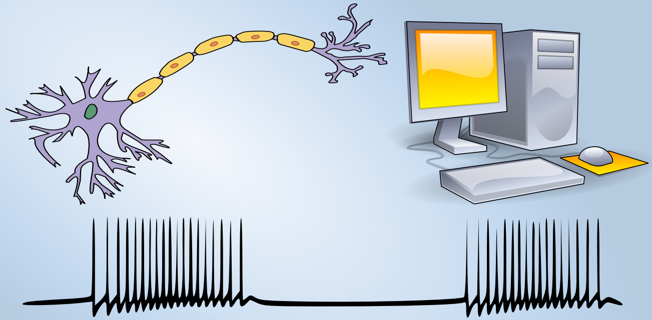
On the metabolically efficient neuronal information transmission

Lubomir Kostal
Institute of Physiology CAS, Prague, Czech Republic

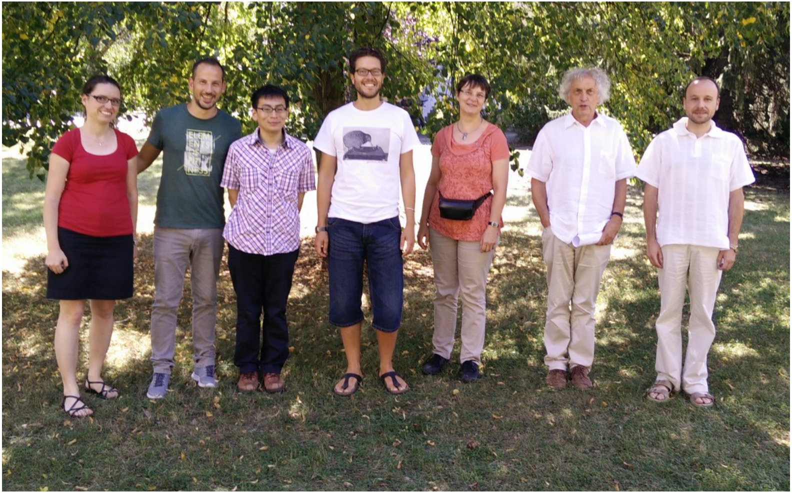
Computational Neuroscience Group, IPHYS, Prague



$$dX_t = \left(-\frac{X_t}{\theta} + \mu\right) dt + \sigma dW_t \qquad H(X) = -\sum_{i=1}^{n} p_i \ln p_i$$

http://comput.biomed.cas.cz

Computational Neuroscience Group, IPHYS, Prague



http://comput.biomed.cas.cz

Outline

1. What is *Computational neuroscience*?
2. Neurons are not *perfect*
3. Neurons are not *reliable*
4. Neurons process information *optimally* . . .
   4.1 Energy-efficient coding

- Thanks to: Ryota Kobayashi, Petr Lansky

Computational Neuroscience

"The aim of computational neuroscience is to explain how electrical and chemical signals are used in the brain to represent and process information."

T. Sejnowski *et al.*: Computational Neuroscience, *Science*, 1988

Cybernetics: Or Control and Communication in the Animal and the Machine

Norbert Wiener, 1948

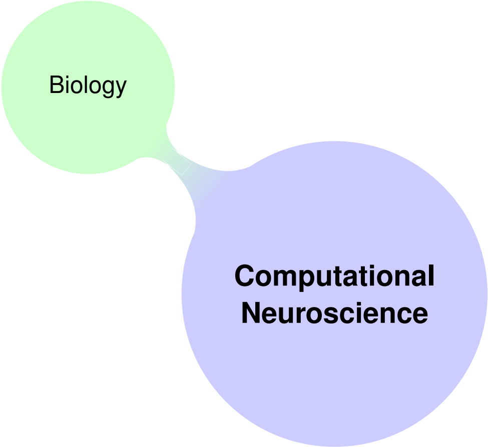
Why? – Progress in *neuroscience* (from molecules to fMRI)
– Progress in *computing power*

But... How the nervous system enables us to *see, remember, plan?*

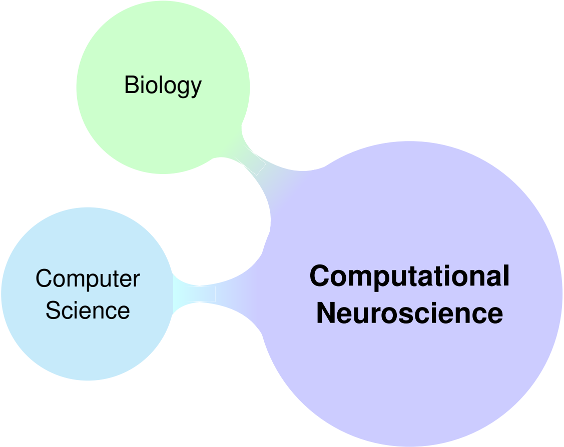
"What are the algorithms used in the brain?"

Computational Neuroscience

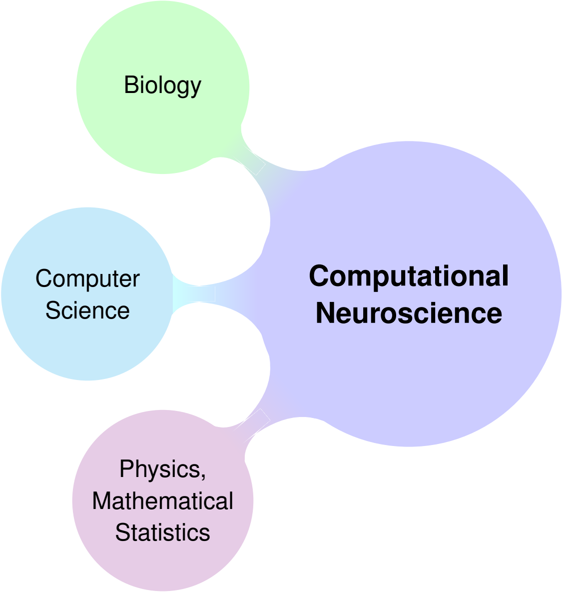
Multidisciplinarity (by definition) ...



