

decide which

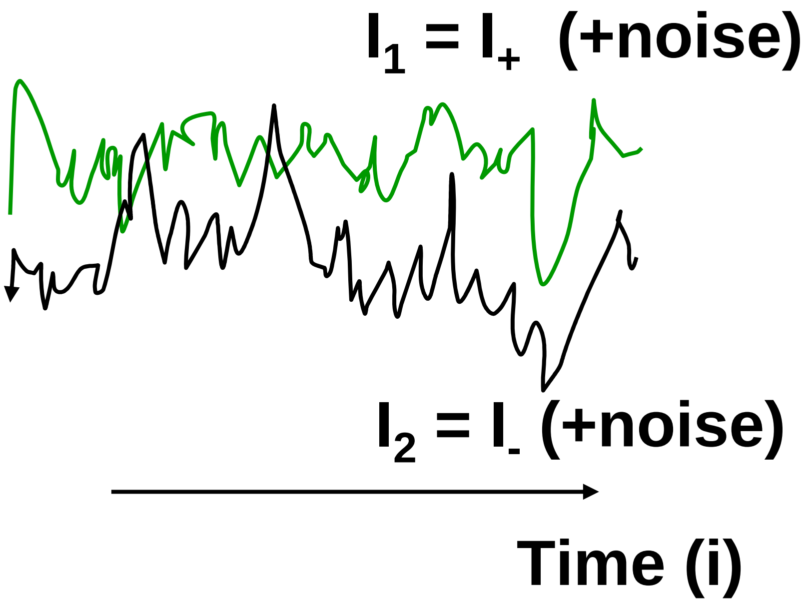
of I_1 or I_2

is larger, i.e. compute

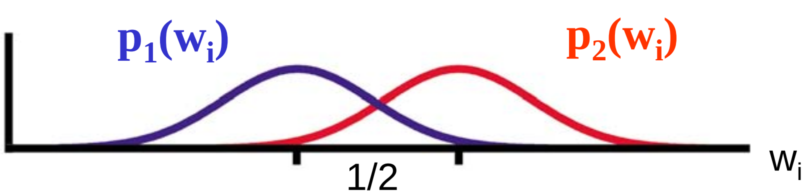
$I_1 - I_2$



Back to problem at hand …



Would like to interpret as SPRT



w_i, increment of input over $\Delta\ t_i$

Issue: input is TWO-D.

Follow Gold/Shadlen, TINS 2001

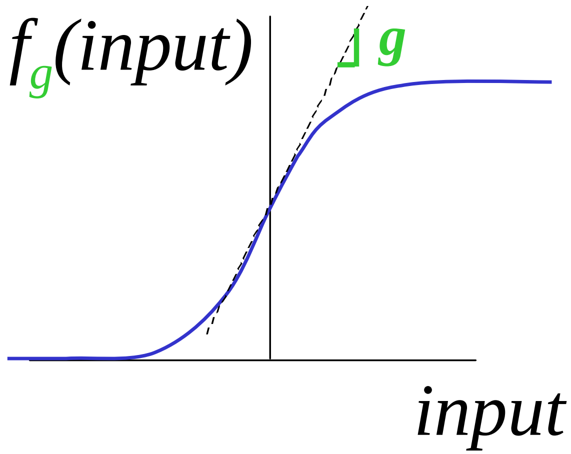
to resolve.

Dynamics?

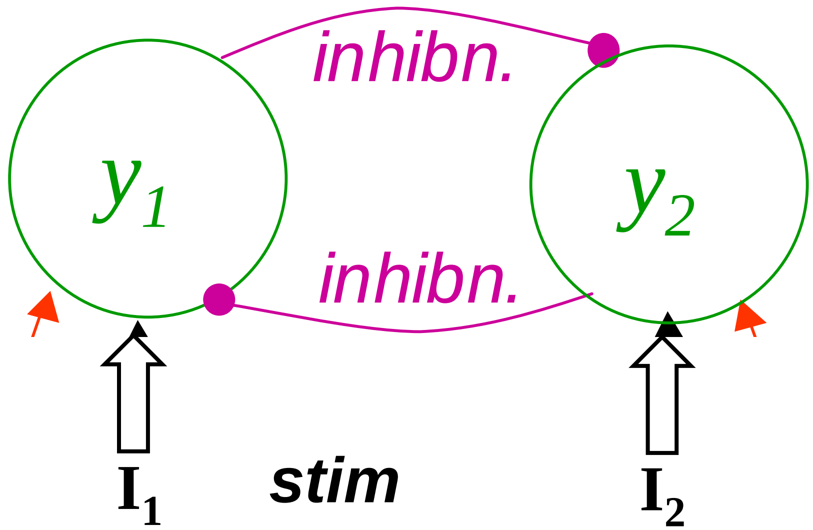
- **Think of neural "units…" described by firing rates y**

