

Adaptive contrast gain control and information maximization ¹

Yuguo Yu* and Tai Sing Lee*,[†]

Center for the Neural Basis of Cognition & Computer Science Department*,[†]
Carnegie Mellon University, Pittsburgh PA 15213, U.S.A.*

Abstract

Contrast gain control is an ubiquitous phenomenon in biological systems. While the static nonlinearities (e.g., saturation and threshold) of neurons can explain a number of contrast gain control phenomena [1-2], the theoretical connection between information maximization and contrast gain control is not completely clear. In this paper, we report that for a neuron with fixed threshold and saturation, information maximization dictates that the gain of the neuronal transfer function and the variance of the input signals should obey an inverse power law. This suggests that in addition to the effect due to static nonlinearity, an additional rescaling of the amplitude gain of the transfer function is needed to achieve information maximization.

Key words: Adaptation, Contrast gain control, Information maximization

1 Introduction

Contrast gain control has been widely observed in many natural systems. It has been observed that neurons of the biological visual systems decrease their gain of the transfer functions when the contrast or variance of the visual input is increased [3-6]. This adaptation in the kernel gain has been suggested to serve to maximize information transmission [7-8]. We have shown that a static neuronal model with fixed threshold and saturation level can already exhibit an adaptation in gain as a function of input variance in a manner that is qualitatively consistent with many observed adaptation phenomena [1]. However, we found that the mutual information, as well as the kernel gain, exhibit a tuning curve that varies as a function of the input variance [2]. That means

¹ This research was supported by NSF CAREER 9984706, NIH Vision Research core grant EY08098, and a NIH grant on the Pittsburgh Supercomputing Center
Corresponding author: Yuguo Yu, yug@cnbc.cmu.edu.

nonlinear dynamics due to static nonlinearity alone cannot maintain a neuron's information transmission rate at an optimal level with the variation in signal variance. Here, we explored theoretically a possible extra adaptation mechanism that might restore information maximization in neurons. We introduce an additional gain rescaling to the static model, while keeping all other parameters (thresholds, saturation levels, membrane time constants) constant. We found it sufficient to produce information maximization that is consistent with some neurophysiological findings [6-7]

2 Model and Results

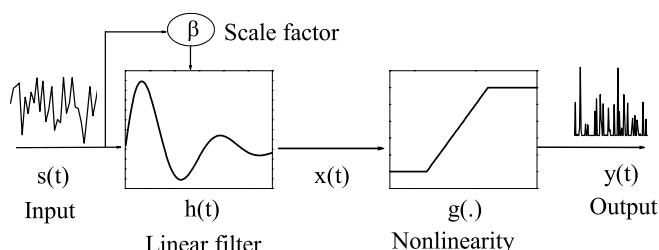


Fig. 1. The LN model consists of a linear filter $h(t)$ followed by a nonlinearity $g(\cdot)$. In the adaptive LN model, the amplitude of the linear filter is scaled by β , which acts as an adaptive mechanism. $x(t)$ is the convolution of the input signal $s(t)$ and the filter $h(t)$. The nonlinearity $g(\cdot)$ operates on $x(t)$ to generate the output $y(t)$.

We use the similar standard linear-nonlinear (LN) cascade model as in Ref.[2] except that we introduce an additional scaling factor $\beta(\sigma)$ see Fig.1. β scales the amplitude of the linear kernel according to the variance of the input signal as an adaptive mechanism. Linear kernel is given by $h(t) = \beta(\sigma) * A * \sin(\pi t/\tau_a) \exp(-t/\tau_b)$ with $A = 1$, $\tau_a = 80$ ms and $\tau_b = 100$ ms. The output of linear kernel $x(t)$ is given by $x(t) = \int_0^{+\infty} h(\tau) s(t - \tau) d\tau$. The nonlinearity $g(\cdot)$ is specified by

$$g(x) = \begin{cases} 0, & \text{if } x < \theta, \\ x - \theta, & \text{if } \theta \leq x < \eta, \\ \eta - \theta, & \text{if } x \geq \eta. \end{cases} \quad (1)$$

where θ is threshold and η is saturation level. Without loss of generality, we use a Gaussian white noise stimulus $s(t)$ with zero mean and SD σ as the input signal. Its probability density function (PDF) is given by $p(s) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{s^2}{2\sigma^2}}$. The linear response $x(t)$ also has a Gaussian distribution with PDF $p(x) =$

