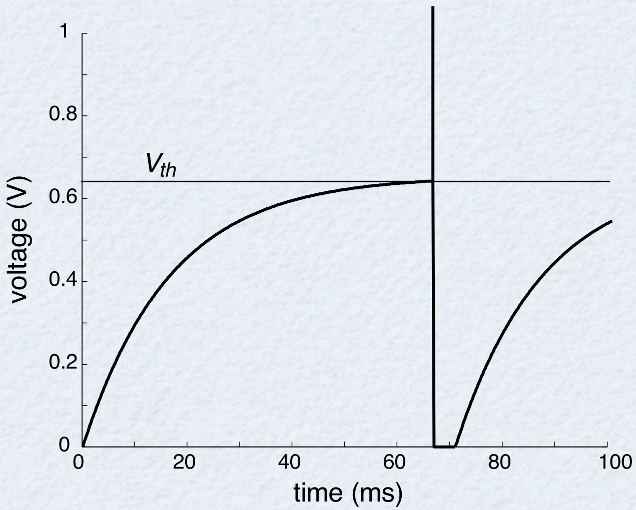
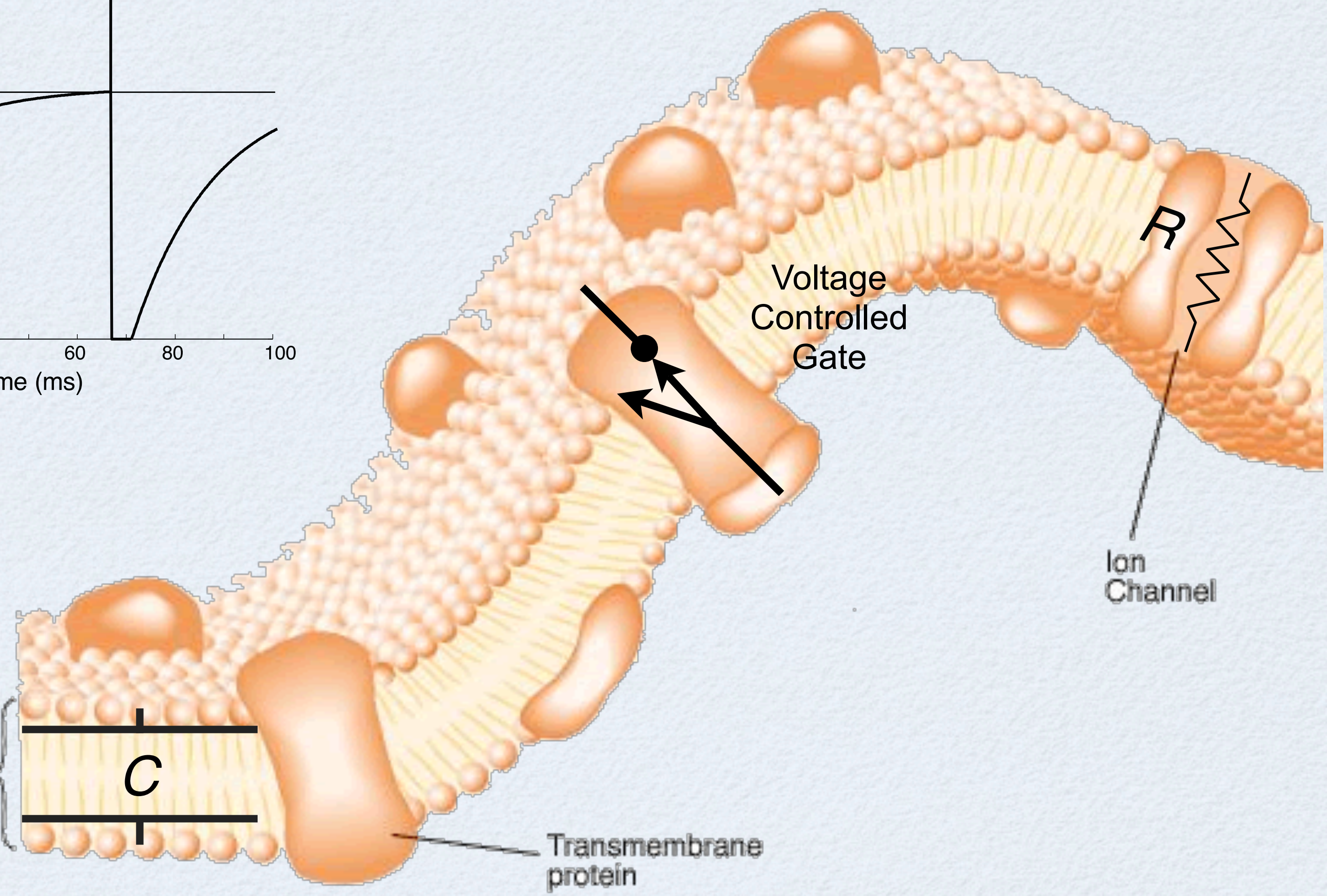
Temporal Representation in Neurons

Leaky integrate-and-fire

- LIF is a good standard because:
    - it is simple
    - it produces spikes
    - it is known to be a limiting case of more complex models
    - the parameters in the model map onto known properties of real neurons

