



# Determine Neuronal Tuning Curves by Exploring Optimum Firing Rate Distribution for Information Efficiency

Fang Han<sup>1,2</sup>, Zhijie Wang<sup>1\*</sup> and Hong Fan<sup>3</sup>

<sup>1</sup> Department of Automation, College of Information Science and Technology, Donghua University, Shanghai, China, <sup>2</sup> Center for Neural Science, New York University, New York, NY, USA, <sup>3</sup> Department of Information Management System, Glorious Sun School of Business and Management, Donghua University, Shanghai, China

This paper proposed a new method to determine the neuronal tuning curves for maximum information efficiency by computing the optimum firing rate distribution. Firstly, we proposed a general definition for the information efficiency, which is relevant to mutual information and neuronal energy consumption. The energy consumption is composed of two parts: neuronal basic energy consumption and neuronal spike emission energy consumption. A parameter to model the relative importance of energy consumption is introduced in the definition of the information efficiency. Then, we designed a combination of exponential functions to describe the optimum firing rate distribution based on the analysis of the dependency of the mutual information and the energy consumption on the shape of the functions of the firing rate distributions. Furthermore, we developed a rapid algorithm to search the parameter values of the optimum firing rate distribution function. Finally, we found with the rapid algorithm that a combination of two different exponential functions with two free parameters can describe the optimum firing rate distribution accurately. We also found that if the energy consumption is relatively unimportant (important) compared to the mutual information or the neuronal basic energy consumption is relatively large (small), the curve of the optimum firing rate distribution will be relatively flat (steep), and the corresponding optimum tuning curve exhibits a form of sigmoid if the stimuli distribution is normal.

## OPEN ACCESS

### Edited by:

Petr Lansky,  
Academy of Sciences of the Czech  
Republic, Czechia

### Reviewed by:

Paolo Massobrio,  
University of Genoa, Italy  
Alexander G. Dimitrov,  
Washington State University  
Vancouver, USA

### \*Correspondence:

Zhijie Wang  
wangzj@dhu.edu.cn

**Received:** 20 October 2016

**Accepted:** 06 February 2017

**Published:** 21 February 2017

### Citation:

Han F, Wang Z and Fan H (2017)  
Determine Neuronal Tuning Curves by  
Exploring Optimum Firing Rate  
Distribution for Information Efficiency.  
*Front. Comput. Neurosci.* 11:10.  
doi: 10.3389/fncom.2017.00010

**Keywords:** neuronal tuning curve, information efficiency, optimum firing rate distribution, mutual information, energy consumption

## INTRODUCTION

Neuronal systems are assumed to be optimized for information encoding after millions of years of natural selection, in the sense that the capacity of the neural systems (or channels in the language of information theory) for information transmission is occupied as much as possible, leaving as little as possible the “power” of the channels being unutilized. This is the “redundancy reduction” hypothesis proposed by Barlow (1959, 1961). Based on this assumption, the optimum models of the neural systems are obtained using information theory (Atick and Redlich, 1990; Borst and Theunissen, 1999; Bethge et al., 2003; Nikitin et al., 2009; McDonnell et al., 2011; Rolls and Treves, 2011; Wei and Stocker, 2016). These optimum models coincide with the existing results well, implying that the way how neural systems process information can be understood and the models of the neural systems can be constructed by the information theory.

As energy consumption of the brain occupies a large part of the total energy consumption of organisms (Erecinska et al., 2004), some studies optimized neural systems by maximizing

the information efficiency, i.e., maximizing the ratio of the mutual information to the energy consumed by the neuron for emitting spikes (Levy and Baxter, 1996; Wang and Zhang, 2007; Torreal dea Francisco et al., 2009; Berger and Levy, 2010; Sengupta and Stemmler, 2014; Sengupta et al., 2014). Some studies (Levy and Baxter, 1996; Attwell and Laughlin, 2001) further elaborated the definition of the information efficiency by taking account of two parts of energy consumption of a neuron. Besides the part of the energy consumed by the neurons to fire spikes (Yu and Liu, 2014), they considered another part of energy consumption, called basic energy consumption, which is relevant to the metabolic cost required to support the living of the neurons or the sub-threshold activity of the neurons. Evidently, the first part of energy consumption is proportional to the number of the spikes, while the second part of the energy consumption has no relationship with the number of the spikes.

Tuning curve describes the relationship between the stimulus and the output of the neuron, which plays a very important role in neuronal information encoding (McDonnell and Stocks, 2008; Nikitin et al., 2009; Yaeli and Meir, 2010; Day et al., 2012; Wang et al., 2013; Han et al., 2015; Yarrow and Series, 2015). Tuning curve is optimized for information encoding to test the “redundancy reduction” hypothesis at the single neuron level (Laughlin, 1981). Considering the energy consumption, tuning curves should be optimized to enable the neurons to encode information efficiently, i.e., to have high ratio of information to energy consumption. To obtain tuning curves in this way, the following three elements need to be considered: the calculation of energy consumption (Levy and Baxter, 1996; Wang and Zhang, 2007; Torreal dea Francisco et al., 2009; Berger and Levy, 2010; Sengupta and Stemmler, 2014; Sengupta et al., 2014), the definition of the information efficiency (Levy and Baxter, 1996; Moujahid et al., 2011; Kostal and Lansky, 2013; Sengupta et al., 2014), and the probability distribution of inputs (Dayan and Abbott, 2001; Nikitin et al., 2009). Most of the existing methods for determining optimum tuning curves assumed simplified situations of above-mentioned three elements, especially the first two. In some methods, energy consumption was not taken into account. For example, early work on the optimization of the tuning curve (Laughlin, 1981) by entropy maximization validates the “redundancy reduction” hypothesis; a general method for determining tuning curves for maximizing mutual information was proposed in McDonnell and Stocks (2008); the optimum tuning curve for maximizing mutual information was found to have a discrete structure (Nikitin et al., 2009); the optimal tuning functions for minimum mean square reconstruction from neural rate responses were derived in Bethge et al. (2003). In other methods the number of spikes was regarded as the energy consumption, and the ratio of the mutual information to the number of the spikes was used as the information efficiency (Moujahid et al., 2011; Kostal and Lansky, 2013). However, in real neural systems, things seem to be more complex. For example, two parts of the energy consumption rather than only spikes should be considered as discussed in the previous paragraph. As for the definition of the information efficiency, one should consider the fact that in some situations, mutual information is more important than energy consumption, while

in some other cases, energy consumption should be underlined. The information and the energy consumption should also be measured in the same scale when calculating information efficiency, despite that information is measured on logarithm scale while energy on linear scale in most existing studies. Furthermore, the probability distribution of the inputs in many cases may take a complex form other than normal distribution. How can we determine the neuronal tuning curves with these complex situations considered?