$\frac{1}{\sqrt{2\pi\sigma_x^2}}e^{-\frac{x^2}{2\sigma_x^2}}$, where σ_x is given by $\sigma_x^2 = \langle x^2(t) \rangle = \sigma^2 \int_0^{+\infty} h^2(\tau) d\tau$, where $\langle \dots \rangle$ denotes time average. β is determined computationally by maximizing the mutual information between $s(t)$ and $y(t)$ with respect to β , for signals of a given σ . This is the adaptive LN model.

We have shown the gain of the effective transfer function of the *static model* varies with the input signal variance σ^2 [1-2]. The ratio α between the gain of the effective kernel and the original linear kernel is given by

$$\alpha = \frac{\int_{\theta}^{\eta} x(x - \theta)p(x)dx + (\eta - \theta) \int_{\eta}^{+\infty} xp(x)dx}{\sigma^2 \int_0^{+\infty} h^2(\tau)d\tau}. \quad (2)$$

α quantifies how much the effective gain of the linear kernel $h(t)$ is affected by threshold θ , saturation η and stimulus standard deviation σ . Mutual information [9] quantifies the ability of a system or a communication channel to convey information and is given by

$$I_m = H(y) - H(y | s) = - \sum_y P(y) \log_2 P(y) + \sum_{s,y} P(s) P(y | s) \log_2 P(y | s), \quad (3)$$

where $H(y)$ is the total entropy of the response, $H(y | s)$ can be defined as noise entropy, accounting for the variability in the response that is not due to variations in the stimulus, but comes from other noise sources. For simplicity, we consider the noiseless case, where $H(y | s) = 0$. In this case, the mutual information is simply equal to the output entropy $I_m = H(y)$.

We are interested in computing the entropy of the output y . The distribution function of y (also refereed to as the cumulative density function) is

$$D(y) = P[Y \leq y] = \begin{cases} 0, & \text{if } y < 0, \\ \int_{-\infty}^{y+\theta} p(x)dx, & \text{if } 0 \leq y < \eta - \theta, \\ 1, & \text{if } y \geq \eta - \theta. \end{cases} \quad (4)$$

In order to compute the entropy of y , we approximate y by a discrete random variable \hat{y} , which we define by $P[\hat{y} = i \Delta y] = P[y \in [(i-1) \Delta y, i \Delta y]]$, where Δy is the quantization resolution of y , and i is an integer from 0 to $\lceil (\eta - \theta)/(\Delta y) \rceil$. Note that $P[y \in (a, b)]$ can be computed analytically as $D(b) - D(a)$. The mutual information, which is equal to entropy for the noiseless case, is then computed as in Eq.[3]. It can be shown that changes in Δy affect the entropy of \hat{y} only by an additive constant [10]. Since we are only interested in comparing the values of mutual information I_m as we change σ , θ , and η , our choice of Δy does not affect our results. Here, we set $\Delta y = 1$.

Figure 2b (blue line) shows mutual information I_m varies with σ , exhibiting a tuning curve, with a global maximum at a particular σ . This maximum is