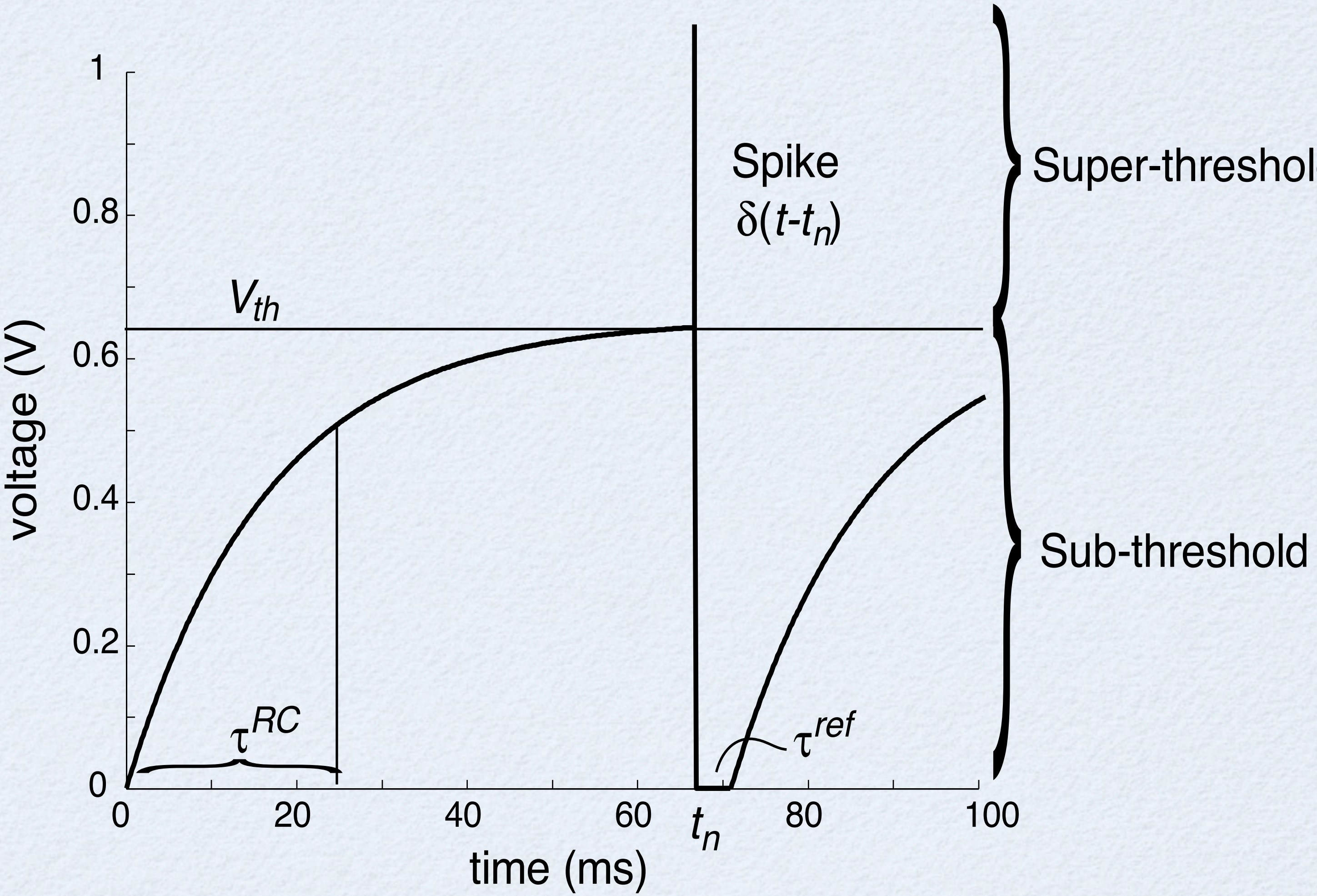
Outside the Neuron



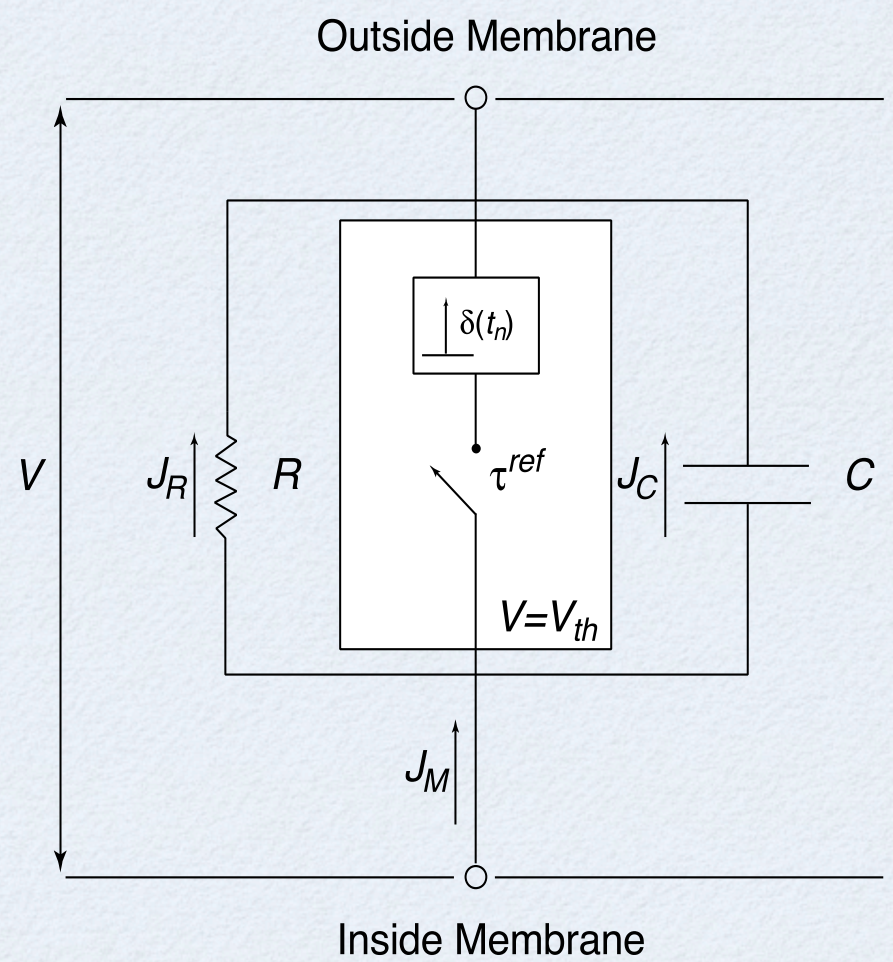


current from dendrites

Inside the Neuron



1
0.8
0.6
0.4
0.2
0
voltage (V)
Vth
Spike
δ(t-tn)
τRC
τref
0
20
40
60
tn
80
100
time (ms)
Super-threshold
Sub-threshold


Outside Membrane
δ(t_n)
τ^ref
V
J_R
R
J_C
C
V=V_th
J_M
Inside Membrane

LIF and bio-plausibility

- the bilipid cell membrane acts like a capacitor
- has a passive flow of ions through the cell membrane ion channels -- a `leak' current in R
- neurons elicit a stereotypical (spike) when the soma voltage passes a threshold
- the spike cannot be repeated for about 1ms after a spike (absolute refractory period).
- for molecular biology details see: http://soma.npa.uiuc.edu/courses/bio303/ch2.html.

Derivation of LIF model

- The current across the membrane can be found by differentiating V=Q/C to give

$$J_C = C\frac{dV}{dt}$$

- The ionic current, J_R, accounts for this passive leak of charge. Ohm's law gives

$$J_R = \frac{V}{R}.$$

- Input from dendrites results in J_M, a membrane current which is the result of J_{bias} and J_{drive}. Kirchoff's law means

$$J_M = J_C + J_R$$

- So, for $\tau^{RC} = RC$

$$\begin{aligned} J_M &= C\frac{dV}{dt} + \frac{V}{R} \\ \frac{dV}{dt} &= -\frac{1}{\tau^{RC}}\left(V(t) - J_M(t)R\right) \end{aligned}$$

Spiking

- This first-order ODE only describes the passive behaviour.

- Once V crosses the neuron threshold, V_{th}, the gate closes for τ^{ref} and a delta function, $\delta(t_n)$, spike is generated

- So τ^{ref} is the absolute refractory period and the spike is a delta function

Solving the ODE

- Start by assuming a solution of the form

$$V(t) = F(t)e^{-t/\tau^{RC}}$$

- Substituting gives

$$\frac{d}{dt}\left(F(t)e^{-t/\tau^{RC}}\right)$$

$$= -\frac{1}{\tau^{RC}}\left(F(t)e^{-t/\tau^{RC}} - J_M(t)R\right)$$

- Solving the ODE with that solution form gives:

$$V(t) = \frac{R}{\tau^{RC}} \int_0^t e^{-(t-t')/\tau^{RC}} J_M(t')dt'.$$

- I.e. the voltage right now (at *t*) depends on all past current input, $J_M(t')$, where each input is weighted by a function that exponentially decays as it gets further away (in the past) from the current time (memory).

- Evaluting this integral for constant J_M gives:

$$V(t) = J_M R \left(1 - e^{-t/\tau^{RC}}\right).$$

LIF Response Function

- Firing rate as a function of time to threshold is

$$a(t_{th}) = \frac{1}{t_{th} + \tau^{ref}}$$

- We can find t_{th}

$$\begin{aligned} V_{th} &= J_M R \left(1 - e^{-t_{th}/\tau^{RC}}\right) \\ t_{th} &= -\tau^{RC} \ln \left(1 - \frac{V_{th}}{J_M R}\right). \end{aligned}$$

- Substituting and re-arranging gives

$$a(x) = \frac{1}{\tau^{ref} - \tau^{RC} \ln \left(1 - \frac{J_{th}}{J_M(x)}\right)}$$

$$Recall : J_M(x) = \alpha x + J^{bias}$$

Weaknesses of the LIF model

- they are point neurons
- time courses of different ion conductances are not modeled
- many physiologically unrealistic assumptions (R constant, J_M static, V_{th} static, no adaptation, etc.)
- One standard complaint with LIF models is that dendrites are linear, but the way $J_M(x)$ is determined is not part of the LIF model

Temporal Coding

- Debate between two views of neural coding:
- 1. rate codes: The firing rate (over about 100ms) of the neuron is what carries information.
- 2. timing codes: The *precise* pattern of spike generation carries information.
- So defined, both contradicted by some evidence
- Luckily, we don't have to choose either kind of code, given our characterization.