It is usually difficult to obtain analytical solutions to the tuning curves when the complex situations discussed in the previous paragraph are considered. This paper proposes a new computer algorithm to deal with these complex situations when optimizing neuronal tuning curves. Firstly, the optimum spike count response distribution (the probability distribution of the numbers of the spikes emitted by the neuron for different inputs, which is explained in details in Section The model for information-efficiency in neuronal encoding system) is analyzed in terms of full entropy and energy consumption, and a combination of exponential-based functions is designed for it. Then the optimum firing rate distribution (see the detailed explanations in Section The model for information-efficiency in neuronal encoding system) for the information efficiency is explored. Based on the analysis of the relationship between the optimum spike count response distribution and the optimum firing rate distribution, a combination of exponential-based functions is designed for the optimum firing rate distribution, after the dependency of the noise entropy on the shape of the functions of the firing rate distribution is analyzed. A forward-backward rapid algorithm with variable step size is proposed for searching the optimum neuronal firing rate distribution function. It is found with the rapid algorithm that a combination of two different exponential functions with two free parameters can describe the optimum firing rate distribution accurately. The rapid algorithm is then used to search the two parameter values of the optimum neuronal firing rate distribution function. Finally, the neuronal tuning curves are calculated based on the optimum neuronal firing rate distribution and the stimuli probability distribution. The paper is organized as follows. The basic concept of the optimum neuronal encoding scheme is described in Section The model for information-efficiency in neuronal encoding system. A new method to search the optimum neuronal firing rate distribution is proposed in Section Search for the optimum neuronal firing rate distribution. The optimum tuning curves are calculated based on the optimum neuronal firing rate distribution in Section The tuning curves based on the optimum neuronal firing rate distribution. Finally, a brief summary is provided in Section Conclusions.

## THE MODEL FOR INFORMATION-EFFICIENCY IN NEURONAL ENCODING SYSTEM

A neuron receives inputs from the environment and encodes them in the spikes of the neuron. For each input  $s$ , the neuron responses with  $n$  spikes within time window  $T$  (we suppose that

neurons must complete the encoding process within this short time period to get a rapid response to the external stimulus) (Bethge et al., 2003). Since neuronal responses are naturally random, the value of  $n$  is different trial by trial even the same input  $s$  is provided, i.e.,  $n$  is a random variable, which is often assumed to obey Poisson distribution for simplification (Dayan and Abbott, 2001; Bethge et al., 2003; Nikitin et al., 2009). The rate of this Poisson distribution is called the firing rate of the neuron. Let  $r = g(s)$  be the tuning curve of the neuron, which we want to determine according to information theory in this paper. Suppose the input  $s$  appears with probability  $P(s)$ , which can often be measured by recording the frequency of the occurrence of the stimulus  $s$  (The frequency of the occurrence of the stimulus between  $s$  and  $s + \Delta s$  is recorded for  $P(s)$  if the stimulus is continuous). If we know the probability density of the firing rate  $p(r)$ , i.e., the firing rate distribution, then we can calculate the tuning curve,  $r = g(s)$ , numerically (see the details in Section The tuning curves based on the optimum neuronal firing rate distribution). How can we get the firing rate distribution  $p(r)$ ? We now discuss this according to the assumption that a neuron is an optimum encoding system (In some situations, the spike sequence emitted by the neuron could be recorded and used to calculate  $p(r)$ . This  $p(r)$  could be used to validate the redundancy reduction hypothesis by comparing it with the  $p(r)$  that is got by information efficiency maximization).

As we discussed in the previous section, the assumption that a neuron is an optimum encoding system means that the neuron can convey a large amount of information on one hand, while consume a small amount of energy on the other hand. Stated in other words, the neuron exhibits optimum information encoding efficiency. The amount of information conveyed by the neuron (mutual information) (Kostal, 2010; Gao et al., 2014) can be described by  $I_m = S_{full} - S_{noise}$ , where  $S_{full}$  is the full entropy of the neuronal response (see Equation 2) and  $S_{noise}$  (see Equation 3) is the noise entropy. As for the energy consumption, the basic energy consumption of the neuron within the period  $T$  is  $E_b$  and the energy consumed by the neuron to emit spikes is  $E_s$ , with total energy consumption  $E = E_b + E_s$ . We measure the energy consumption in units of number of spikes. For example,  $E_b = 2$  means that the basic metabolic cost required to support the living of the neuron within  $T$  is equivalent to the energy consumed by the neuron to emit 2 spikes.

A simple but widely used definition of the information efficiency,  $I_E$ , is the ratio of mutual information to the energy consumed by neurons to emit spikes, i.e.,  $I_E = \frac{I_m}{E_s}$  (Levy and Baxter, 1996; Moujahid et al., 2011; Sengupta and Stemmler, 2014). This definition is extended in this paper, which is described by Equation (1).

$$I_E = (2^{I_m} - 1)/(E_s + E_b)^c \quad (1)$$

This extension is motivated by the following three considerations. The first is that energy consumption is composed of two parts in real neurons as discussed previously. Thereby  $E = E_s + E_b$  (Levy and Baxter, 1996; Attwell and Laughlin, 2001) instead of  $E_s$  is used in the calculation of  $I_E$ . The second is that in some situations, mutual information (energy consumption) may be

considered more important than energy consumption (mutual information). The exponential  $c$  in Equation (1) is used to model this relative importance in these situations. When  $c < 1$ , mutual information is considered more important; when  $c > 1$ , energy consumption is more focused on. The third is that mutual information is measured in bits, therefore it should be transformed into linear representation as energy consumption. That's why we use  $2^{I_m} - 1$  for calculation of  $I_E$ .

Suppose the inputs are discretized as  $s(i) = i\Delta s$ , where  $i = 1, 2, \dots, M$ ,  $\Delta s = 0.0001$  and  $M = 10,000$  (the stimuli are normalized in this paper). The probability of the occurrence of  $s(i)$  is  $P(s(i)) = (p(s)|s(i))\Delta s$ . The firing rates of neurons are discretized as  $r(i) = r(i-1) + \Delta r$ ,  $i = 1, 2, \dots, N$ ,  $\Delta r = 0.1$  and  $N = 5,000$ . The probability of the occurrence of  $r(i)$  is  $P(r(i)) = (p(r)|r(i))\Delta r$ . Suppose the probability that the neuron emits  $n$  spikes within time window  $T$  is  $P(n)$  (we name  $P(n)$  the spike count response distribution in this paper), then the full entropy of the neuron response is

$$S_{full} = \sum_n P(n) \log_2 P(n), \quad (2)$$

and the noise entropy of the firing rate of the neuron is

$$\begin{aligned} S_{noise} &= \sum_j \sum_n P(s(j))P(n|s(j)) \log_2 P(n|s(j)) \\ &= \sum_j \sum_n P(r(j))P(n|r(j)) \log_2 (P(n|r(j))), \end{aligned} \quad (3)$$

where  $P(n|s(j))$  is the conditional probability and  $r(j) = g(s(j))$ .

## SEARCH FOR THE OPTIMUM NEURONAL FIRING RATE DISTRIBUTION

If the optimum firing rate distribution for information efficiency and the input distribution are given, one can calculate the tuning curve numerically (it will be discussed in Section The tuning curves based on the optimum neuronal firing rate distribution). In this section, we will discuss how to determine the optimum firing rate distribution for information efficiency. To achieve this, we first explore the optimum spike count response distribution for the entropy efficiency (entropy efficiency considers full entropy and energy consumption) in Section Optimum spike count response distribution for entropy efficiency. In Section Optimum neuronal firing rate distribution for entropy efficiency, we determine optimum neuronal firing rate distribution for the entropy efficiency based on the assumption that the number of the firing of the neuron obeys Poisson distribution when firing rate is fixed. We find using the non-linear least square method that the optimum firing rate distribution for the entropy efficiency has the same form of function as the optimum spike count response distribution but having different parameter values. In Section Optimum neuronal firing rate distribution for information efficiency, to extend the entropy efficiency to information efficiency, the noise entropy is included in the framework (note that entropy efficiency considers full entropy

but information efficiency considers mutual information) and we found that the form of the firing rate distribution function for the maximum information efficiency is the same as that for entropy efficiency. Finally, in Sections Algorithm for searching the optimum neuronal firing rate distribution and Searching the optimum neuronal firing rate distribution, a rapid algorithm is proposed to calculate the optimum firing rate distributions.