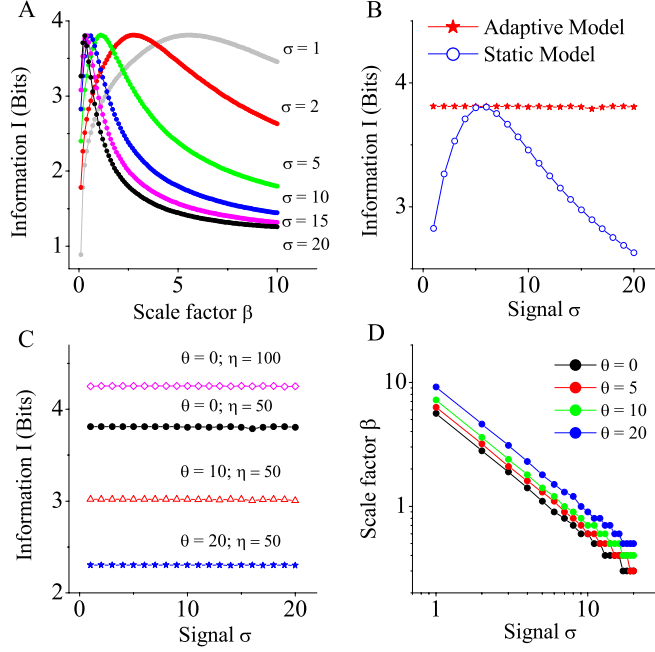


Fig. 2. (a) For an adaptive LN model with threshold $\theta = 0$ and saturation $\eta = 50$, mutual information I as a function of rescaling factor β for signals of different σ 's. (b) The mutual information I as a function of σ for the adaptive LN model and the static LN model. (c) Mutual information I as a function of σ for various model parameters with β chosen to maximize I for each σ . (d) Optimal β as a function of σ for four given values of threshold in a log-log plot exhibits an inverse power-law.

determined by the threshold and the saturation levels [2] The lower (or higher) is the threshold (saturation), the larger the maximum of information can be transmitted (see Fig.2c). Can rescaling the gain of the linear kernel β restores the information transmission to its maximum possible level at different σ 's? This is desirable because the capacity of the system will then be fully utilized regardless of the input variance. We found that rescaling β alone can indeed achieve this purpose. Fig. 2a shows that for each σ , mutual information varies with the scaling factor β and that there always exists an optimal β to restore the information transmission rate of the system to the maximum level which is constrained only by the threshold and the saturation level (Fig 2c). This gain rescaling allows the system's information transmission rate to be maintained at optimal regardless of the input variance (Fig 2b). The optimal β decreases monotonously with an increase in σ . Interestingly, in the log-log plot, β and σ displays a inverse power law relationship, the fitting slope is -1 (Fig. 2d).

The response gain γ , i.e., the amplitude of the measured effective kernel, now equal to $\alpha * \beta$ (see Eq.[2]), also displays a scale invariant power law relationship with a scale -1 with σ (Fig.3a). Noted that, without gain rescaling, i.e., $\beta = 1$ is fixed, the gain γ varies with σ , displaying a bell-like tuning

curve, which emerges simply from static-nonlinearity. Gain rescaling is crucial for maintaining mutual information constant with variation in σ . There is experimental evidence supporting this inverse-power law between effective kernel amplitude gain and the signal contrast in the neurophysiological literature. Fig.3b is adapted from Fig.5a in Ref.[6], which shows that response gain of cat simple cell varying with the contrast of monocular input also in power law relationship in the log-log plot. A contrast gain slope of -1 indicates that the gain control system is 100% effective, whereas a slope of 0 indicates that the system is completely ineffective. The cells in Fig. 3b showed a slope around -0.75 suggesting the neurons might be trying to re-scale gain adaptively in order maximize information transmission.