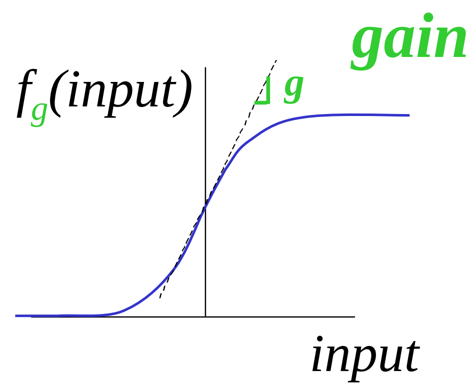
  which approach equilibrium rates *f(input)* with time constant τ_m.

$$\tau_m \frac{dy}{dt} = -y + f\,(input)$$



Two inhibiting neural populations



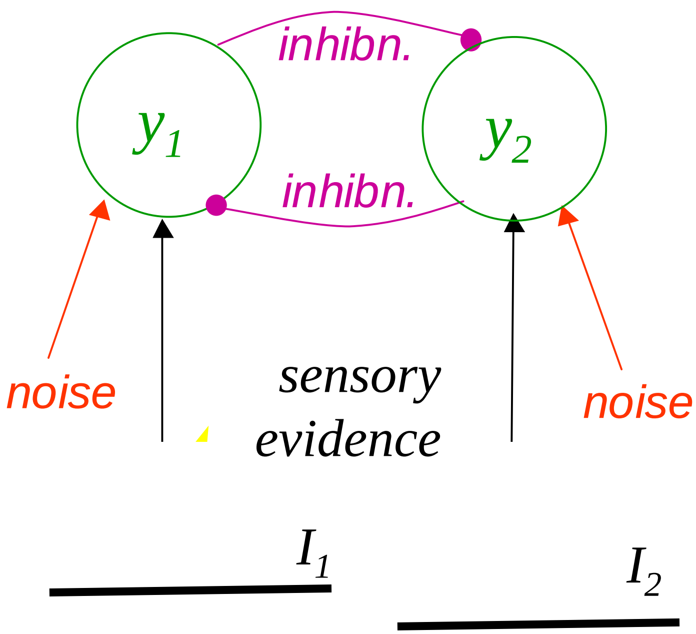


Firing rates (*y_1, y_2*) of competing neural pops...
approach *f(input)* with time constant τ_m.

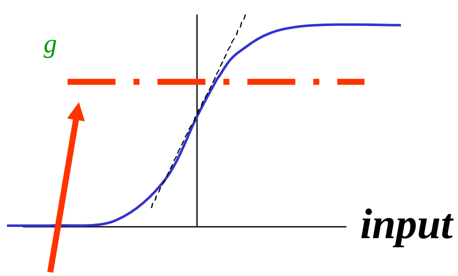
$$\boxed{\begin{aligned}\tau_m \frac{dy_1}{dt} &= -y_1 + f\left(-\beta y_2 + I_1(t)\right)\\ \tau_m \frac{dy_2}{dt} &= -y_2 + f\left(-\beta y_1 + I_2(t)\right)\end{aligned}}$$

Two-alternative choice task

Firing rates (y_1, y_2) of competing neural pops...