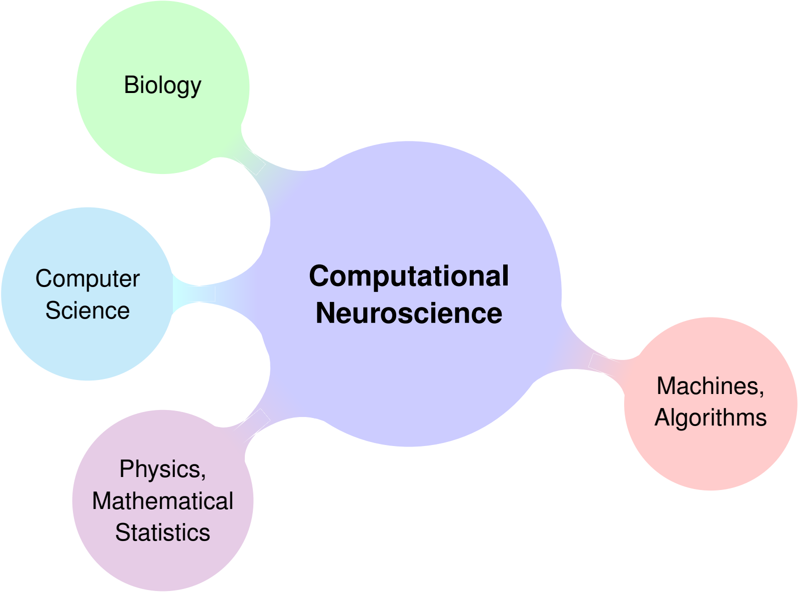
Multidisciplinarity (by definition) . . .



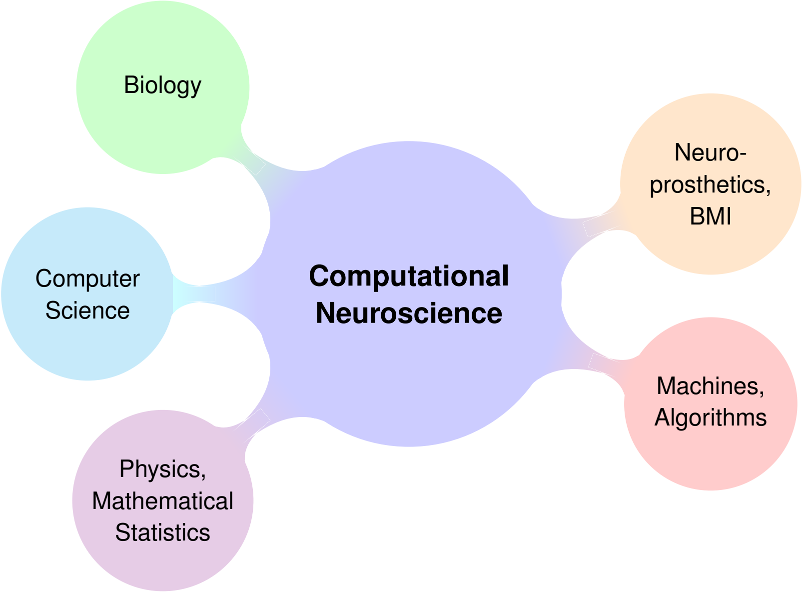
Multidisciplinarity (by definition) . . .



. . . *Fundamental Research* + Applications



. . . *Fundamental Research* + Applications



Computational neuroscience

- Effective theories (quantitative description)
- Since 1980's: dramatic increase (journals, conferences, labs, ...)

1. Models of neurons (networks, systems)
2. Coding, information processing and transmission

- Sensory neurons: stimulus coding $\Rightarrow$ artificial systems
- Applications: technology (HW, algorithms), *bio-inspired computing*

- Comput. Neuroscience *vs.* Artificial Neuronal Networks

Selected topics (Dept. Comput. Neurosci, Prague)

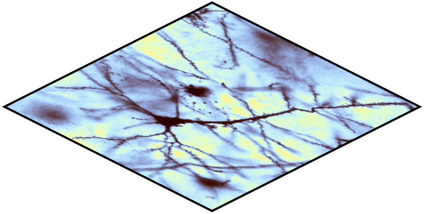
- **Focus: *neural coding***
- Neuronal spiking regularity measures and methods of their non-parametric estimation
- Models of insect pheromone receptor neurons (specific moth: *A. polyphemus*)
- Applications of statistical estimation theory
  - CRB achievability, “practical” sample sizes and *threshold effect* in ML estimation
- Applications of information theory
  - Efficient coding hypothesis in olfaction
  - Optimal processing of weak signals
  - Effect of metabolic cost on information transfer

Selected topics (Dept. Comput. Neurosci, Prague)

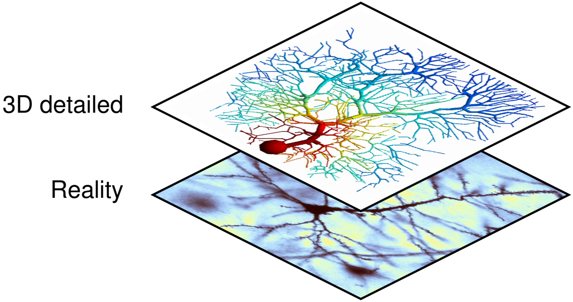
- **Focus: *neural coding***
- Neuronal spiking regularity measures and methods of their non-parametric estimation
- Models of insect pheromone receptor neurons (specific moth: *A. polyphemus*)
- Applications of statistical estimation theory
  - CRB achievability, “practical” sample sizes and *threshold effect* in ML estimation
- Applications of information theory
  - Efficient coding hypothesis in olfaction
  - Optimal processing of weak signals
  - Effect of metabolic cost on information transfer