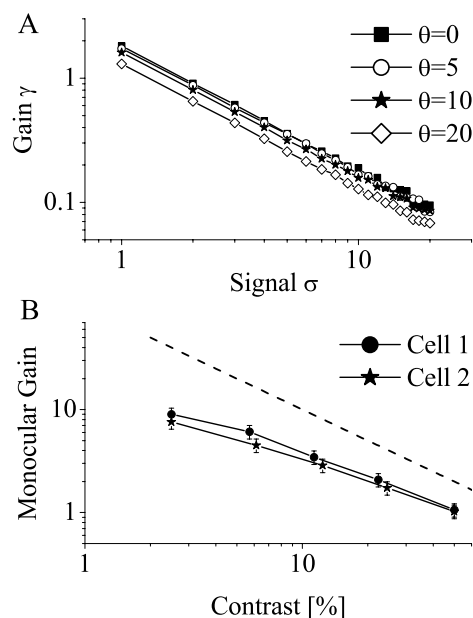


Fig. 3. Gain γ as a function of σ for four given values of threshold with β optimized for information transmission. (b) Monocular gain in two cat simple cells as a function of stimulus contrast (from 2.5% to 50%) in the log-log axes; for comparison, a line with slope of -1 is plotted in dashed line. Adapted from Fig 5. of Ref. [5].

3 Discussion

The power law relationship observed experimentally and the theoretical findings of this paper suggests that there might exists an intrinsic adaptation mechanism in single neurons or network to maximize the information encoding. Such an adaptation, which is independent of static nonlinearity, involves mainly changing of the amplitude of the linear transfer function. Biophysically, a single neuron itself is a circuit, whose transfer function is controlled by the ionic conductances and capacitances. Recent experiments showed that

the activation of Na^+ -activated and Ca^{2+} -activated K^+ currents inside single cells or the adjustment of voltage dependent conductance might play an important role in contrast adaptation [11-12]. These mechanisms might serve to control the gain of the linear kernel part of the LN cascade.

In summary, we concluded that while static nonlinearity, such as threshold and saturation, is sufficient to modulate gain as a function of input variance, the information capacity of the neurons cannot be fully utilized until the gain of the linear kernel is also adjusted intelligently according to the input variance. Thus, the contrast gain adaptation phenomena observed in neurophysiological experiments can potentially be factored into a component due simply to the static nonlinearities of the neurons [1-2] and another component due to an adaptive rescaling of the gain in the linear transfer function in the LN cascade for the purpose of information maximization.

References

- [1] Y.G. Yu and T.S. Lee, Dynamical mechanisms underlying contrast gain control in single neurons *Phys. Rev. E* 68 (2003) 1901-1907.
- [2] Y.G. Yu, T.S. Lee, and B. Potetz, Neuronal design principles for contrast gain control and information maximization. *Phys. Rev. Lett.* Submitted, (2003).
- [3] R.M. Shapley and J.D. Victor, The contrast gain control of the cat retina. *Vision Res.* 19 (1979) 431-434.
- [4] D. Chander and E.J. Chichilnisky, Adaptation to temporal contrast in primate and salamander retina. *J. Neurosci.* 21 (2001) 9904-9916.
- [5] S.A. Baccus and M. Meister, Fast and slow contrast adaptation in retinal circuitry. *Neuron* 36 (2002) 909-919.
- [6] A.M. Truchard, I. Ohzawa, and R.D. Freeman, Contrast gain control in the visual cortex: monocular versus binocular mechanisms. *J. Neurosci.* 21 (2000) 287-299.
- [7] N. Brenner, W. Bialek, and R. de Ruyter van Steveninck, Adaptive rescaling maximizes information transmission, *Neuron*, 26 (2000) 695-702.
- [8] P. Adorjan, C. Piepenbrock, and K. Obermayer, Contrast adaptation and infomax in visual cortical neurons. *Rev. Neurosci.* 10 (1999) 181C200.
- [9] C.E. Shannon and W. Weaver, *The Mathematical Theory of Communication* (Univ. of Illinois Press, Urbana, IL, 1949).
- [10] P. Dayan and L. F. Abbott, *Theoretical Neuroscience* (MIT Press, Cambridge, 2001), Chap. 4, pp.129.
- [11] M.V. Sanchez-Vives, L.G. Nowak, and D.A. McCormick, Cellular mechanisms of long lasting adaptation in visual cortical neurons in vitro. *J. Neurosci.* 20 (2000) 4286 C4299.
- [12] M. Stemmler and C. Koch, How voltage-dependent conductances can adapt to maximize the information encoded by neuronal firing rate. *Nature Neurosci.* 2 (1999) 521-527.