Temporal encoding

- Temporal encoder, just like population encoder

$$a(x(t)) = G[J(x(t))] = \sum_n \delta(t - t_n),$$

- Where

$$J(x(t)) = \alpha \tilde{\phi} x(t) + J^{bias}$$

- Need a temporal decoder, $h(t)$

Temporal Decoding

- In short, we want to find the optimal linear h(t) in

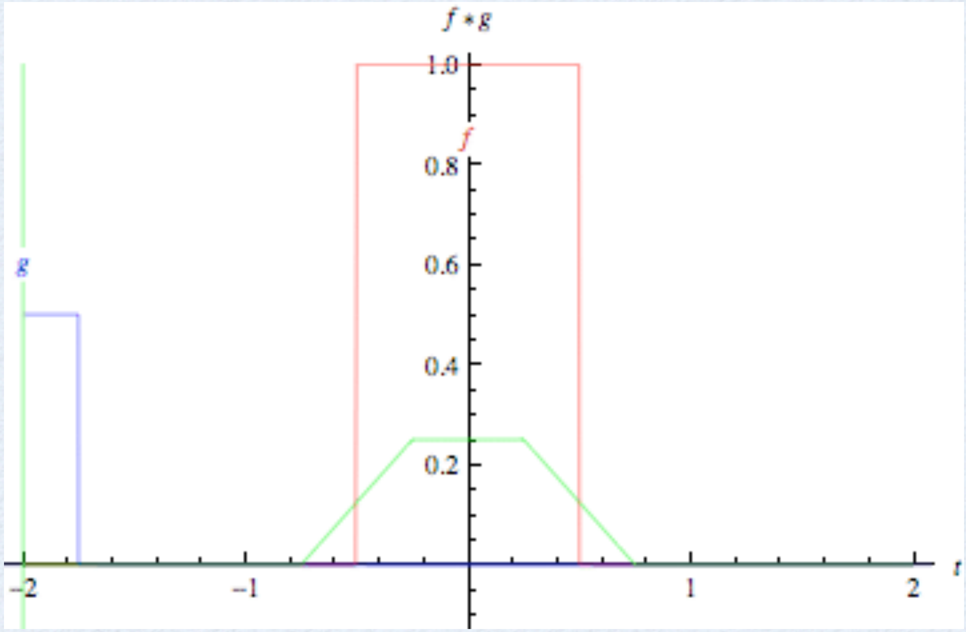
$$\hat{x}(t) = \int_0^T h(t - t') \sum_n \delta(t' - t_n) dt'$$

- I.e., we're trying to find the impulse response of a linear system that would give the best estimate of x(t) given the spike train

- Evaluating gives: $$\hat{x}(t) = \sum_n h(t - t_n).$$

Convolution

$$r(t) = f(t) * g(t) = \int_{-\infty}^{\infty} g(t-\tau) f(\tau) d\tau$$



- The response of a linear system (which could be a decoder), is just such a convolution of the impulse response with the input

Neural convolutions

$$\hat{x}(t) = \int_0^T h(t-t') \sum_n \delta(t'-t_n)dt'$$

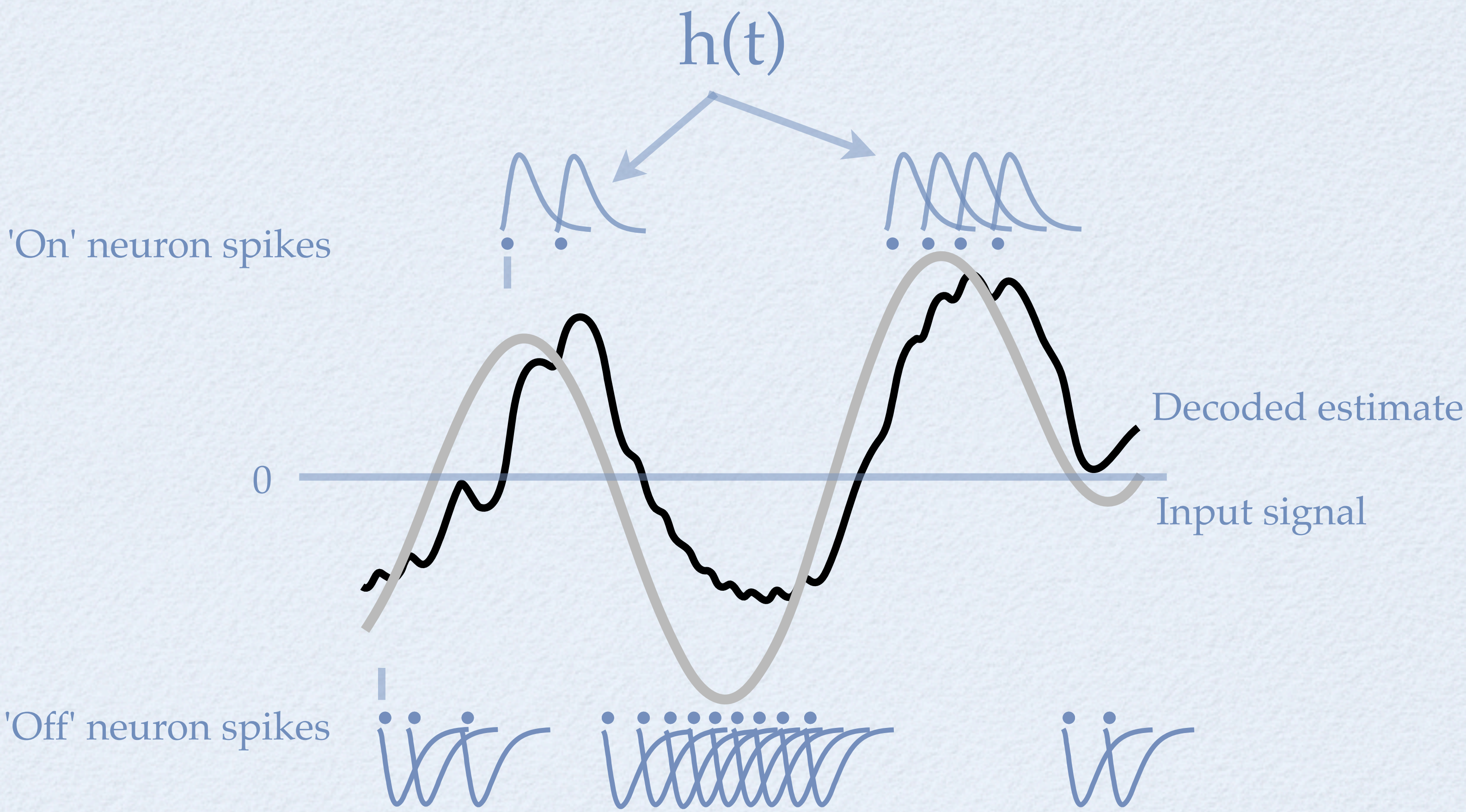
$$= \sum_n \int_0^T h(t-t')\delta(t'-t_n)dt'$$

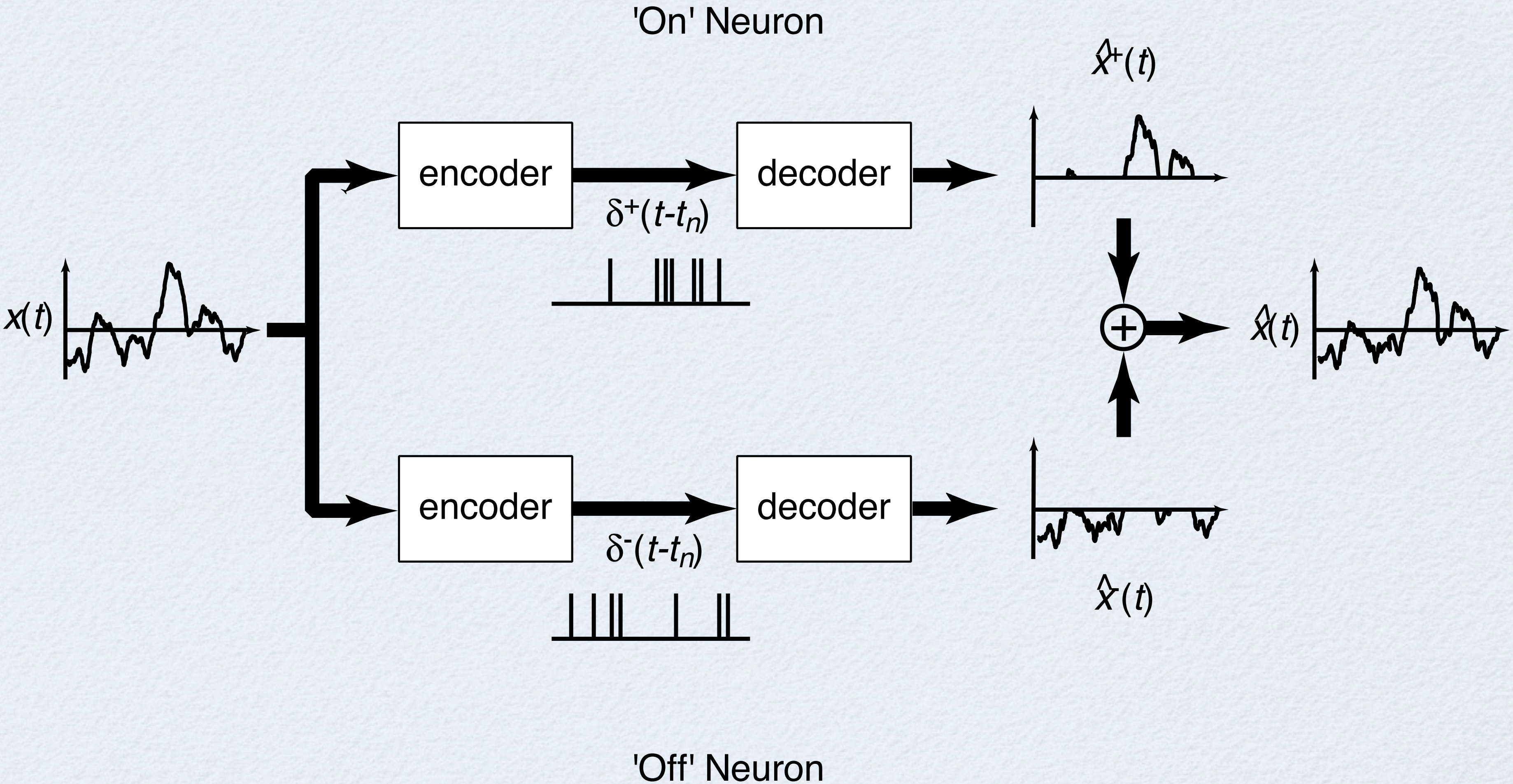
$$\text{for one n } = \int_0^T h(t-t')\delta(t'-s)dt'$$