### Optimum Spike Count Response Distribution for Entropy Efficiency

Let us begin with a simple case, entropy efficiency, which only considers full entropy of the spike count neuronal response and energy consumption. Similar to the information efficiency, we define entropy efficiency,  $IS$ , as  $IS = (2^{S_{full}} - 1)/(E_s + E_b)^c$ . As we know, the neuronal response should be uniformly distributed for the maximum entropy, i.e.,  $P(n) = 1/n_{max}$  where  $n_{max}$  is the maximum spike number of the neuron within the measuring time window  $T$ . But maximum entropy does not mean maximum entropy efficiency, because the value of energy consumption is relatively large if  $P(n) = 1/n_{max}$  because  $E_s = \sum_n P(n)n$ . As

$\frac{dE_s}{dP(n)} = n$ , to reduce energy consumption, the function  $f(n) = P(n)$  should be smaller with larger  $n$ . Namely, to decrease the energy consumption, the function  $f(n)$  should decrease hardly to 0, i.e.,  $f(n)$  should decrease monotonously, and the rate of the decrease should also decrease with the increase of  $n$ . Thereby we can conceive that exponential-like function can describe such functions that lead to small energy consumption and large entropy. Since  $\sum_n f(n) = 1$ , the simplest function of  $f(n)$  is

$$f(n) = \alpha e^{-\alpha n}. \tag{4}$$

The entropy will be smaller when  $f(n)$  deviates more from uniform distribution. Therefore, if the parameter  $\alpha$  is large then  $f(n)$  will decay rapidly, which will result in less energy consumption ( $E_s$ ) but small full entropy (see **Figure 1**); otherwise if  $\alpha$  is small, then  $f(n)$  will decay slowly, which will result in large energy consumption ( $E_s$ ) and large full entropy (see **Figure 1**). If  $\alpha$  is very small,  $f(n)$  will be almost flat leading to maximum full entropy. Therefore, both energy consumption and full entropy decrease with the increase of  $\alpha$ , but the rates of the decrease are different, which enables the optimum value of  $\alpha$  leading to maximum information efficiency (see Section Algorithm for searching the optimum neuronal firing rate distribution for details).

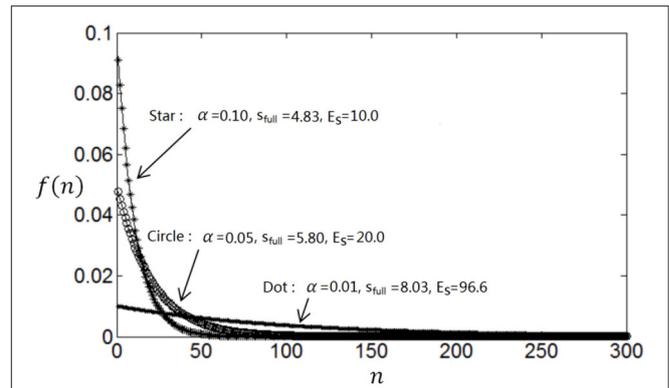
The functions  $f(n) = \alpha e^{-\alpha n}$  have only one free parameter  $\alpha$ . This restrains the possible shapes of the spike count response distribution functions. To make the possible shapes of the distribution functions more various, we use a function that is a combination of two exponential functions with two free parameters as Equation (5).

$$f(n) = (\alpha e^{-\alpha n} + \beta e^{-\beta n})/2. \tag{5}$$

We can also use a function with three free parameters as Equations (6) or (7).

$$f(n) = (\alpha e^{-\alpha n} + \beta e^{-\beta n} + \gamma e^{-\gamma n})/3, \tag{6}$$

$$f(n) = \alpha e^{-\alpha n} \gamma + \beta e^{-\beta n} (1 - \gamma). \tag{7}$$



**FIGURE 1 |** Dependency of full entropy and energy consumption on the parameter  $\alpha$ .

Note that  $n$  is an integer, the values of  $\gamma$  in Equation (7) lie between 0 and 1, and for all these functions,  $\sum_n f(n) = 1$ . We

design the function as such a form that both full entropy and  $E_s$  decrease monotonically with the increase of either  $\alpha$  or  $\beta$ . This allows us to design a rapid algorithm (see Section Algorithm for searching the optimum neuronal firing rate distribution) to search the optimum parameters.

### Optimum Neuronal Firing Rate Distribution for Entropy Efficiency

If the response of the neuron is non-random, then average spike number over multiple observations (firing rate) is identical to the spike number measured in a single observation. Thereby, firing rate distribution is identical to the spike count response distribution. However, neuronal response is random, which usually follows Poisson distribution (Dayan and Abbott, 2001; Bethge et al., 2003; Nikitin et al., 2009). It is interesting to see that the function of firing rate distribution takes the same form as that of spike count response distribution but with different parameters, when this randomness is taken into account. That is, if we use a two-free-parameter function, the function of firing rate distribution may be described by Equation (8) [note that unlike the neuronal response function of Equation (5) where  $n$  is an integer,  $r$  is a continuous variable in Equation (8)]:

$$f(r) = (\alpha e^{-\alpha r} + \beta e^{-\beta r})/2. \tag{8}$$

This can be confirmed by the nonlinear least square method. For example, the curve with blue dot in **Figure 2** is the spike count response distribution produced by Equation (5) with parameters  $\alpha = 0.1$  and  $\beta = 0.5$ . We denote this curve as  $f(n)$ . We search the best firing rate distribution denoted as  $f(r)$ , which can produce a spike count response distribution fitting the blue dot curve ( $f(n)$ ) most perfectly using the nonlinear least square method. We got a function of Equation (8) with  $\alpha = 0.0935$  and  $\beta = 0.5882$  for the firing rate distribution, which is shown by the curve with green star in **Figure 2**. The curve with red square, which we denote as  $\tilde{f}(n)$ , is the spike count response distribution that is produced by this firing rate distribution (note that if the firing rate  $r$  is fixed, the

neuronal spike count response  $n$  obeys Poisson distribution). Red square curve ( $\tilde{f}(n)$ ) fits blue dot curve ( $f(n)$ ) perfectly, implying that the optimum firing rate distribution for the entropy efficiency can also be described by the continuous version of Equation (5), i.e., Equation (8). We use  $Err = \sum_n (f(n) - \tilde{f}(n))^2$  to characterize the distance of the two neuronal spike count response functions. It is shown from  $Err = 1.88 \times 10^{-5}$  for the red square curve and blue dot curve that the two curves are almost identical. More experimental results are shown in **Table 1**. The first line of **Table 1** shows  $f(n)$  described by Equation (5) with various parameter values of  $\alpha$  and  $\beta$ ; the second line shows  $f(r)$  which is described by the continuous version of Equation (5) (Equation 8) but having different parameter values of  $\alpha$  and  $\beta$ ; and the third line shows  $Err$  (note that  $Err$  is not the distance between  $f(n)$  and  $f(r)$ , instead it is the distance between  $f(n)$  and  $\tilde{f}(n)$ ). The results in **Table 1** confirm that if the optimum neuronal spike count response can be described by Equation (5), the optimum firing rate can also be described by continuous version of Equation (5) (Equation 8).