Neuronal models

Reality



Neuronal models

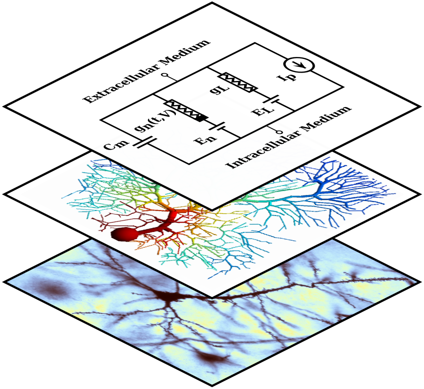


Neuronal models

Hodgkin-Huxley

3D detailed

Reality



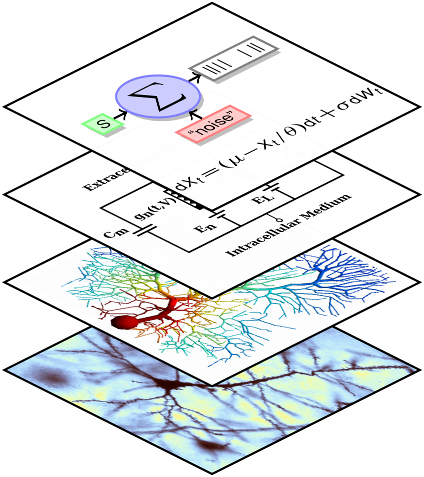
Neuronal models

Integrate & Fire

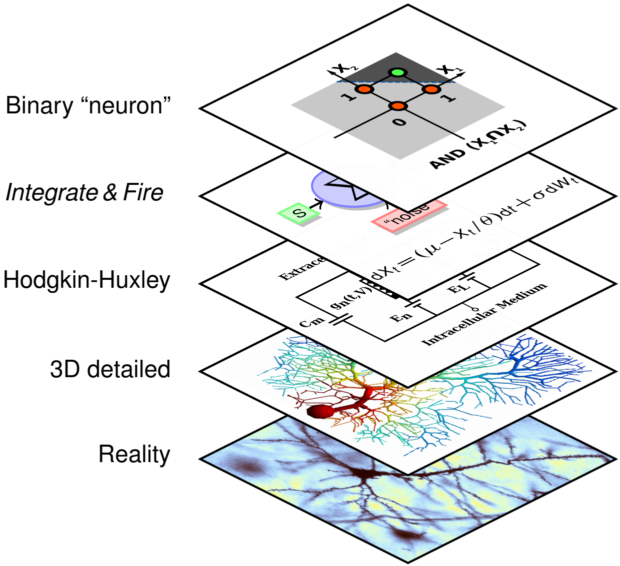
Hodgkin-Huxley

3D detailed

Reality



Neuronal models



Neuronal models

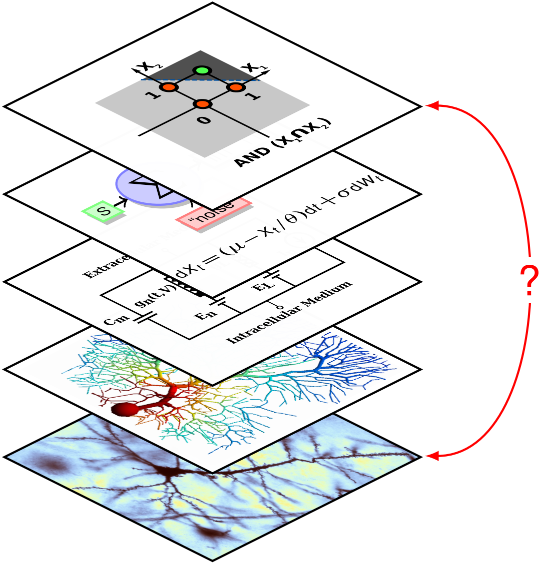
Binary “neuron”

Integrate & Fire

Hodgkin-Huxley

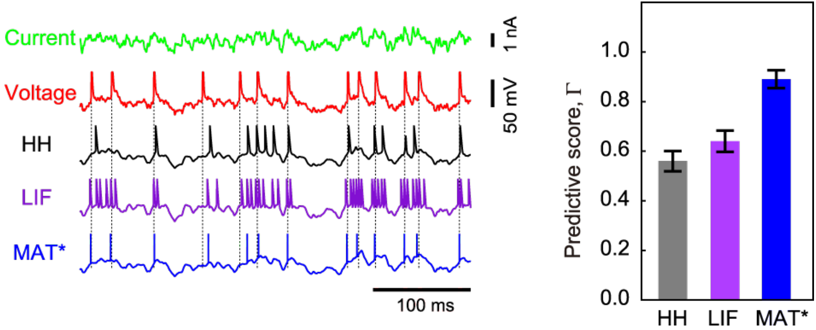
3D detailed

Reality



Spike prediction: how "good" are neuronal models?

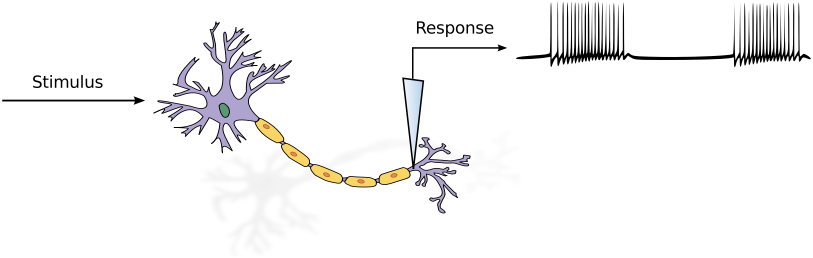
Ryota Kobayashi: *Quantitative Neuron Modeling* (2007, 2008, 2009)



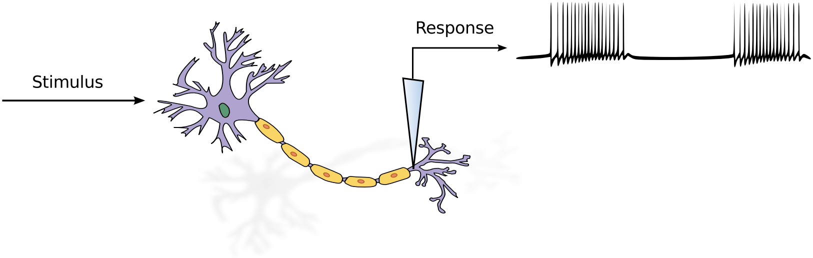
(Rat motor cortex)

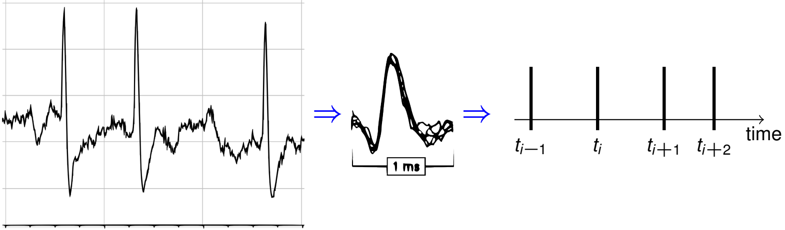
Kobayashi et al., *Front. Comput. Neurosci.* (2009)
Jolivet et al., *J. Neurosci. Methods* (2008); Gerstner & Naud, *Science* (2009)

Neuronal code: basic assumptions

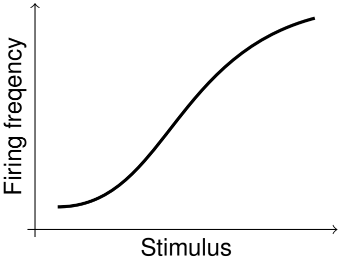


Neuronal code: basic assumptions



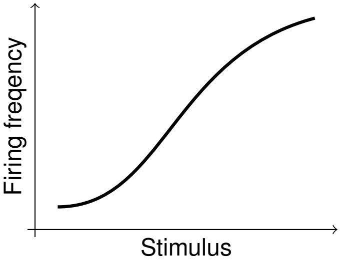


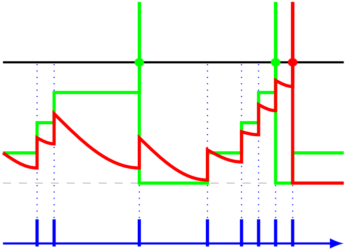
Frequency *vs.* temporal coding



1. Frequency: Adrian (1926), number of pulses (AP) per unit time

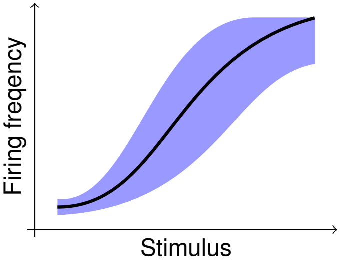
Frequency *vs.* temporal coding

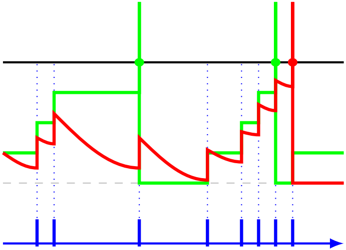




1. Frequency: Adrian (1926), number of pulses (AP) per unit time
2. Temporal: Perkel & Bullock (1968), intervals between APs,