$f(t\text{-}\tau)$ $g(\tau)$

$$V(t) = \frac{R}{\tau^{RC}} \int_0^t e^{-(t-t')/\tau^{RC}} J_M(t')dt'.$$

Temporal Representation




'On' Neuron
$\hat{x}^{+}(t)$
encoder
decoder
$\delta^{+}(t-t_n)$
$x(t)$
$\hat{x}(t)$
encoder
decoder
$\delta^{-}(t-t_n)$
$\hat{x}^{-}(t)$
'Off' Neuron

Finding h(t)

- We need to define our domain ('x' in the population case)

- How about a Fourier series

$$x(t;\mathbf{A}) = \sum_{n=-(N-1)/2}^{(N-1)/2} A_n e^{i\omega_n t}$$

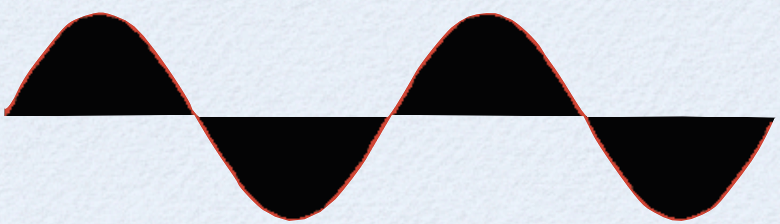
- Where the coefficients are chosen somehow

$$\rho(\mathbf{A}) = \prod_{n=0}^{(N-1)/2} \rho(A_n) = \prod_{n=0}^{(N-1)/2} \frac{1}{\sqrt{2\pi\mathcal{P}_n}} e^{-A_n^2/2\mathcal{P}_n}$$

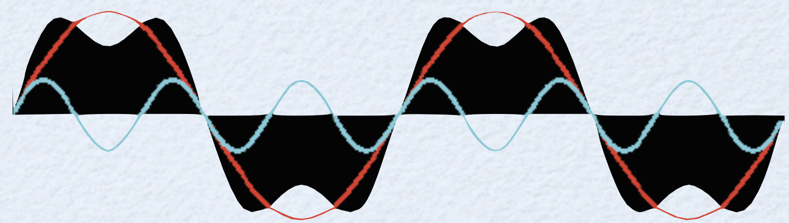
Fourier transform

- Any function can be expressed as a series of sines and cosines

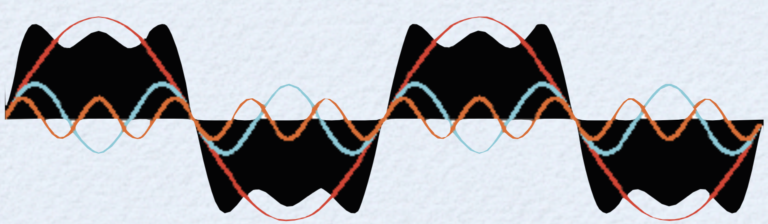
f



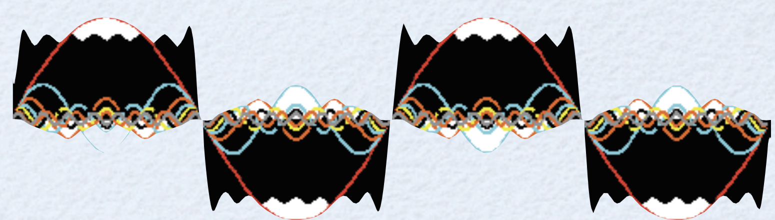
f+3f



f+3f+5f



f+3f+5f+...+15f



Fourier series

$$f(t) = \frac{1}{2}a_0 + \sum_{n=1}^{\infty} a_n \cos(nt) + \sum_{n=1}^{\infty} b_n \sin(nt)$$

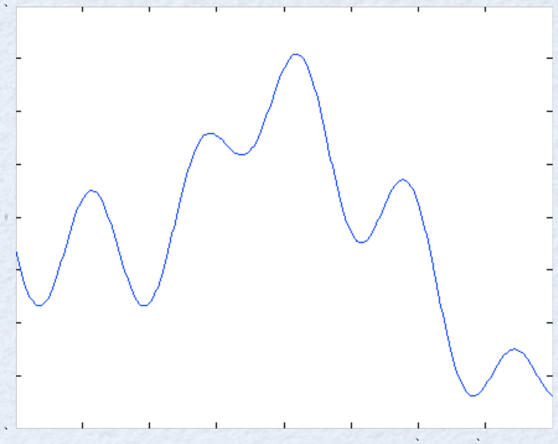
- since $\quad e^{int} = \cos(nt) + i\sin(nt)$

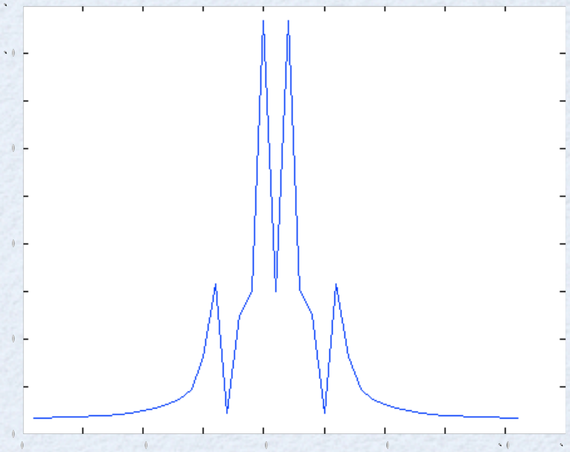
$$f(t) = \sum_{n=-\infty}^{\infty} A_n e^{int}$$

- A_n are complex
- A_n are complex conjugates around zero if $f(t)$ is real.