### Optimum Neuronal Firing Rate Distribution for Information Efficiency

Since mutual information of the neuronal response is the actual amount of information encoded in the output of the neuron, it is more important to discuss information efficiency than entropy efficiency. Based on the entropy efficiency discussed in

the previous two subsections, we consider the noise entropy in the following study and discuss the optimum firing rate distribution for information efficiency, i.e., determine the form of the function of the neuronal firing rate distribution that results in more mutual information but less energy consumption. Let us first consider what kind of neuronal firing rate distribution can lead to less noise entropy. We first divide the noise entropy  $S_{noise}$  into components with each component associated with individual firing rate  $r$ . According to Equation (3),  $S_{noise} = \sum_j \sum_n P(r(j))P(n|r(j))\log_2 P(n|r(j))$ . Let

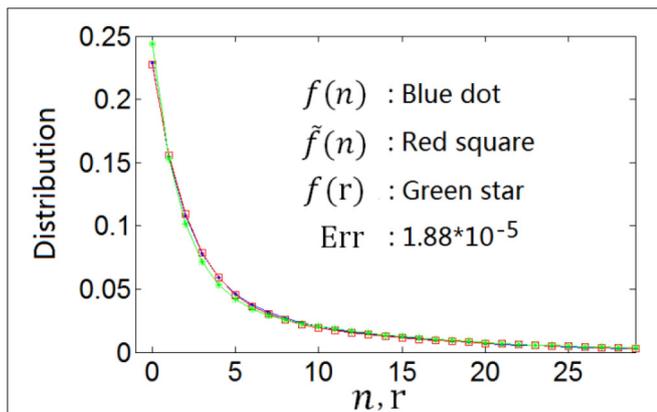
$$S_{noise} = \sum_j P(r(j))N(r), \tag{9}$$

where  $N(r)$  represents the increase rate of the noise entropy to the probability  $P(r(j))$ , we have

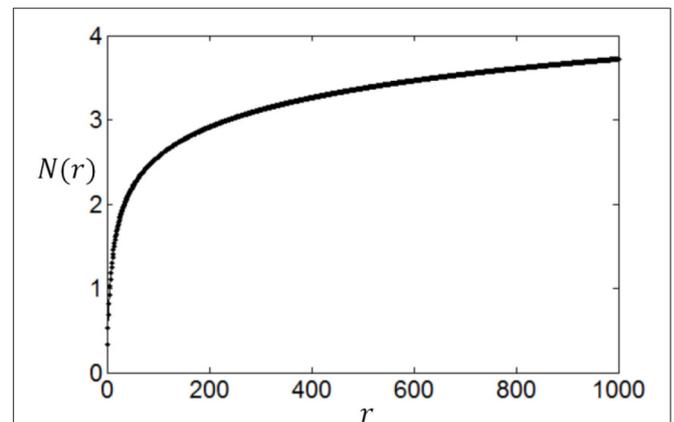
$$N(r) = \sum_n P(n|r(j)) \log_2 P(n|r(j)), \tag{10}$$

where  $P(n|r(j))$  is the possibility that the neuron emits  $n$  spikes when the firing rate of the neuron is  $r(j)$ .  $P(n|r(j))$  follows Poisson distribution as  $P(n|r(j)) = e^{-r(j)}r(j)^n/n!$ .

We compute  $N(r)$  with different  $r$  and plot the dependency of  $N(r)$  on  $r$  in **Figure 3**. It can be seen from **Figure 3** that the noise entropy component  $N(r)$  increases with  $r$ . Therefore, we can infer according to Equation (9) and **Figure 3** that if the value of  $P(r)$  decreases with the increase of  $r$  then  $S_{noise}$  will be small, very similar to the fact that if the value of  $P(r)$  decreases with the increase of  $r$  then energy consumption will be



**FIGURE 2 |** The optimum spike count response distribution (blue dot) described by Equation (5) and the firing rate distribution (green star) searched by the nonlinear least square method which produces the spike count response distribution (red square) fitting the optimum spike count response perfectly.



**FIGURE 3 |** The dependency of increase rate of noise entropy on the firing rate.

**TABLE 1 |**  $f(n)$  with various values of  $(\alpha, \beta)$ ,  $f(r)$  with values of  $(\alpha, \beta)$  and  $Err$ .

$f(n)$	(0.10, 0.50)	(0.20, 0.50)	(0.30, 0.50)	(0.40, 0.50)	(0.20, 0.60)	(0.30, 0.80)	(0.30, 0.90)
$f(r)$	(0.09, 0.59)	(0.21, 0.62)	(0.34, 0.64)	(0.49, 0.65)	(0.21, 0.76)	(0.32, 1.12)	(0.31, 1.27)
$Err$	$1.9 \times 10^{-5}$	$8.6 \times 10^{-7}$	$1.1 \times 10^{-7}$	$8.5 \times 10^{-8}$	$3.9 \times 10^{-6}$	$4.4 \times 10^{-6}$	$1.1 \times 10^{-5}$

small. Thus, we conclude that a firing rate distribution function capable of producing high full entropy, low noise entropy and less energy consumption should have a shape like that in **Figure 1** (or **Figures 7–9**). Namely, we can use the functions of Equations (4)–(7) to describe the optimum firing rate distribution function for the maximum information efficiency.

It is worthy of noting that by treating the noise entropy in this way, we can see the effect of the noise entropy on the shape of the optimum firing rate distribution clearly, thereby we can extend the application of our method to other neural response models. Our method applies for any neural response models as long as  $N(r)$  has a shape like **Figure 3**, i.e.,  $N(r)$  increases rapidly and then slowly with the increase of  $r$ . For example, we found that a Poisson process with a stochastic refractory period (Bair et al., 1994) and a negative binomial distribution (Goris et al., 2014) produce  $N(r)$  that have similar shape as **Figure 3**. Thereby our method is also applicable to these two neural response models.

### Algorithm for Searching the Optimum Neuronal Firing Rate Distribution

We first give a short summary of our discussion in previous sections. Our idea is to determine the shape of the optimum firing rate distribution by heuristically analyzing the effect of the full entropy, energy consumption and noise entropy on the shape of the optimum firing rate distribution. As it is relatively easier to observe the effect of the full entropy and energy consumption on the shape of the optimum spike count response distribution, we first determine the shape of optimum spike count response distribution with only full entropy and energy consumption taken into account and use a combination of exponential functions to describe the optimum spike count response distribution. Then we confirm by using the nonlinear least square method that the optimum spike count response distribution described by the combination of exponential functions can be produced by a firing rate distribution which can also be described by a combination of exponential functions. We further discuss the effect of the noise entropy on the shape of the optimum firing rate distribution by calculating  $N(r)$  described in Section Optimum neuronal firing rate distribution for information efficiency. Thus, the shape of the optimum of the firing rate distribution is determined. It is interesting to see by numerical calculation in Section Searching the optimum neuronal firing rate distribution that the combination function of a pair of exponents can describe the optimum firing rate distribution well enough when Poisson process is adopted, therefore we use a family of low-parametric functions (combination functions of a pair of exponents) in this paper. Of course, the combination functions of two exponents may not be good enough when other neuronal response models are adopted. Three exponents or other kinds of functions may need to be used to describe the optimum firing rate distribution in other cases.

Poisson process is adopted for the neuron response, therefore optimum firing rate distribution can be described by the functions of Equations (4)–(7). It is worthy of noting that although a family of low-parametric functions is used in the paper, it is difficult to solve the problem analytically.