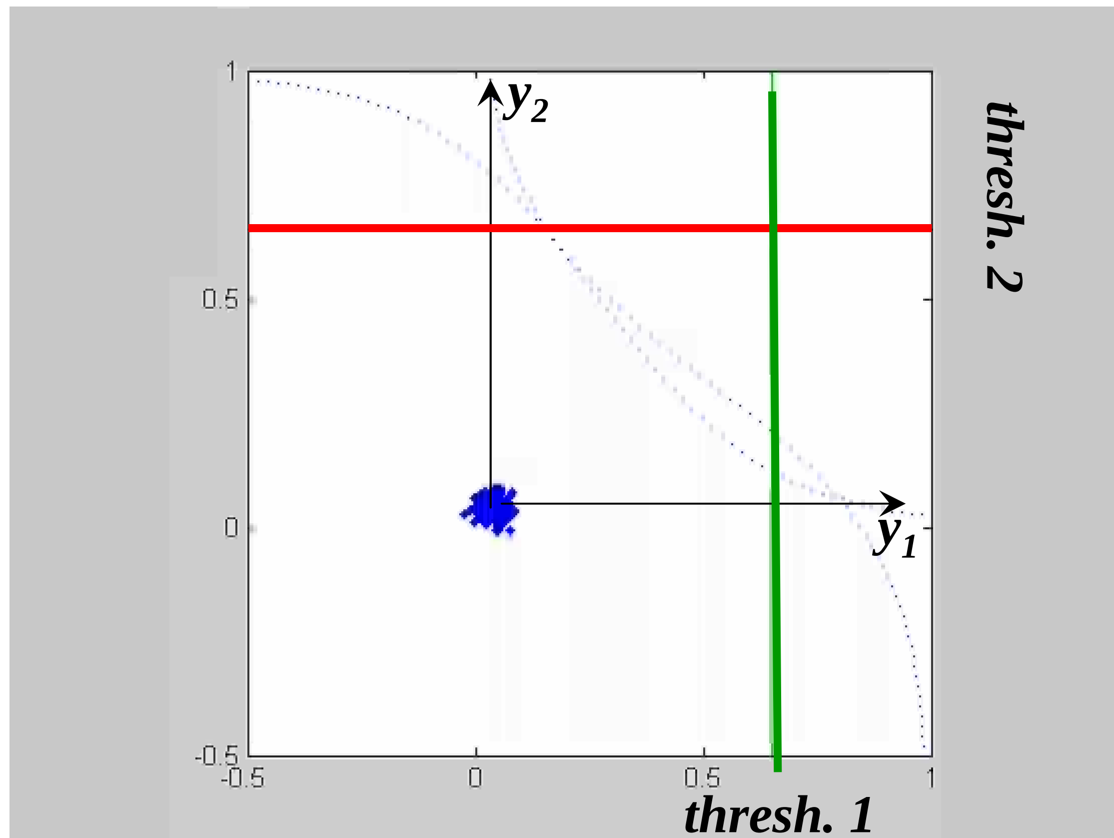


approach functions of their inputs,



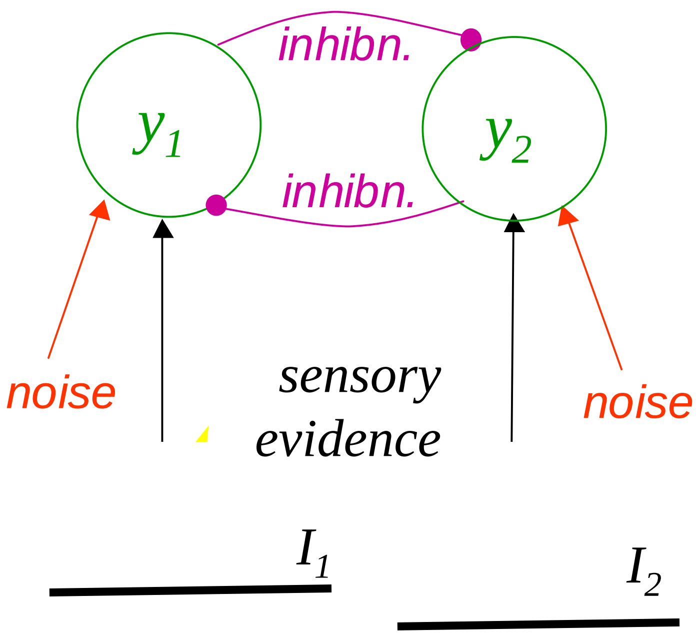
crossing decision thresholds along the way.

Decision 1 or 2 made when firing rate y_1 or y_2 *crosses threshold*

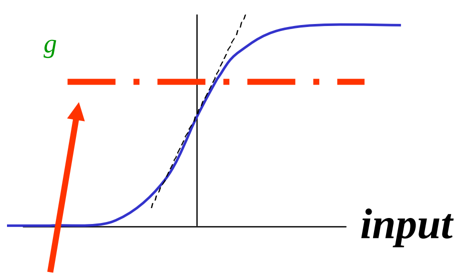


Two-alternative choice task

Firing rates (y_1, y_2) of competing neural pops...