An example

- y = sin(x) +2*cos(2*x)+ .5*sin(5*x)+cos(8*x);
- plot(x,y)
- plot(fftshift(abs(fft(y))))





Minimizing

- So, we want to minimize:

$$\left\langle [x(t;\mathbf{A}) - h(t) * R(t;\mathbf{A})]^2 \right\rangle$$

- In frequency space:

$$E = \left\langle \frac{1}{2\pi} \left| x(\omega;\mathbf{A}) - h(\omega) R(\omega;\mathbf{A}) \right|^2 \right\rangle_{\mathbf{A},\omega}$$

Monte Carlo

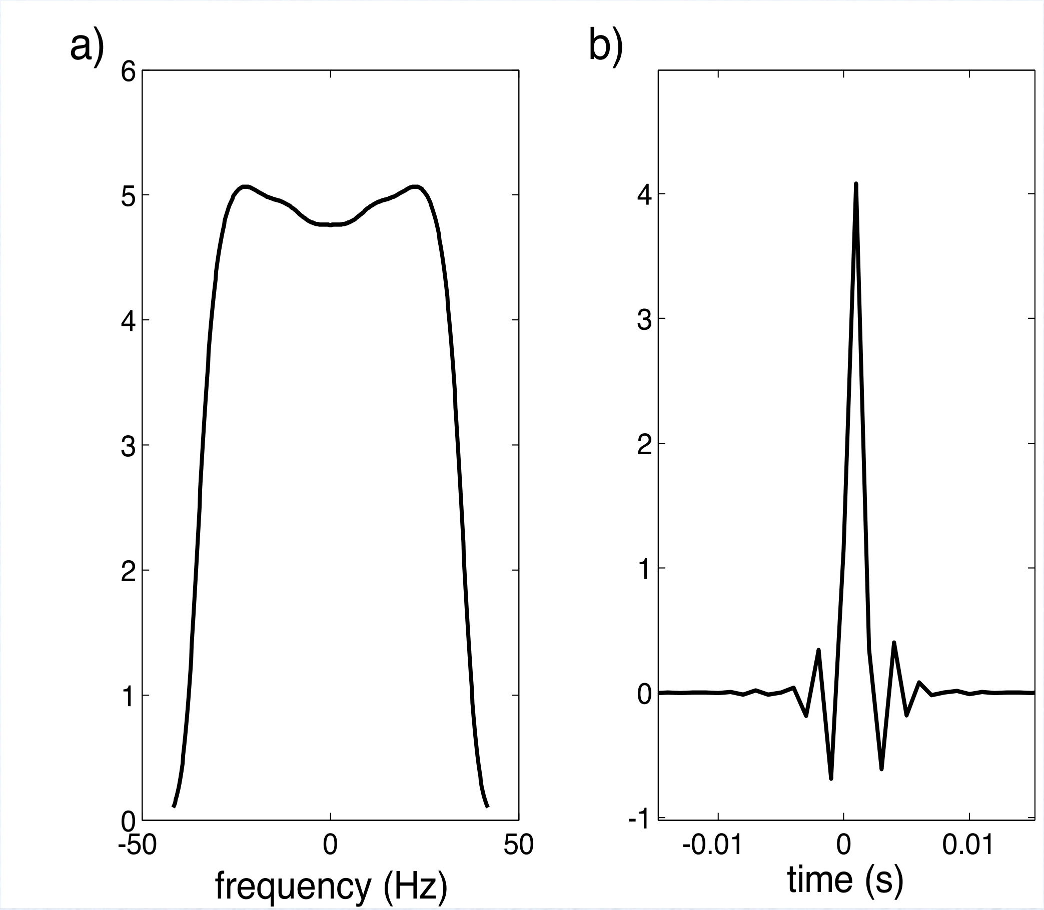
- We can't solve this directly, so we estimate

$$E = \frac{1}{N^{\alpha}} \sum_{\alpha} \sum_{n} \left| A^{\alpha}(\omega_n) - h(\omega_n) R(\omega_n; \mathbf{A}^{\alpha}) \right|^2 .$$

- with a bunch of samples (i.e., windowing)
- This allows us to solve for the decoder

$$h(\omega) = \frac{\langle A(\omega) R^*(\omega) \rangle_{\mathbf{A}}}{\left\langle \left| R(\omega; \mathbf{A}) \right|^2 \right\rangle_{\mathbf{A}}}$$