Frequency *vs.* temporal coding



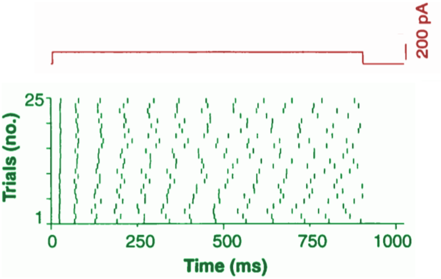


1. Frequency: Adrian (1926), number of pulses (AP) per unit time
2. Temporal: Perkel & Bullock (1968), intervals between APs,

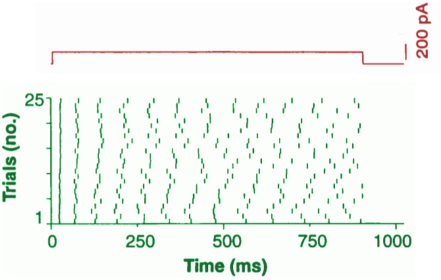
- *Variability, noise*: Stein et al., *Nat. Rev. Neurosci.* 2005

Reliability of neuronal response

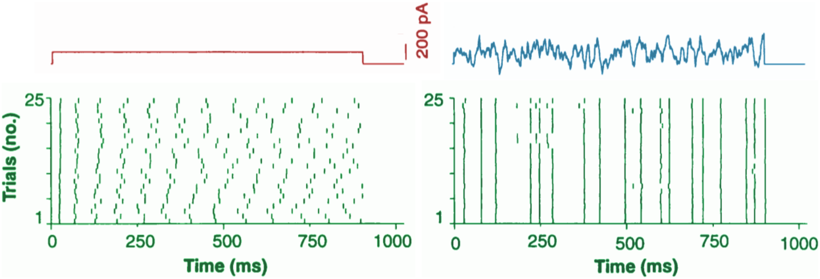
Reliability of neuronal response



Reliability of neuronal response

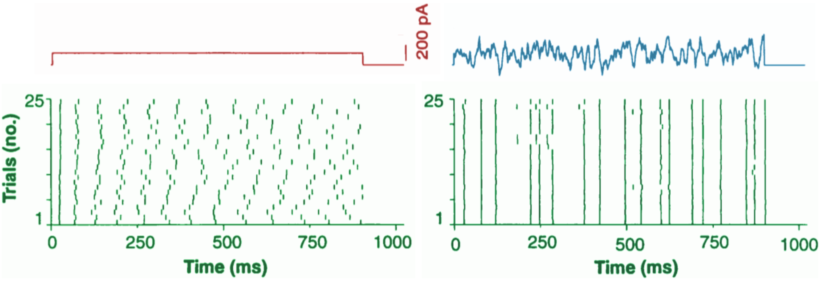


Reliability of neuronal response



Mainen and Sejnowski, *Science* (1995)

Reliability of neuronal response



Mainen and Sejnowski, *Science* (1995)

1. Encoding of *elementary* (simple) *vs. complex* stimuli
   - rate-level, dose-response, tuning *curves*
   - bottom-up approach?
2. “*Convenient*” *vs. natural* stimulation
   - parameterization (dimensionality, Gaussian processes)
   - experimental setup (*e.g.*, insect olfaction)

Efficient coding hypothesis

Horace B. Barlow, 1961

Neurons are adapted, through both evolutionary and developmental processes, to the *statistical characteristics* of their *natural* stimulus.

Efficient coding hypothesis

Horace B. Barlow, 1961

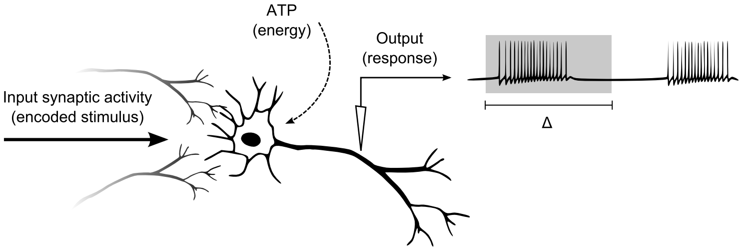
Neurons are adapted, through both evolutionary and developmental processes, to the *statistical characteristics* of their *natural* stimulus.

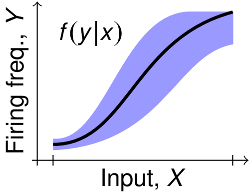
Methods: *information theory*, estimation theory, ...

- Optimality conditions, *infomax* (Linsker 1987)
- Retinal neurons (Laughlin 1981, Laughlin *et al.* 1996)
- “Scale” of neuronal performance (Rieke *et al.*, 1996)

Extensions: metabolic cost, decoding feasibility, realistic models, ...

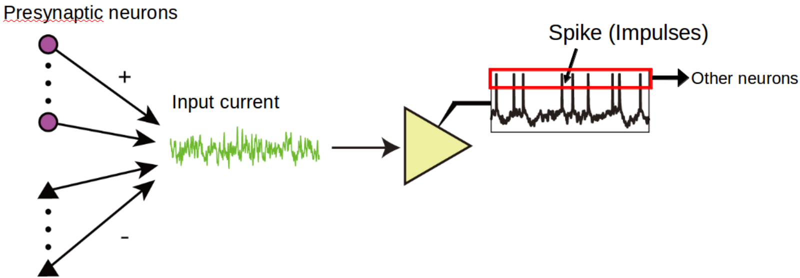
Energy-efficient neural coding?