This is because the object to be optimized (maximized) is Equation (1), i.e., the optimum firing rate distribution is searched to achieve both high exponential-weighted information and low power-weighted energy consumption concurrently. This optimization problem is different from that of mutual information maximization when energy consumption is fixed. Therefore, we next discuss how to design algorithms to search the parameter values of the optimum firing rate distribution functions for the maximum information efficiency in this subsection. Firstly, let us consider the simplest neuronal firing rate distribution function with only one parameter  $\alpha$ . If  $\alpha$  gets larger,  $p(r)$  decreases more sharply with the increase of  $r$ , resulting in a smaller amount of energy consumption, full entropy [note that the flatter of the curve of  $p(r)$ , the larger the full entropy], and noise entropy [note that noise entropy is small if  $p(r)$  is small for large values of  $r$  according to **Figure 3**]. **Figures 4, 5** show that the values of mutual information and energy consumption decrease monotonically with increasing the parameter  $\alpha$ . The larger  $\alpha$  gets, more slowly these quantities decrease with  $\alpha$ . Finally, they approach to their limits, respectively. Therefore, we can design a variable step size scheme to search the optimum parameters.

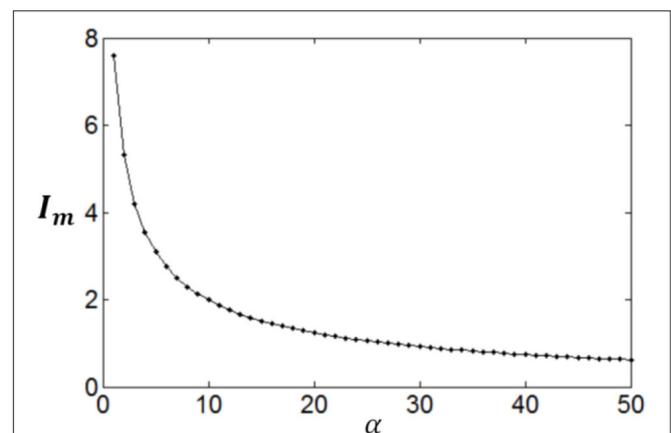


FIGURE 4 | The change of mutual information with the parameter  $\alpha$ .

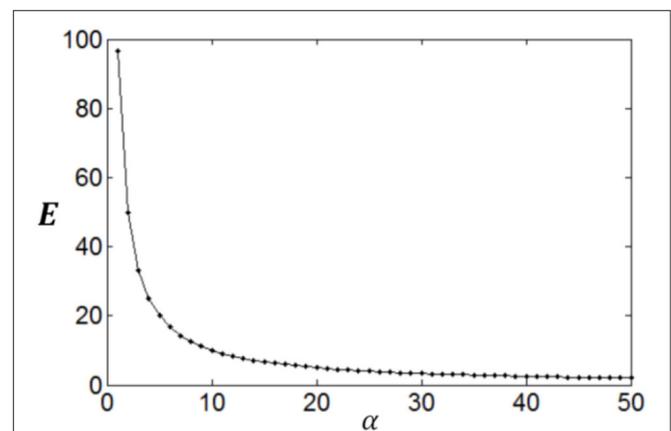


FIGURE 5 | The change of energy consumption with the parameter  $\alpha$ .

Specifically, we can use small step sizes when  $\alpha$  is small, but use relatively large step sizes when  $\alpha$  is large. In this paper, we let  $\alpha(i) = \alpha(i-1) + \alpha(i-1)/\log_2(200 \times \alpha(i-1))$  and  $\alpha(1) = 0.01$ .

Furthermore, it can be seen that although both mutual information and energy consumption decrease with the increase of  $\alpha$ , the decrease rates are different. This is shown in **Figure 6**. When  $\alpha$  is small, the decrease rate of energy consumption is more rapid, meaning that information efficiency increases with the increase of  $\alpha$ . With the increase of  $\alpha$ , the difference between the decrease rates of energy consumption and mutual information becomes smaller and smaller, finally vanishes. Accordingly, the increase rate of information efficiency becomes smaller and smaller, and finally reaches the maximum value as shown in **Figure 6**. To search this maximum point and the corresponding parameter  $\alpha$ , we develop a forward-backward algorithm with forward phase and backward phase described as follows.

Forward phase: starting from the initial point  $\alpha(i)$  and the corresponding information efficiency  $I_E|_{\alpha=\alpha(i)}$ , we evaluate the information efficiency when  $\alpha = \alpha(L+i)$  with  $L > 1$ ,  $I_E|_{\alpha=\alpha(L+i)}$ . If  $I_E|_{\alpha=\alpha(L+i)} > I_E|_{\alpha=\alpha(i)}$ , then evaluate  $I_E|_{\alpha=\alpha(2L+i)}$ , and so on. This process is repeated until  $I_E|_{\alpha=\alpha(QL+i)} < I_E|_{\alpha=\alpha((Q-1)L+i)}$ .

Backward phase: Starting from the point  $\alpha(QL+i)$ , which is obtained in the forward phase, we calculate  $I_E|_{\alpha=\alpha(QL+i-1)}$ . If  $I_E|_{\alpha=\alpha(QL+i-1)} > I_E|_{\alpha=\alpha(QL+i)}$ , then calculate  $I_E|_{\alpha=\alpha(QL+i-2)}$ , and so on. This process is repeated until  $I_E|_{\alpha=\alpha(QL+i-U)} < I_E|_{\alpha=\alpha(QL+i-U+1)}$ . We then get the optimum parameter  $\alpha_{opt} = \alpha(QL+i-U+1)$  and the maximum information efficiency  $I_{E_{max}} = I_E|_{\alpha=\alpha(QL+i-U+1)}$ .

For a firing rate distribution function with two free parameters,  $\alpha_1$  and  $\alpha_2$ , another feature of the distribution function, the symmetry of the distribution function (see Equation 5), can be used to further reduce the computational complexity. Therefore, the algorithm of the search for the two parameters for the optimum information efficiency can be described as follows (the algorithm for a firing rate distribution function with three free parameters is similar to that with two free parameters).

Step 1: Discretize  $\alpha_1$  and  $\alpha_2$  using variable step size scheme discussed in the first paragraph in this subsection. We get discretized values,  $\alpha_1(1) = \alpha_2(1) = 0.01$ ,  $\alpha_1(2) = \alpha_2(2) = 0.02, \dots$ ,  $\alpha_1(4) = \alpha_2(4) = 0.042, \dots$ ,  $\alpha_1(65) = \alpha_2(65) = 51$  (we suppose that the values of the parameters are less than 51) for both  $\alpha_1$  and  $\alpha_2$ .

Step 2: Let  $i = 1$ ;

Step 3: Set  $\alpha_1 = \alpha_1(i)$ ;

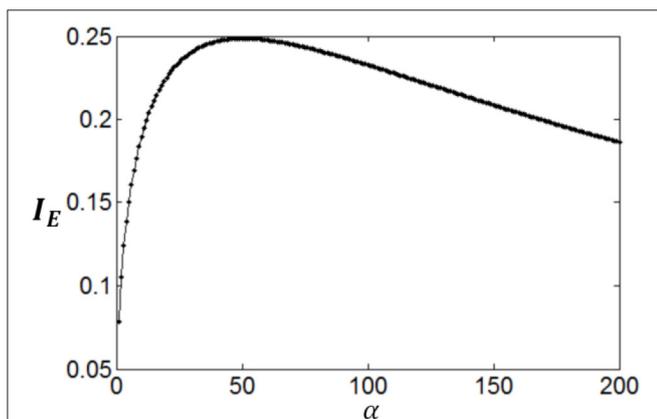
Step 4: Set the initial value of  $\alpha_2 = \alpha_2(i)$  (note we do not set  $\alpha_2 = \alpha_2(1)$  due to the symmetry property);

Step 5: Use the forward and backward scheme for the parameter  $\alpha_2$ . Record the largest information efficiency and the corresponding parameter values of  $\alpha_1$  and  $\alpha_2$ .