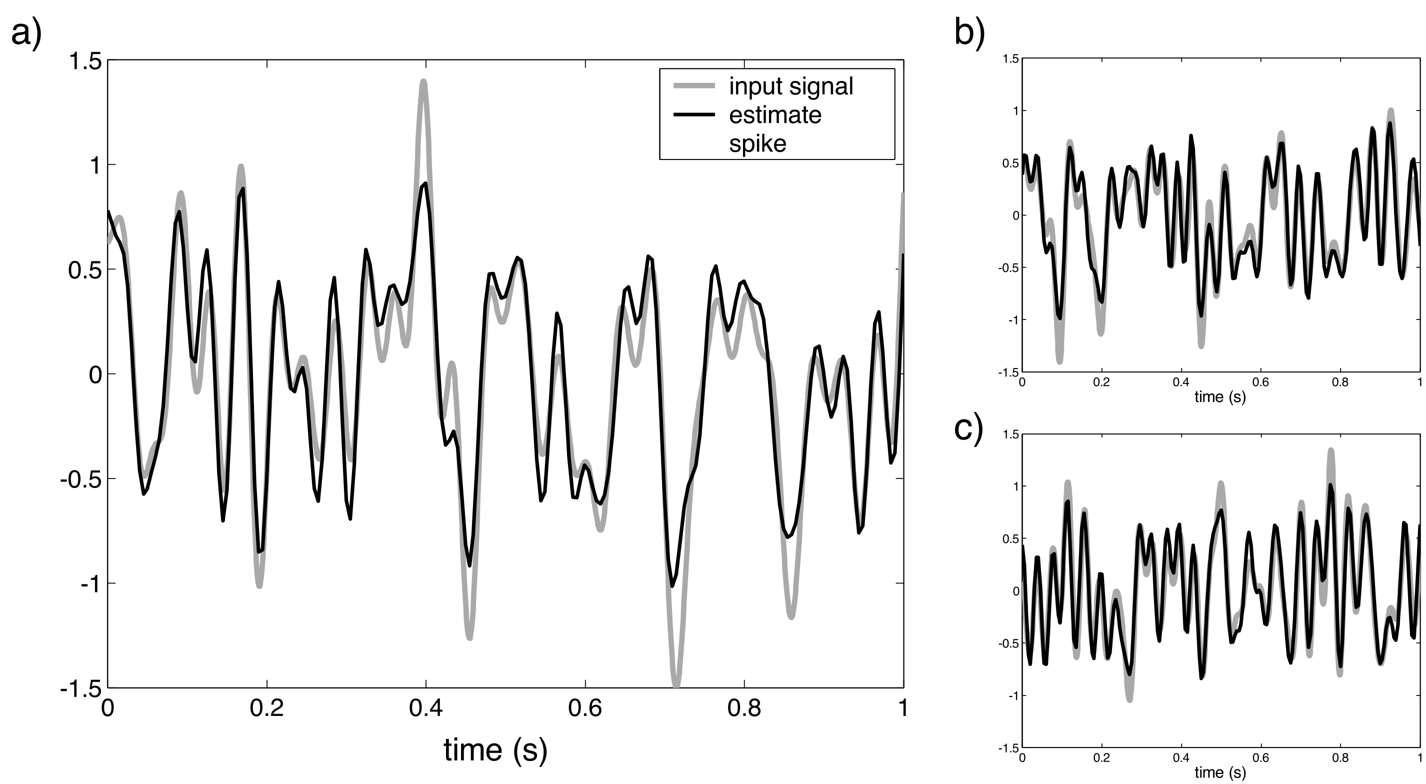
A sample filter



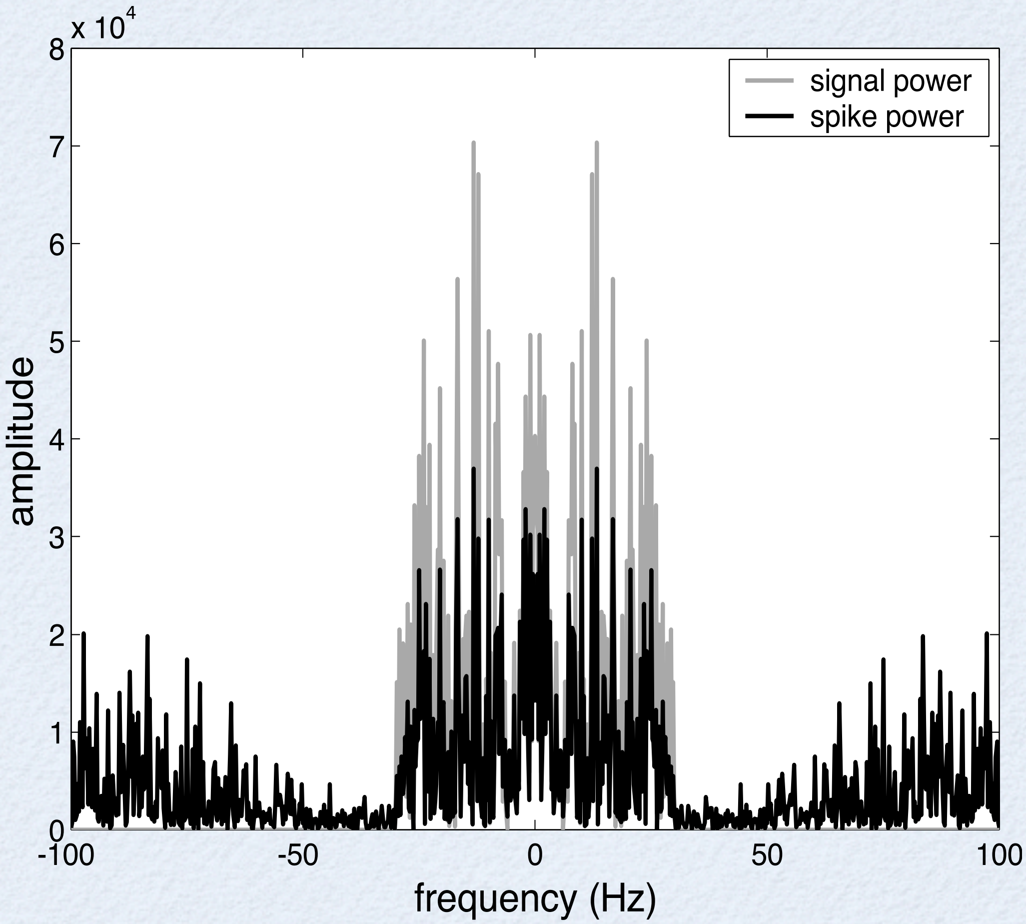
About this filter

- The filter was found for bandlimited (0-30Hz) Gaussian white noise on 4 s of data.
- The filter is not symmetrical in time but skewed in the positive direction
- The filter has a negative time component (non-causal).

Decoding examples

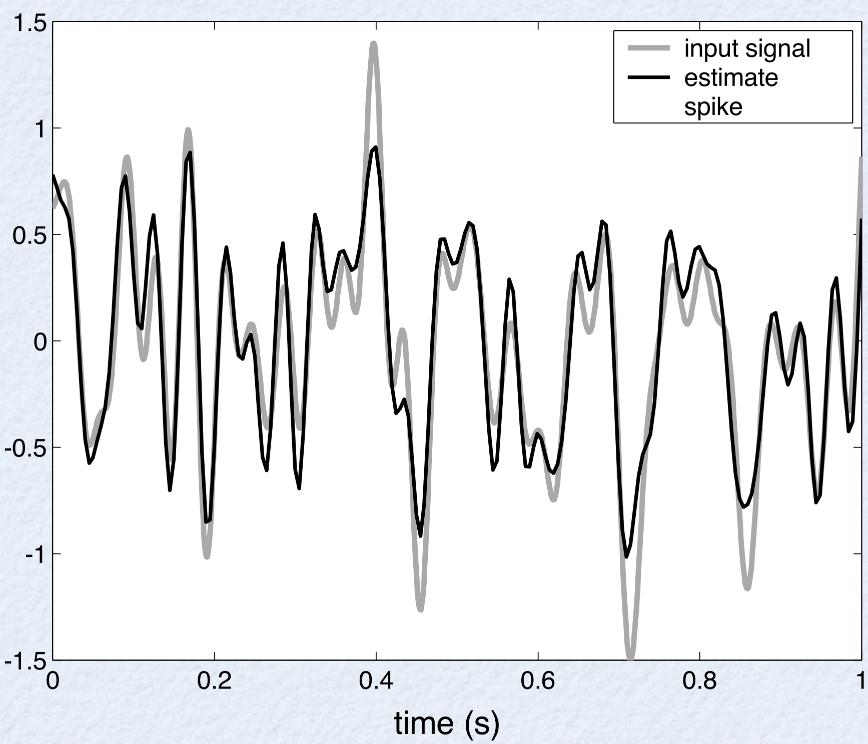


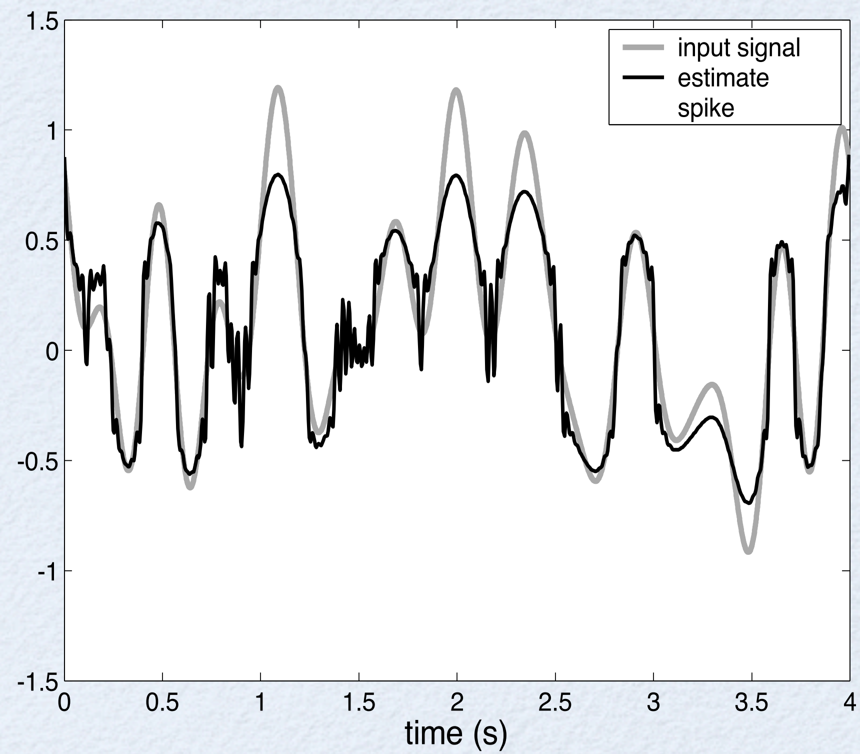
Power Spectrum Comparison



Rate/timing code

- Same h(t), different signals (note x-axis)





Correlation times

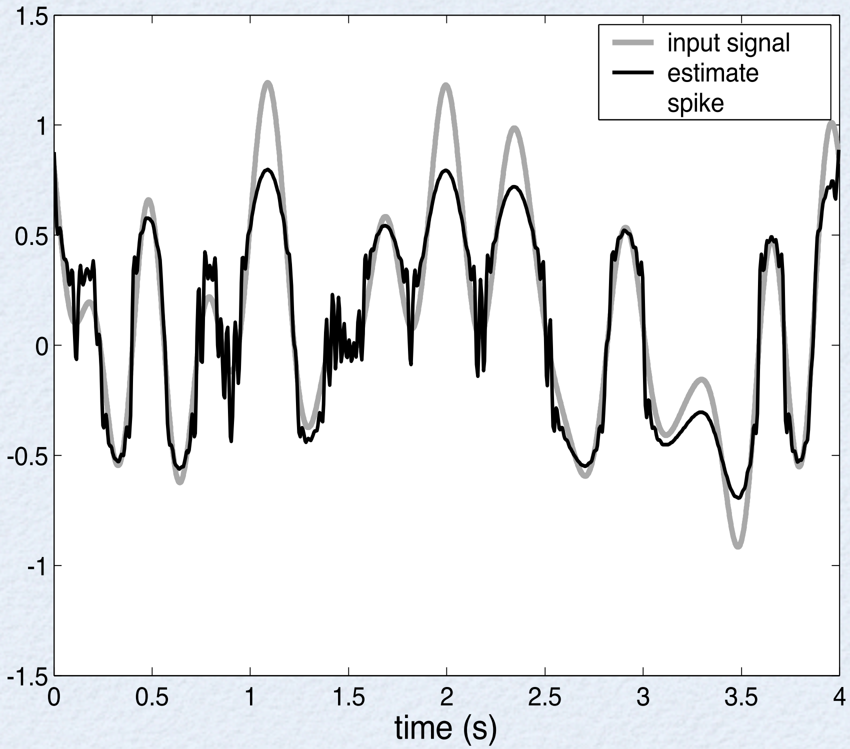
- One way to think about the difference between these functions is in terms of correlation time
- Correlation time: the maximum time window for which the autocorrelation function is above half its maximum value.