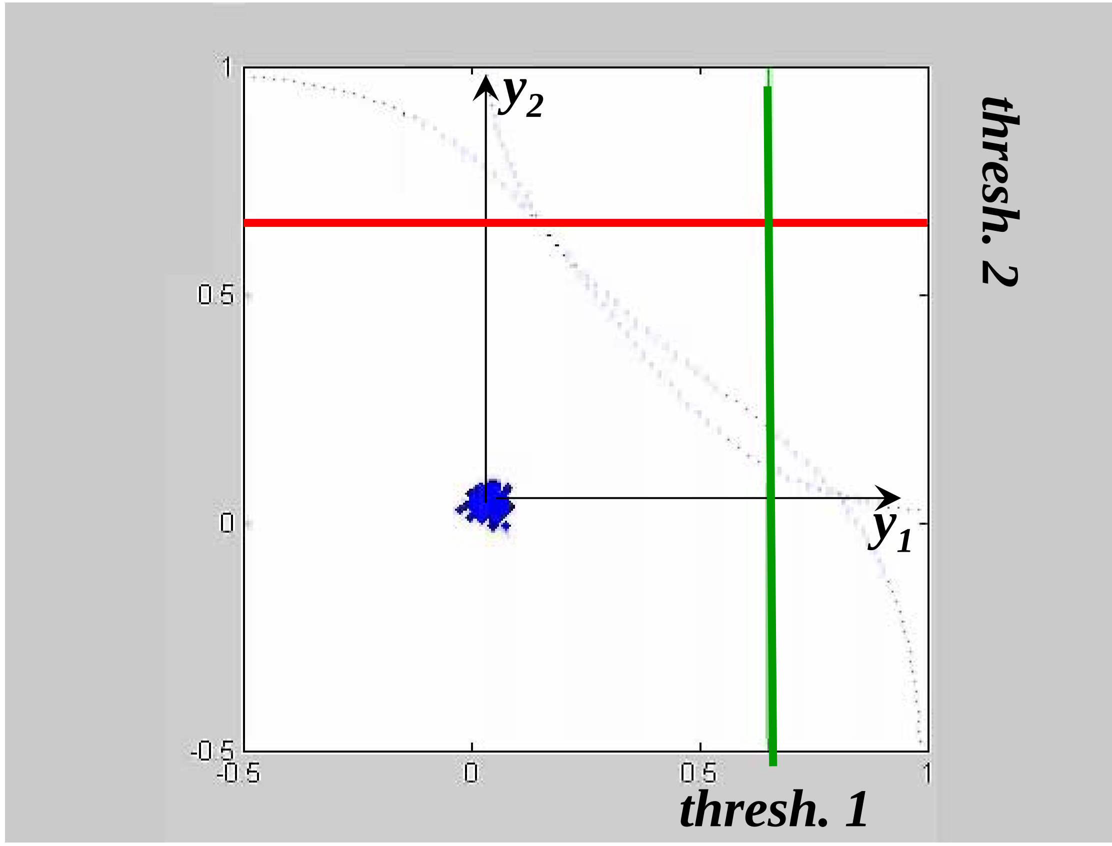


approach functions of their inputs,

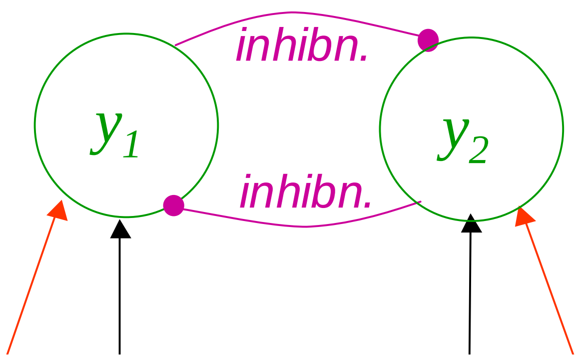


crossing decision thresholds along the way.