- Input (exc. conductance, X) – Output (response, Y)
  - Rate coding: #APs in Δ
  - Temporal coding, ...
  - Model parameters, type of stimulation, ...
- Model: $f(y|x)$✔ but $p(x)$**?**
- Efficiency: **Energy** $\times$ **Information** $\times$ **...**

Investigated neuronal model (*cortical excitatory*)



- Extended Hodgkin-Huxley + *point-conductance* (*stochasticity*)
  - Adaptation (I_M)
  - Balanced input, $\lambda_E \propto \lambda_I$
  - Excitatory and inhibitory conductances, $\langle g_{E,I} \rangle \propto \lambda_{E,I}$
  - Effective reversal potential: V_r (Miura *et al.*, 2007)
- Input, $x \equiv \langle g_E \rangle$: mean excitatory conductance (input parameter)
- Output, $y \equiv \#\text{APs}/\Delta$: firing rate

Metabolic cost of neuronal activity & efficiency

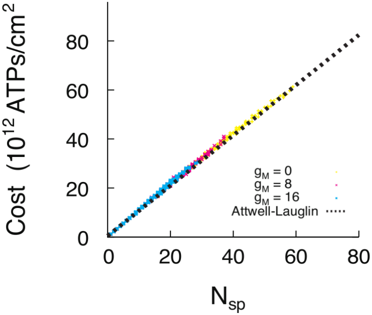
- *Empirical* metabolic cost given $X = x$ (Attwell & Laughlin, 2001)

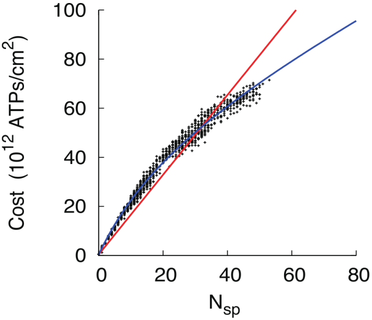
$$w(x) = \kappa \times \big(\langle \#\text{APs in } \Delta\rangle \,|\, x\big) + \beta\,\Delta$$

$[\kappa = 7.1 \times 10^8$ ATPm, $\beta = 4.4 \times 10^8$ ATPm/s$]$

- *Theoretical* (model) cost: only small corrections (RK)✔

- Excitatory *vs.* inhibitory neurons





Methods: mutual information $I(X;Y)$

- Given the "model" $f(y|x)$ and input distribution $X \sim p(x)$:

$$I(X;Y) = \int_X \int_Y p(x) f(y|x) \log_2 \frac{f(y|x)}{\int_Z f(y|z) p(z)\,\mathrm{d}z} \mathrm{d}y\,\mathrm{d}x$$

- Interpretation of $I(X;Y)$ – why $I(X;Y)$?

Methods: mutual information $I(X;Y)$

- Given the "model" $f(y|x)$ and input distribution $X \sim p(x)$:

$$I(X;Y) = \int_X \int_Y p(x) f(y|x) \log_2 \frac{f(y|x)}{\int_Z f(y|z) p(z)\,\mathrm{d}z}\,\mathrm{d}y\,\mathrm{d}x$$

- Interpretation of $I(X;Y)$ – why $I(X;Y)$?

  1. General: Average statistical dependence (nonlinear) between X and Y
     take care: varying $p(x)$ to $\uparrow I(X;Y)$ may $\uparrow \mathrm{MSE}[\hat{x}(y)]$ too!
  2. "Information transfer" (Shannon) between X and Y: *reliable* transfer, amount in bits
     – conditions are not automatically guaranteed!

Methods continued: capacity-cost function $C(W)$

- *Efficient coding hypothesis*: find the ultimate bounds
- Vary the input distribution $p(x)$ to maximize $I(X;Y)$

$$C(W) = \max_{p(x):W_p \leq W} I(X;Y)$$

- "Interpretation" of $C(W)$: maximum information transfer, provided that the average metabolic cost does not exceed W
- Classical channel capacity: $C = C(W \to \infty)$

Methods continued: capacity-cost function $C(W)$

- *Efficient coding hypothesis*: find the ultimate bounds
- Vary the input distribution $p(x)$ to maximize $I(X;Y)$

$$C(W) = \max_{p(x):W_p \leq W} I(X;Y)$$

- "Interpretation" of $C(W)$: maximum information transfer, provided that the average metabolic cost does not exceed W
- Classical channel capacity: $C = C(W \to \infty)$

- Optimal "trade-off" between information and cost (if it exists), the efficiency:

$$E = \max_W \frac{C(W)}{W},$$

  i.e., $1/E$ is the *minimal cost* of 1 bit

Methods: information optimization

- $L[p]$: functional over convex and compact set $\mathcal{F}$ (Smith, 1971)

$$\delta_g L[p_0] = \lim_{\varepsilon \downarrow 0} \frac{L[(1-\varepsilon)p_0 + \varepsilon g] - L[p_0]}{\varepsilon}, \quad \varepsilon \in [0,1]$$

- $L[p]$ diff. and convex $\cap$: unique maximum exists at p_0:

$$\delta_g L[p_0] \leq 0 \quad \text{for all } g \in \mathcal{F}$$

Methods: information optimization

- $L[p]$: functional over convex and compact set $\mathcal{F}$ (Smith, 1971)

$$\delta_g L[p_0] = \lim_{\varepsilon \downarrow 0} \frac{L[(1-\varepsilon)p_0 + \varepsilon g] - L[p_0]}{\varepsilon}, \quad \varepsilon \in [0, 1]$$

- $L[p]$ diff. and convex $\cap$: unique maximum exists at p_0:

$$\delta_g L[p_0] \leq 0 \quad \text{for all } g \in \mathcal{F}$$

- Finding $C(W)$: convex $\cap$ in $p(x)$, e.g.,

$$\text{''}\, L[p] = I(X;Y) - \lambda \int p(x) w(x)\, \mathrm{d}x \text{''}$$

Optimality conditions

- $I[p] = I(X;Y)$ evaluated for $X \sim p(x)$, $N[p] = \int_X p(x)\,\mathrm{d}x - 1$, $\Gamma[p] = \int_X w(x)p(x)\,\mathrm{d}x - W$, $\lambda \geq 0$ and β are multipliers

$$L[p] = I[p] - \lambda\Gamma[p] - \beta N[p],$$

- Directional diff δ_g (after some manipulations)

$$\delta_g L[p] = \int_X g(x)\left[D_{\mathrm{KL}}[f(y|x) \,\|\, p(y)] - 1 - \beta - \lambda w(x)\right]\mathrm{d}x,$$

- Hence $\delta L[p] \leq 0$ implies (with '=' for $p(x) > 0$)

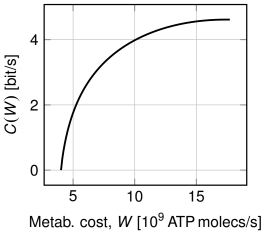
$$w(x) \geq \frac{1}{\lambda}\int f(y|x)\log_2 \frac{f(y|x)}{\int p(\tilde{x})f(y|\tilde{x})\,\mathrm{d}\tilde{x}}\,\mathrm{d}y - \frac{1-\beta}{\lambda},$$