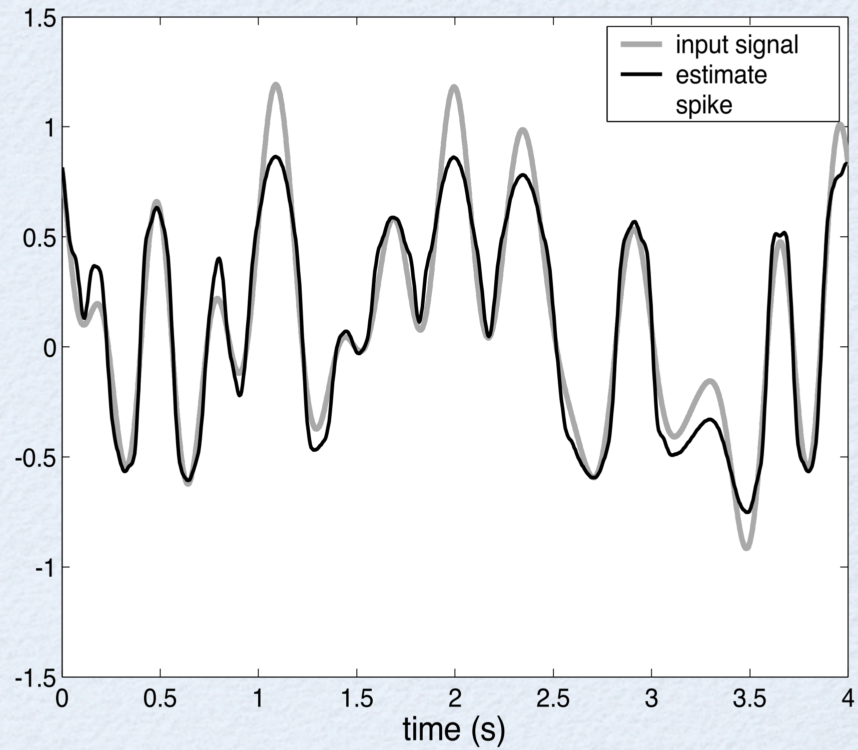
$$A(\tau) = \int_{-\infty}^{\infty} f(t) f(t+\tau)\, dt$$

- Future values of the signal are predictable out to that window
- So generally, a wider bandwidth means a shorter correlation time

Correlation time

- Different correlations times for finding h(t) make some, but not a huge difference

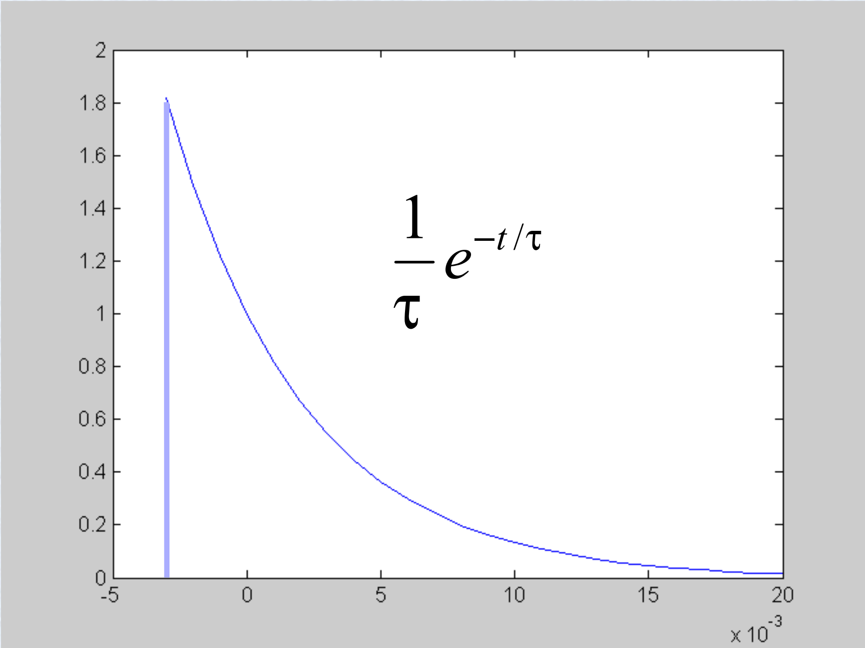


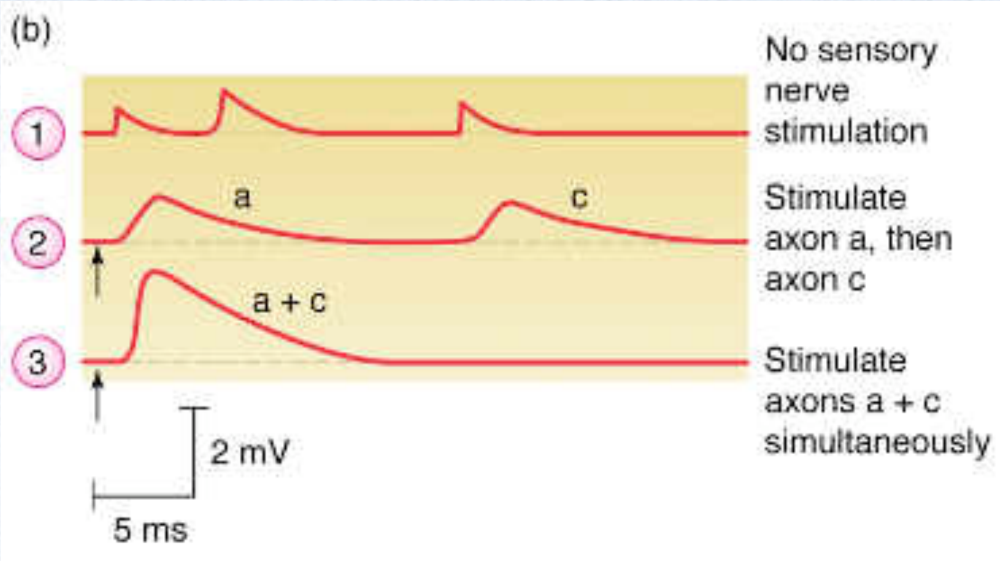


On optimal filters

- Optimal decoder extends just over 5 ms or so.
- They are non-causal... hence not used by brains
- Nevertheless, they provide:
  - a means of comparing models to real neurons
  - bounds for non-optimal decoding
  - justification of linear filtering