Decision 1 or 2 made when firing rate y_1 or y_2 *crosses threshold*



Firing rates (y_1, y_2) of competing neural pops...



$$\frac{dy_1}{dt} = -y_1 + f\left(-\beta y_2 + I_1 \quad + \eta(t)\right)$$
$$\frac{dy_2}{dt} = -y_2 + f\left(-\beta y_1 + I_2 \quad + \eta(t)\right)$$

Linearize:

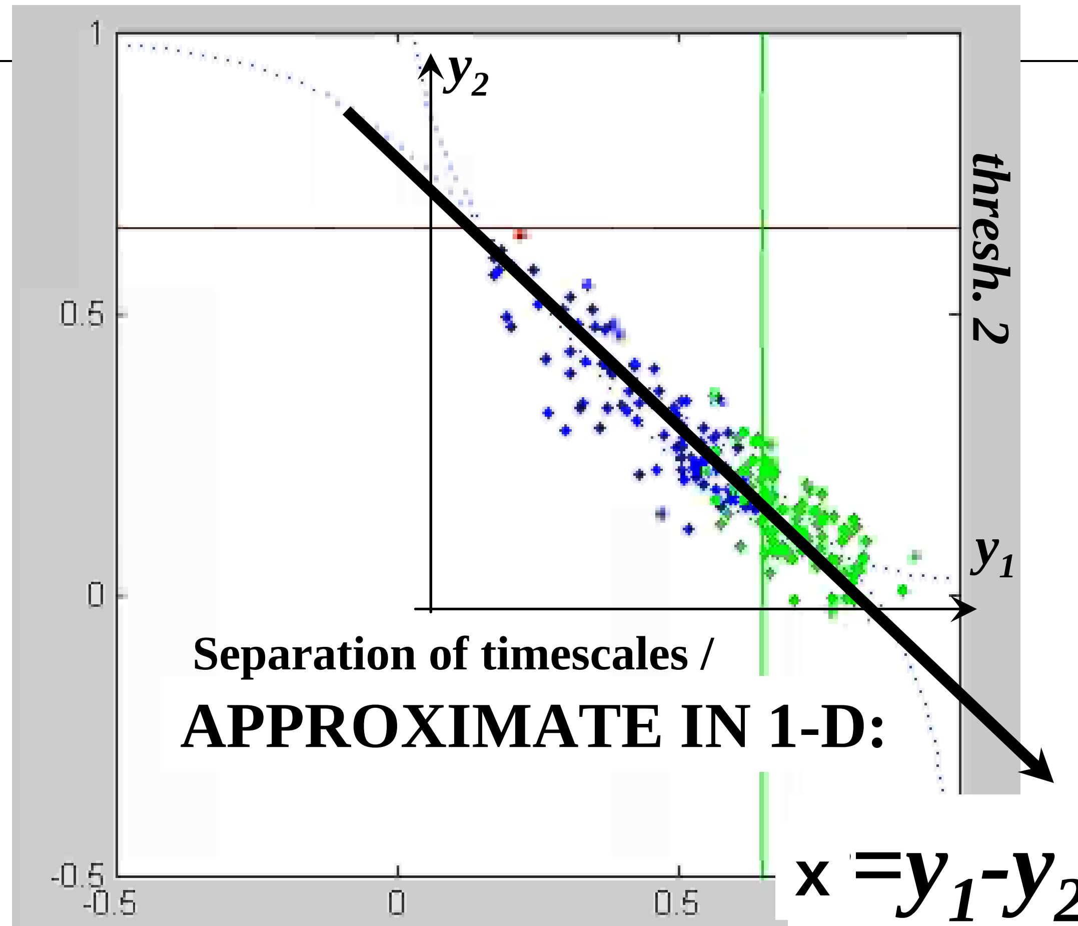
$$\frac{dy_1}{dt} = -y_1 + g * \left(-y_2 + I_1 \quad + \eta_1\right)$$
$$\frac{dy_2}{dt} = -y_2 + g * \left(-y_1 + I_2 \quad + \eta_2\right)$$

Subtract:

$$\frac{d\,\mathrm{x}}{dt} = -\,\mathbf{x} + g\,(\,\mathrm{x} + I_1 - I_2) + g\eta(t) \qquad (1)$$