Methods of solution and results

- Numerical methods only, but the model may be arbitrary
- Algorithm: Cutting-plane + LP
  (Huang & Meyn, 2005), Eulerian $\times$ Lagrangian CM
- Results:
  - obtain $C(W)$ and the "optimal" input distribution $p^*(x)$ for each W
  - efficiency E and the corresp. *optimal cost* W^*
  - postsynaptic firing rate dist.: $f^*(y) = \int f(y|x)p^*(x)\,\mathrm{d}x$
- Sensitivity to perturbations of $p^*(x)$
- $p^*(x)$ is *discrete* (typical for finite-range inputs)

Results: max. information, efficiency, PS firing rates

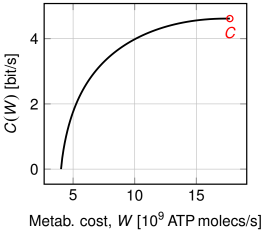
(a) Capacity-cost, $C(W)$



($V_r = -60\,\mathrm{mV}$, $g_M = 8\,\mathrm{mS/cm^2}$)

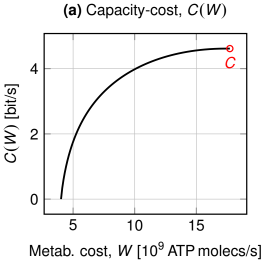
Results: max. information, efficiency, PS firing rates

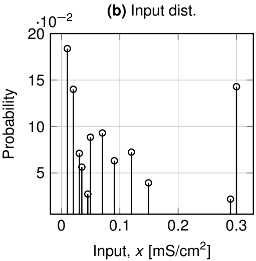
(a) Capacity-cost, $C(W)$



($V_r = -60\,\mathrm{mV}$, $g_M = 8\,\mathrm{mS/cm^2}$)

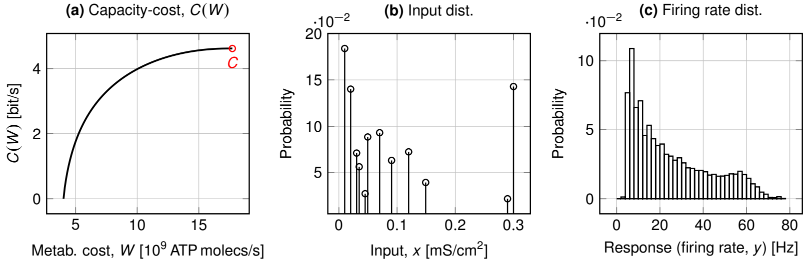
Results: max. information, efficiency, PS firing rates





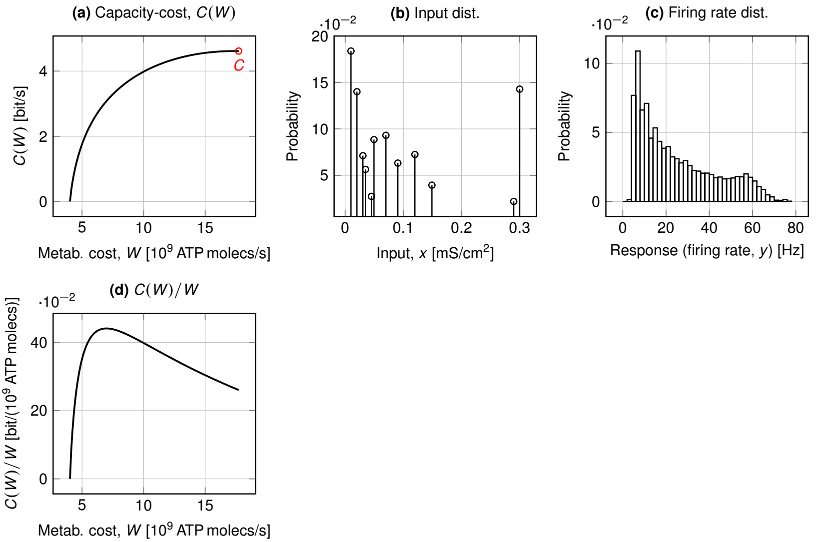
($V_r = -60$ mV, $g_M = 8$ mS/cm^2)

Results: max. information, efficiency, PS firing rates



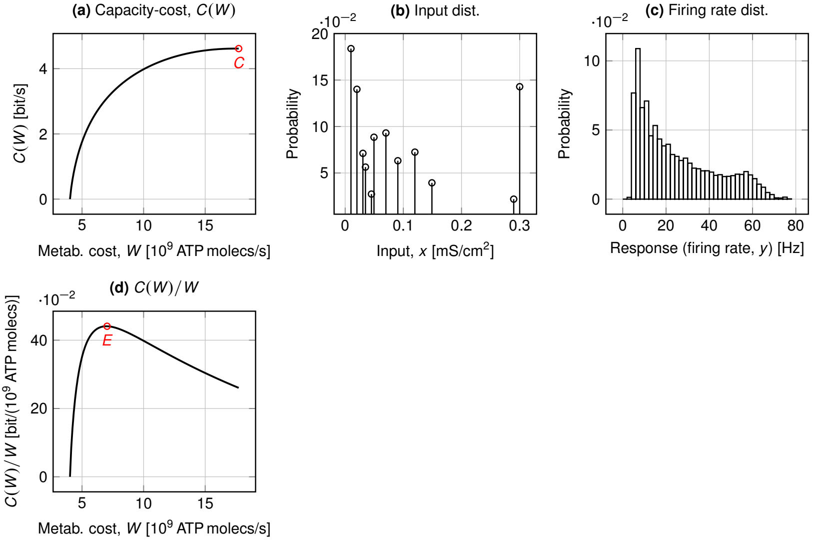
($V_r = -60$ mV, $g_M = 8$ mS/cm^2)

Results: max. information, efficiency, PS firing rates



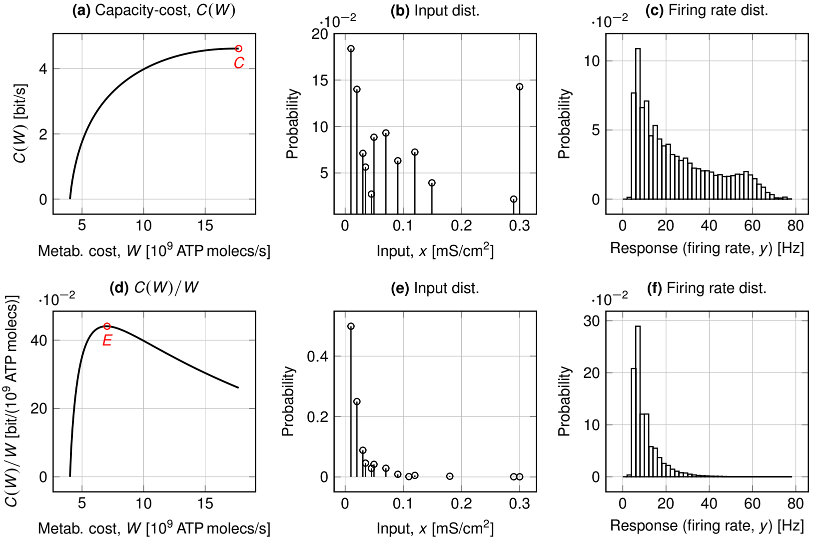
($V_r = -60\,\text{mV}$, $g_M = 8\,\text{mS/cm}^2$)