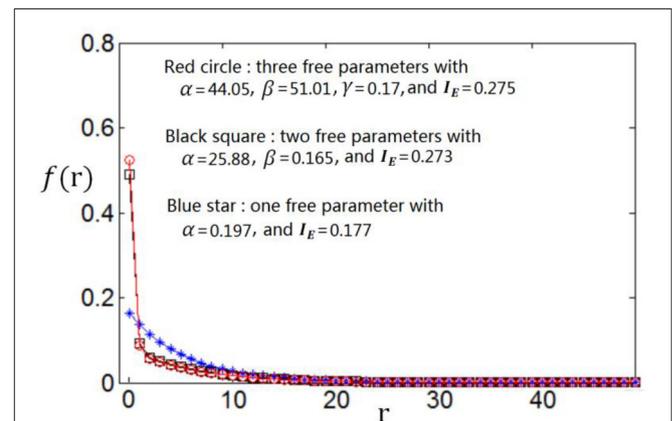
Step 6: Let  $i \leftarrow i + 1$ . If  $i \leq 64$  then goto Step 3.

### Searching the Optimum Neuronal Firing Rate Distribution

We use the proposed algorithm to search the optimum firing rate distribution for the maximum information efficiency. Among the four classes of firing rate distribution functions (Equations 4–7), we first determine which one is the best for searching the maximum information efficiency. It is found from **Figures 7–9** that the maximum information efficiency searched using firing rate distribution functions described by two free parameters is much higher than that using the functions described by only one free parameter. But the maximum information efficiency searched using firing rate distribution functions with three free parameters [see **Figures 7, 8** where Equation (6) is used and **Figure 9** where Equation (7) is used] is almost the same as that searched using the functions with two free parameters. Therefore, we believe that firing rate distribution functions described by two free parameters are good enough to describe the optimum firing rate distribution, and additionally they are of light computational complexity due to the few numbers of free parameters. Thus, we use the firing rate distribution of Equation (5) with two free parameters in the remainder of this paper.



**FIGURE 6 |** The change of information efficiency with the parameter  $\alpha$ ;  $c = 1$  and  $E_b = 1$ .



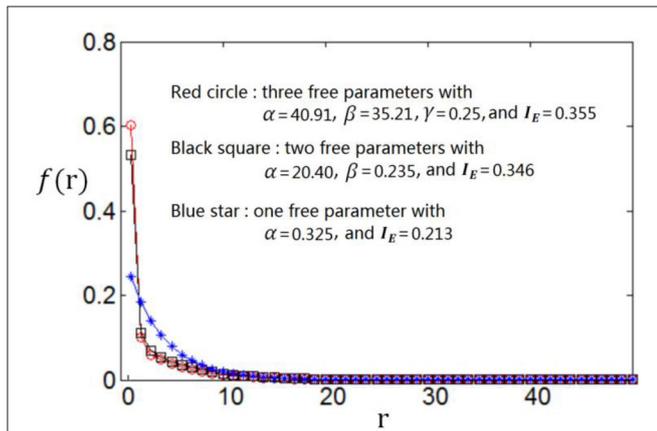
**FIGURE 7 |** Comparisons of the three kinds of firing rate distribution functions with different numbers of free parameters. Equation (6) is used for the firing rate distribution with three free parameters.  $E_b = 2$  and  $c = 1$ .

As the optimum firing rate distribution is different when a different definition of the information efficiency is used, we next explore the dependency of the optimum firing rate distribution on the parameters of the information efficiency. **Figure 10** shows that when exponential index  $c$  becomes smaller, the optimum firing rate distribution will become flatter. This can be explained as follows. A firing rate distribution being relatively flat will result in large mutual information. Therefore, considering that small value of  $c$  means that energy consumption is not very important in evaluating the information efficiency, this relatively flat firing rate distribution will lead to maximum information efficiency, though this relatively flat distribution implies large amount of energy consumption [note  $E_s = \sum_n P(n)n$ ]. **Table 2** shows the dependency of the parameter values of the optimum firing rate distribution functions,  $\alpha$  and  $\beta$ , on the parameter value of the

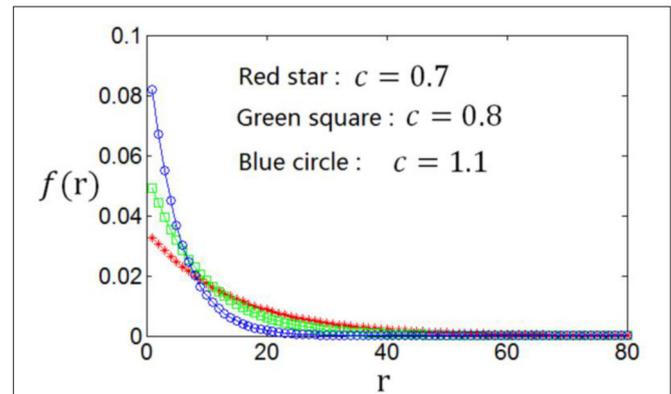
information efficiency  $c$ . **Figure 11** shows that if basic energy consumption of neurons ( $E_b$ ) is increased, the optimum firing rate distribution will become flat. This is not strange because large value of  $E_b$  means that the energy consumed by the neurons to emit spikes is relatively not important. Therefore, the optimum firing rate distribution is close to the one that leads to maximum mutual information, i.e., the optimum firing rate distribution is somehow flat. **Table 3** shows the dependency of the parameter values of the optimum firing rate distribution functions,  $\alpha$  and  $\beta$ , on the parameter values of the information efficiency  $E_b$ .

### THE TUNING CURVES BASED ON THE OPTIMUM NEURONAL FIRING RATE DISTRIBUTION

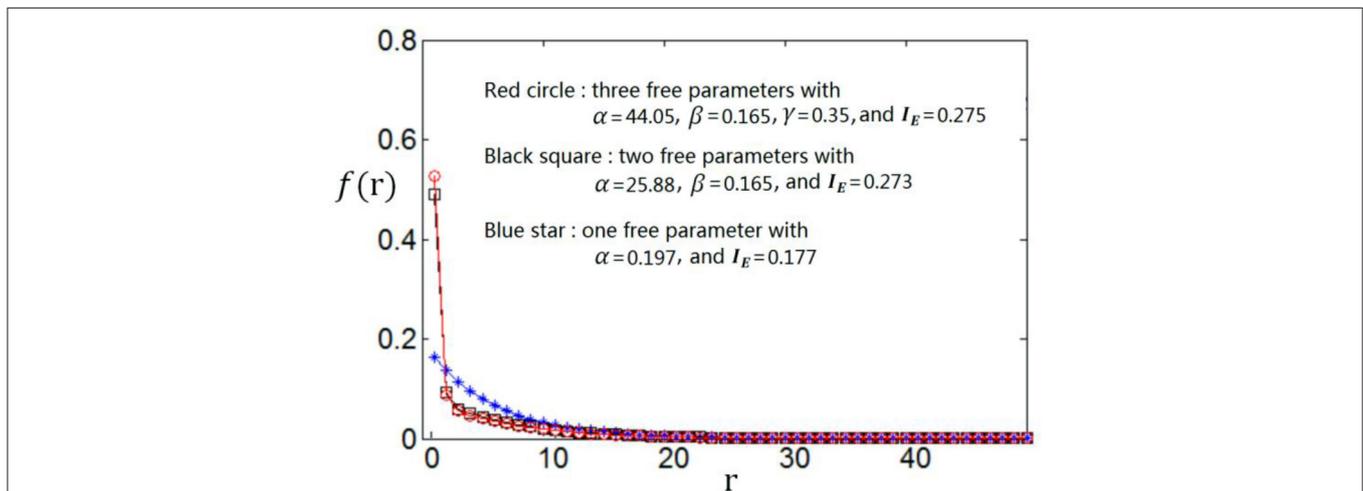
We can numerically compute the tuning curve  $g(\cdot)$  ( $r = g(s)$ ) easily if the probability distribution of the stimuli  $s$ ,  $P(s(i))$ ,