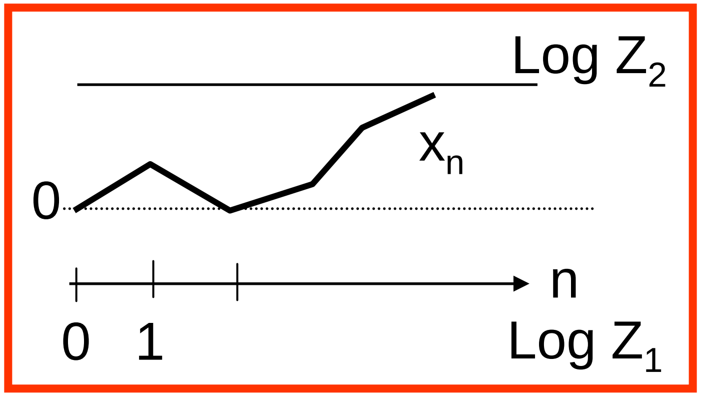
TUNE GAIN to $g = 1$, define drift $A = I_1 - I_2$

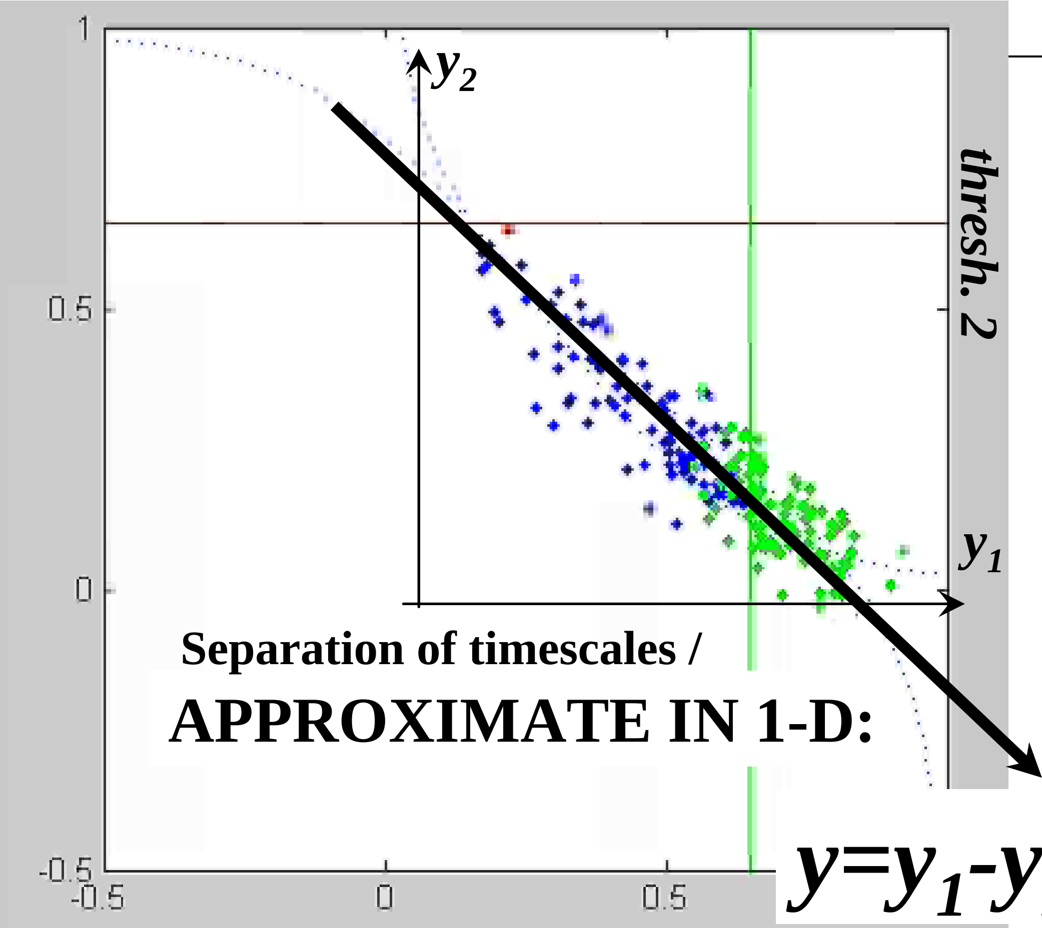
$$\frac{d\,\mathrm{x}}{dt} = A + \eta(t) \qquad (2)$$



recover SPRT, optimal decision strategy

Firing rates (y_1, y_2) of competing neural pops...





$$\frac{d\,x}{dt} = -\,x + g\,(\,x + I_1 - I_2) + g\eta(t) \qquad (1)$$

TUNE GAIN to $g = 1$, define drift $A = I_1 - I_2$

$$\frac{d\,x}{dt} = A + \eta(t) \qquad (2)$$

recover SPRT,

optimal decision strategy