Typical PSCs



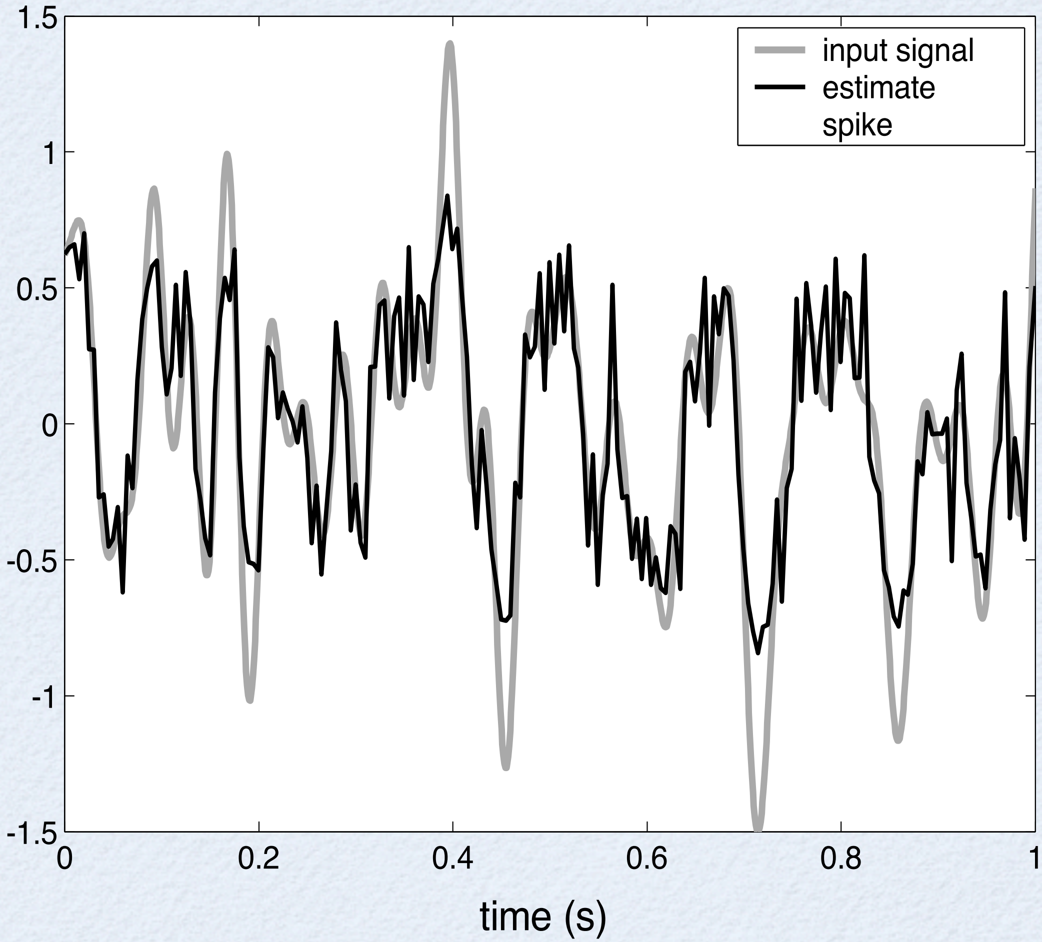


From: http://www.zoo.utoronto.ca/berry/zoo252/l9/lect9.htm

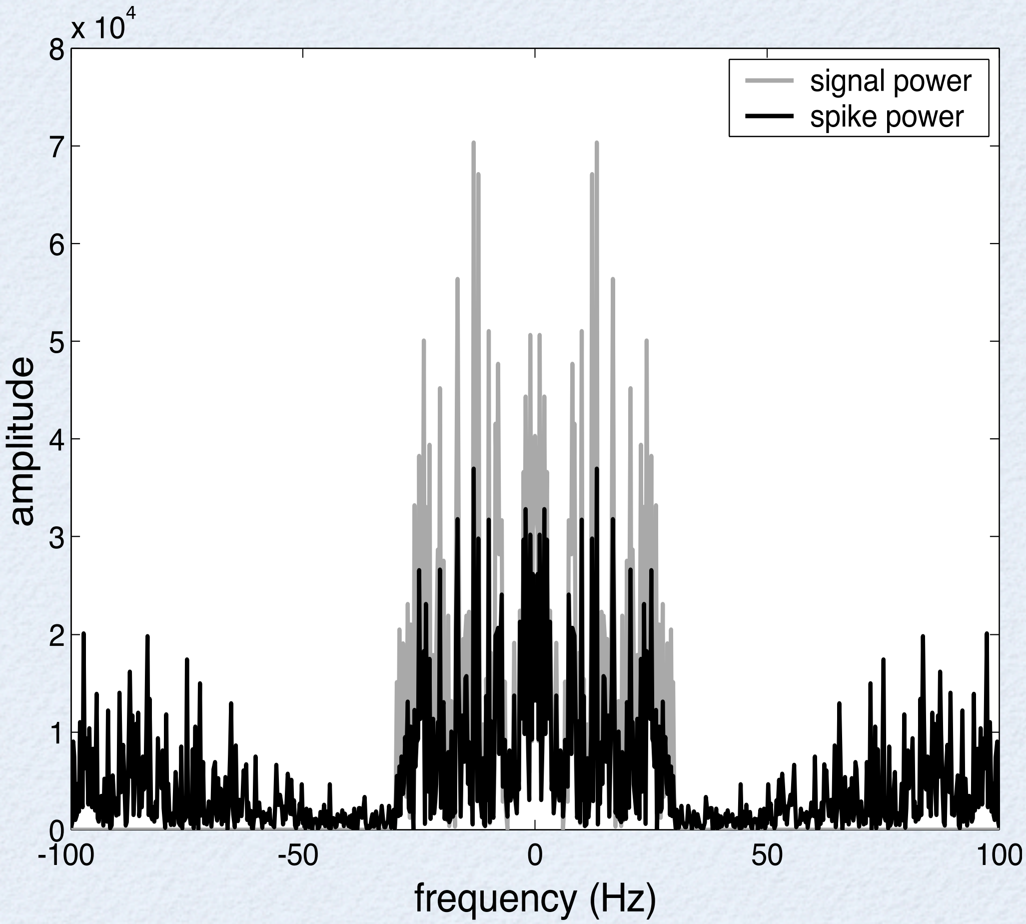
Post-synpatic currents

- PSCs are ellicited upon the reception of neurotransmitters by dendrites.
- PSCs look surprisingly like optimal filters (although they are strictly causal).
- They can be used as temporal filters...

Decoding with PSCs



Power spectrum comparison



PSCs vs Optimal filters

- Results in small reduction in information transmission, but a huge gain in plausibility
- Reductions in coding accuracy can be made up for by including more neurons.
- Similarity suggests linear decoding is a good characterization of neural representation
- So, all future models use PSC filtering instead of optimal filtering

Information in neural systems

- measures are surprisingly consistent across many different neural systems
  - cricket cercal system, between about 150 and 300 b/s. (between 1.1 and 3 bits per spike)
  - bullfrog sacculus rates of about 3 b/s
  - motion-selective H1 neurons; about 3 b/s
  - salamander retina, rate of about 3.4 b/s
  - primate visual area V5 1-2 b/s
- highest transmission rates: bullfrog auditory neurons, rates as high as 7.8 b/s (natural stimuli, broadband, 1.4 b/s)

Lessons

- these are impressively high transmission rates, approach the optimal possible coding efficiencies
- in frog sacculus, cricket cerci, bullfrog audition, and electric fish, the codes are 20-60% efficient
- efficiency increases for natural stimuli
- methods provide alower bound on the code, efficiencies could be higher

Information theory

- Self-information is defined as

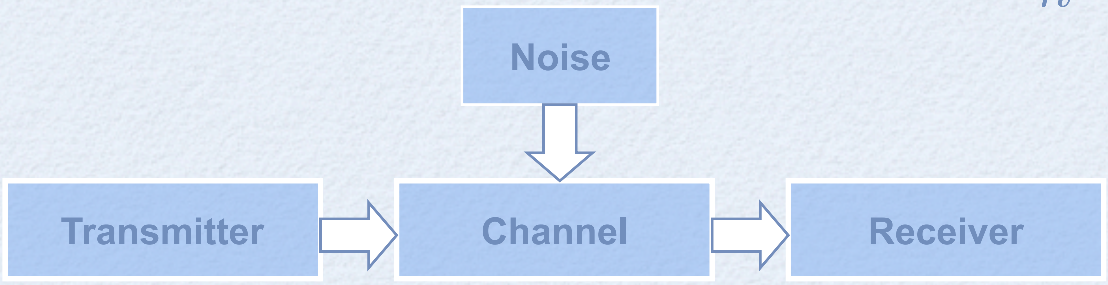
$$I(x) = log_2(1/p(x)) = -log_2(p(x))$$

- So, for independent events,

$$I(x_1 x_2) = I(x_1) + I(x_2)$$

- Entropy

$$H(x) = E(I(x)) = -\sum_{n} p(x_i) log_2 p(x_i)$$



Information in model neurons

- Technically, since the models are deterministic, there is no source of uncertainty.
- However, we have a source of error (linear decoding for nonlinear encoding).
- Can calculate information by looking at sources of error and using the Hartley-Shannon theorem:

$$I = \frac{1}{2}\log_2(1 + SNR)$$

Information and SNR

- Signal is explained power, noise is unexplained power (or variance), so

$$SNR = \frac{\left\langle \hat{x}^2 \right\rangle_x}{MSE}$$

- Where

$$\begin{aligned} MSE &= \left\langle x^2 \right\rangle_x - \left\langle \hat{x}^2 \right\rangle_x \\ \left\langle \hat{x}^2 \right\rangle_x &= \left\langle h^2 R^2 \right\rangle_x \end{aligned}$$

Information rate

- Information wrt input signal x per frequency channel:

$$I = \frac{1}{2}\log_2\left[\frac{\left\langle x^2\right\rangle_x}{\left\langle x^2\right\rangle_x - \frac{\langle x\cdot R(x)\rangle_x^2}{\langle R^2(x)\rangle_x}}\right]$$

- So we can compute the rate as

$$InfoRate = \frac{1}{2}\frac{\Delta\omega}{2\pi}\sum_n I_n$$

Optimal & PSC decoding