Results: max. information, efficiency, PS firing rates



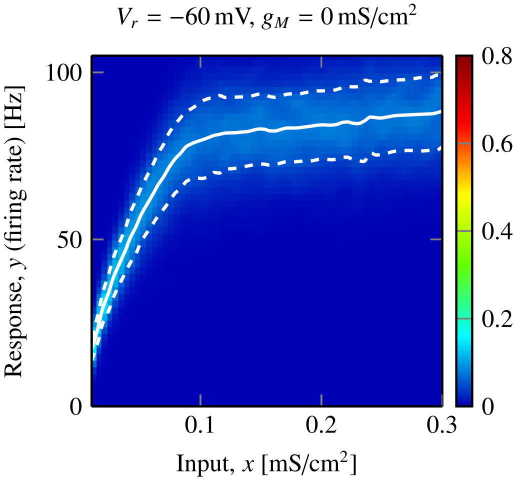
($V_r = -60\,\text{mV}$, $g_M = 8\,\text{mS/cm}^2$)

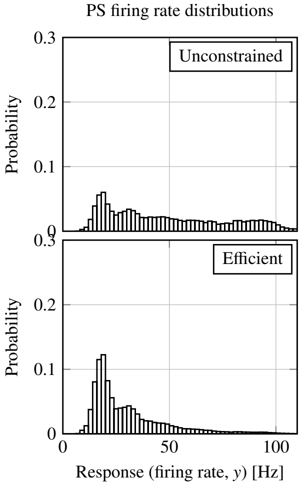
Results: max. information, efficiency, PS firing rates



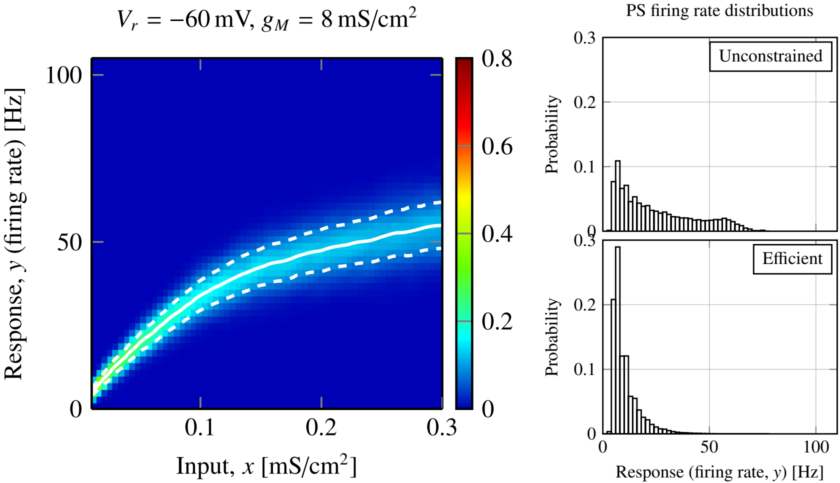
($V_r = -60\,\text{mV}$, $g_M = 8\,\text{mS/cm}^2$)

Neuronal model: increasing adaptation, g_M

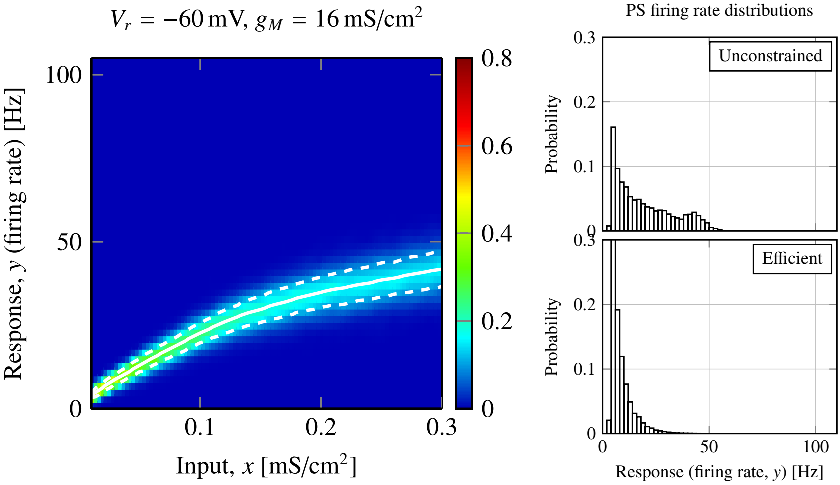




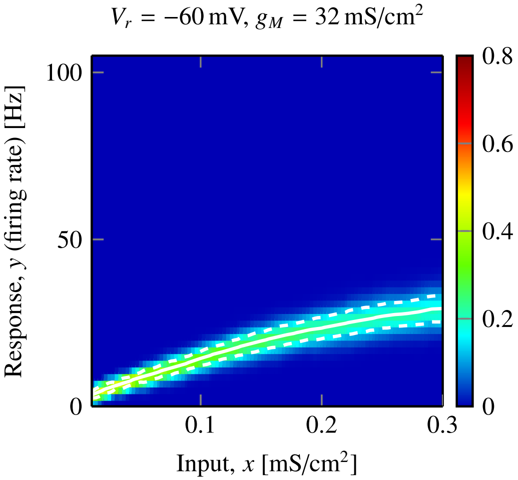
Neuronal model: increasing adaptation, g_M