**FIGURE 8 | Comparisons of the three kinds of firing rate distribution functions with different numbers of free parameters.** Equation (6) is used for the firing rate distribution with three free parameters.  $E_b = 1$  and  $c = 1$ .



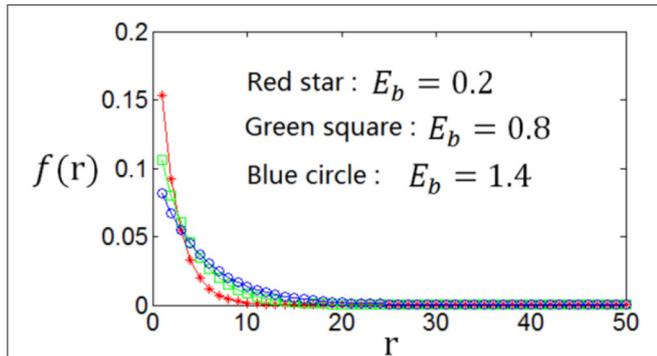
**FIGURE 10 | Dependency of the optimum firing rate distribution on the parameter of the exponential index ( $c$ ) of the information efficiency  $E_b = 2$ .**



**FIGURE 9 | Comparisons of the three kinds of firing rate distribution functions with different numbers of free parameters.** Equation (7) is used for the firing rate distribution with three free parameters.  $E_b = 2$  and  $c = 1$ .

**TABLE 2 | Dependency of the values of  $\alpha$  and  $\beta$  (parameters of the optimum firing rate) on the value of  $c$  (parameter of the information efficiency).**

$c$	0.65	0.70	0.75	0.80	0.90	1.0	1.1	1.2
$(\alpha, \beta)$	(0.06, 47)	(0.07, 41)	(0.09, 38)	(0.11, 33)	(0.14, 28)	(0.16, 26)	(0.20, 24)	(0.23, 22)
$I_E$	0.546	0.484	0.434	0.392	0.324	0.273	0.232	0.198



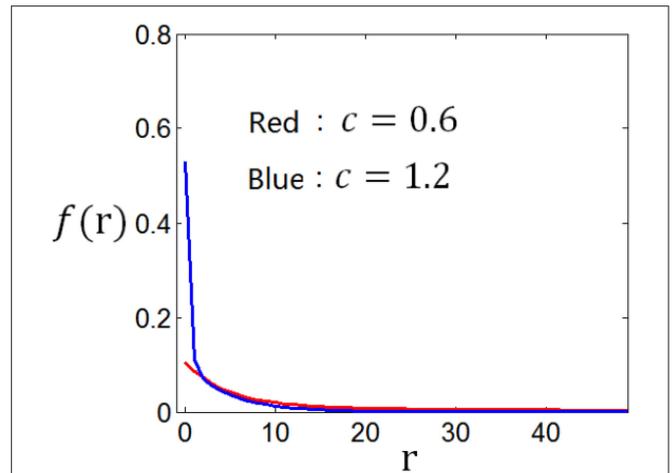
**FIGURE 11 | Dependency of the optimum firing rate distribution on the parameter of the basic energy consumption ( $E_b$ ) of the information efficiency.  $c = 1$ .**

**TABLE 3 | Dependency of the values of  $\alpha$  and  $\beta$  (parameters of the optimum firing rate) on the value of  $E_b$  (parameter of the information efficiency).**

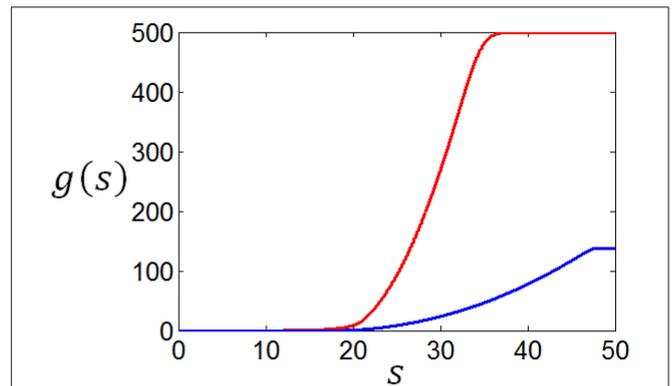
$E_b$	0	0.2	0.4	0.6	0.8	1.2	1.4	1.8
$(\alpha, \beta)$	(0.98, 20)	(0.51, 19)	(0.38, 20)	(0.32, 20)	(0.28, 20)	(0.23, 22)	(0.20, 22)	(0.16, 24)
$I_E$	0.633	0.504	0.442	0.400	0.370	0.326	0.310	0.283

and the distribution of the firing rate  $r$ ,  $P(r(j))$ , are given. Let us make a summation of the probability of the stimuli,  $P(s(1)), P(s(2)), \dots, P(s(M))$ , one by one until we get a number  $d_1$  ( $d_1 > 1$ ) such that  $\sum_{i=1}^{d_1} P(s(i)) \leq P(r(1))$  but  $\sum_{i=1}^{d_1+1} P(s(i)) > P(r(1))$ . Then  $r(1) = g(s(d_1))$ ; similarly, we get  $d_2$  ( $d_2 > d_1$ ) such that  $\sum_{i=d_1+1}^{d_2} P(s(i)) \leq P(r(2))$  but  $\sum_{i=d_1+1}^{d_2+1} P(s(i)) > P(r(2))$ . Then  $r(2) = g(s(d_2))$ ; repeat this process and we obtain numerically the tuning curve  $r(j) = g(s(d_j))$ ,  $j = 1, 2, \dots, M$ .

Using the method discussed in the previous paragraph, we calculate the tuning curves given different stimuli distributions and different firing rate distributions. Suppose the stimuli obey normal distribution, and the firing rate distributions are taken to be the two curves in Figure 12, we get corresponding two tuning curves in Figure 13. We can see from Figures 12, 13 that if energy consumption is not important compared to the mutual information (see the red curve in Figure 12, which is the optimum firing rate distribution with  $c = 0.6$  in the definition of information efficiency), the tuning curve takes a sigmoid form (see red curve in Figure 13), which has been widely used in computational neuroscience.