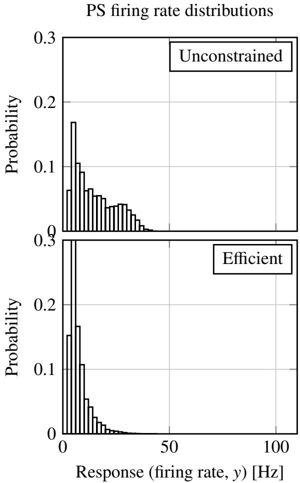


Neuronal model: increasing adaptation, g_M

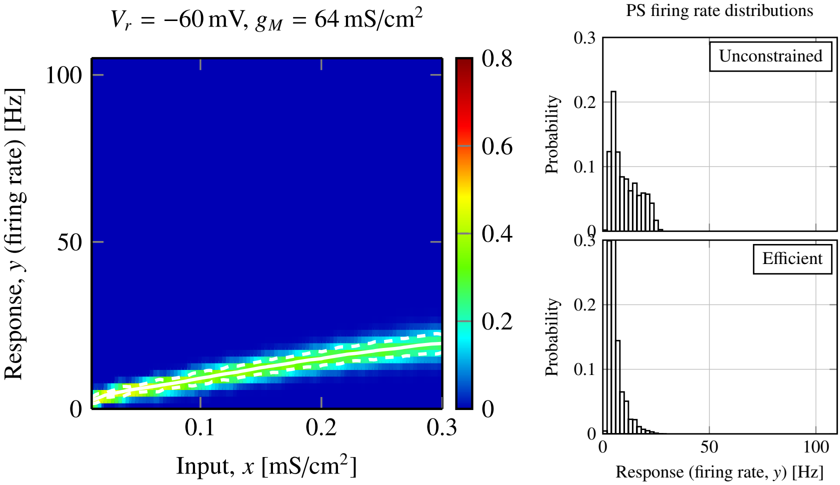


Neuronal model: increasing adaptation, g_M



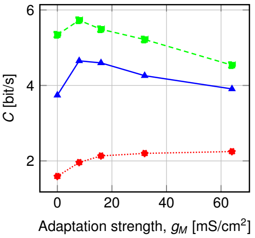


Neuronal model: increasing adaptation, g_M

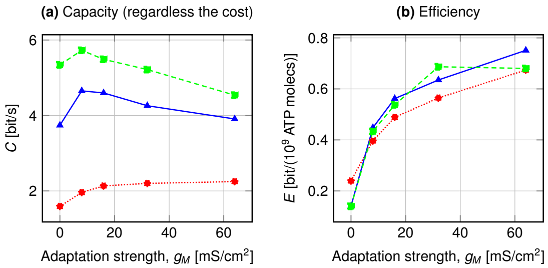


Results: adaptation *vs.* information efficiency

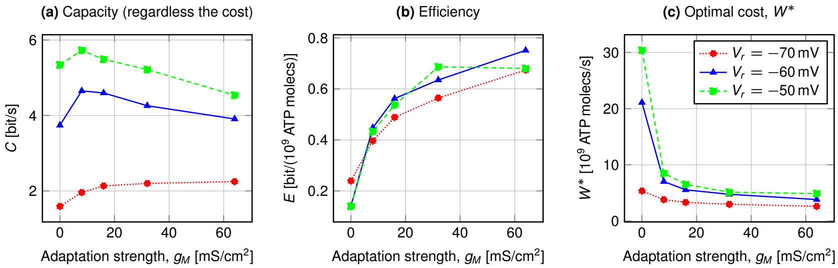
(a) Capacity (regardless the cost)



Results: adaptation *vs.* information efficiency



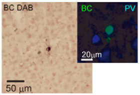
Results: adaptation *vs.* information efficiency

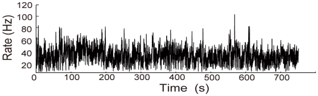


- Preliminary conclusions:
  - Adaptation facilitates the efficiency
  - Adaptation reduces the optimal metabolic workload

Comparison with experimental data

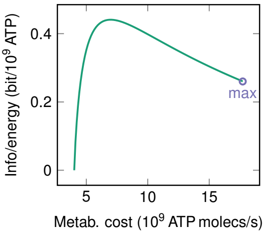
Experimental data:
in vivo (recordings $> 30\,$min.),
layer 2-6 (sensorimotor cortex),
pyramidal and inter-neurons



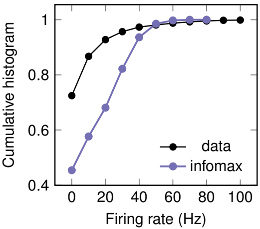


Data: Dr. Tomoki Fukai (RIKEN)

Information/energy

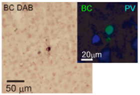


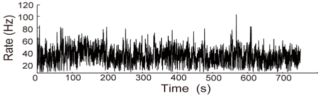
Pyramidal neuron (example)



Comparison with experimental data

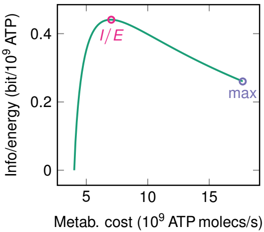
Experimental data:
in vivo (recordings $> 30\,\text{min.}$),
layer 2-6 (sensorimotor cortex),
pyramidal and inter-neurons



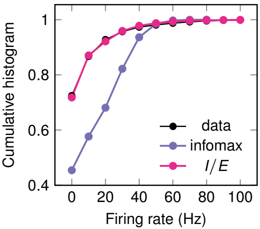


Data: Dr. Tomoki Fukai (RIKEN)

Information/energy

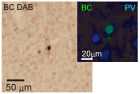


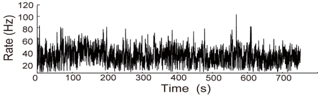
Pyramidal neuron (example)



Comparison with experimental data

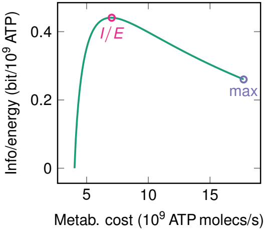
Experimental data:
in vivo (recordings > 30 min.),
layer 2-6 (sensorimotor cortex),
pyramidal and inter-neurons



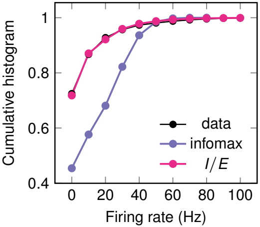


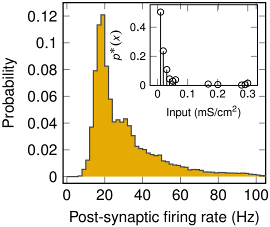
Data: Dr. Tomoki Fukai (RIKEN)

Information/energy



Pyramidal neuron (example)





– Predicted PSFR histograms match the data

– Sparse synaptic activity (Brecht *et al.*, *Nature*, 2004)

– Log-normal distribution of synaptic weights (Song *et al.*, *PLoS Biol.*, 2005)

– Low muscarinic-K conductance ($\leq$ 8 mS/cm^2)

Conclusions

- Metabolic cost considerations seem to have significant impact on the results
- Spike-frequency adaptation improves metabolic efficiency of information transmission,
  e.g., *cost per single bit decreases with increasing* g_M
- For discussion:
  - Interpretation of information-theoretic quantities
  - *Achievability* of information transfer
  - Separated *vs.* joint source-channel coding