**FIGURE 12 | Firing rate distributions used to calculate the neuronal tuning curves.**



**FIGURE 13 | The neuronal tuning curves corresponding to the firing rate distributions in Figure 12 and normal stimuli distribution.**

In the case where energy consumption is underlined (see the blue curve in Figure 12 which is the optimum firing rate distribution for the information efficiency with  $c = 1.2$ ), the tuning curve (the blue one in Figure 13) take a form that it is below the one (the red one in Figure 13) that corresponding to the case when energy consumption matters little, and the top part of the sigmoid function is cut off.

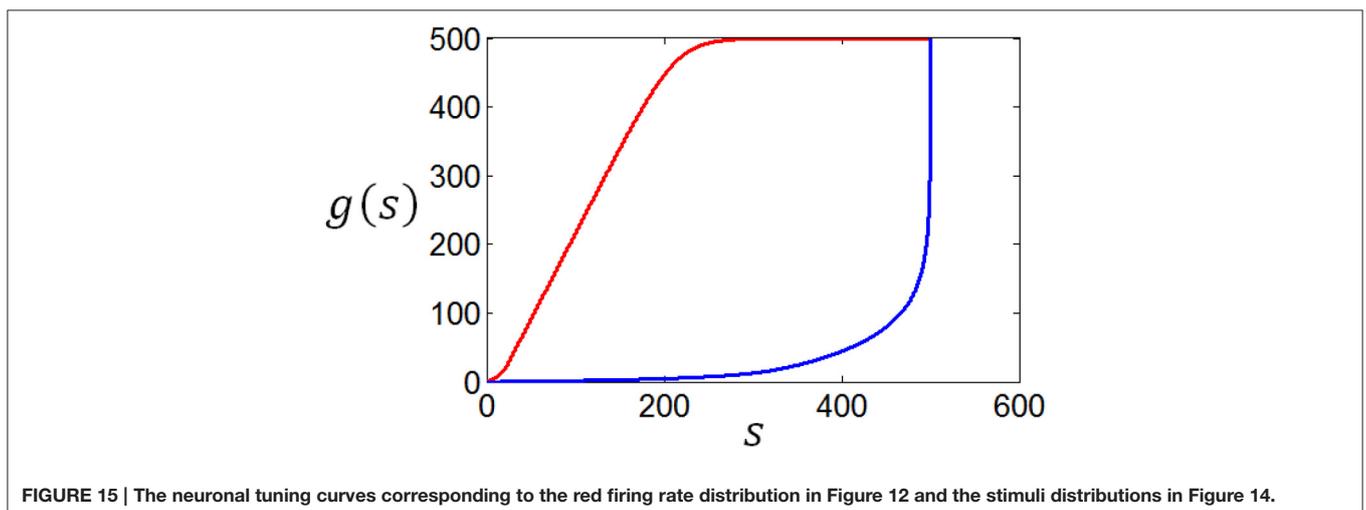
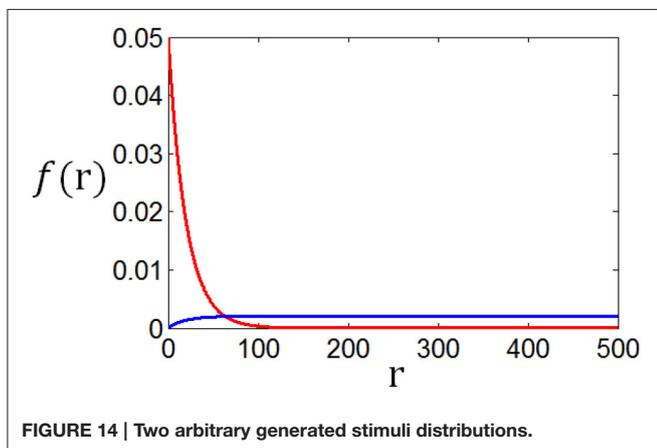
As a matter of fact, the tuning curves can be calculated for any stimuli distributions. Figure 14 shows two arbitrary curves of stimuli distribution, and the two curves in Figure 15 are the corresponding tuning curves.

## CONCLUSIONS

Neural systems are assumed to be optimized through biological evolution for millions of years for information processing, i.e., they are optimized to achieve highest information efficiency. Using this assumption, one can study many features of neuronal

systems. The input–output relationships (tuning curves),  $r = g(s)$ , have been explored in this paper. As one can calculate the tuning curve numerically if optimum firing rate distribution for information efficiency and the stimuli distribution are given, we focused on exploring the optimum firing rate distribution for the information efficiency in this paper. Firstly, a new definition for the information efficiency,  $I_E = (2^{I_m} - 1)/(E_s + E_b)^c$ , has been given. The energy consumption consists of two components, the basic energy consumption  $E_b$  and the spike emission energy consumption  $E_s$ . The relative importance of the energy consumption in the definition of the information efficiency is also modeled by the parameter  $c$ . Then, four main results concerning the optimum firing rate distribution have been obtained. (1) Contrast to the fact that the spike count response distribution should be flat for the maximum full entropy, the function of the spike count response distribution should decrease rapidly to 0 for the minimum spike emission energy consumption ( $E_s$ ). Therefore, the function of the optimum spike count response distribution for the entropy efficiency should exhibit a shape that decreases gradually to 0. This kind of functions can be described by a combination of the exponential functions. (2) Using the

nonlinear least square method, we found that the optimum firing rate distribution function for the entropy efficiency has the same form of function as the optimum spike count response distribution. In other words, the function of the optimum firing rate distribution for the entropy efficiency can also be described by the combination of the exponential functions. Furthermore, we found that the dependency of the noise entropy on the shape of the firing rate distribution function is similar to that of the spike emission energy consumption, i.e., the function of the firing rate distribution should decrease rapidly to 0 to achieve the minimum noise entropy. Therefore, it can be concluded that a firing rate distribution function capable of producing high full entropy, low noise entropy and less energy consumption should have a shape decreasing gradually to 0, which can be described by a combination of exponential functions. (3) We developed a rapid algorithm with variable step size and forward-backward scheme to search the parameter values of the optimum firing rate distribution function. It has been found by this algorithm that firing rate distribution functions described by two free parameters,  $f(r) = (\alpha e^{-\alpha r} + \beta e^{-\beta r})/2$ , are accurate enough to describe the optimum firing rate distribution. Due to the small number of free parameters, they are of light computational complexity for the search for the optimum parameter values. (4) The dependency of the optimum firing rate distribution functions on the parameters of the information efficiency has been explored. It has been found that if exponential index  $c$  is decreased (the energy consumption is relatively neglected), the optimum firing rate distribution will become relatively flat. And if the basic energy consumption of neurons ( $E_b$ ) is increased, the optimum firing rate distribution will also become relatively flat. On the other hand, if exponential index  $c$  is increased (the energy consumption is relatively underlined) or  $E_b$  is decreased, the optimum firing rate distribution will become relatively steep. Finally, we designed an algorithm to calculate the tuning curves when the firing rate distribution and an arbitrary stimulus distribution are provided. By this algorithm, we found that in the case of normal stimuli distribution, the tuning curve exhibits a form of sigmoid if energy consumption is relatively



neglected or the neuronal basic energy consumption is relatively large.

## AUTHOR CONTRIBUTIONS

All authors listed, have made substantial, direct and intellectual contribution to the work, and approved it for publication.

## REFERENCES

- Atick, J. J., and Redlich, A. N. (1990). Towards a theory of early visual processing. *Neural Comput.* 2, 308–320. doi: 10.1162/neco.1990.2.3.308
- Attwell, D., and Laughlin, S. B. (2001). An energy budget for signaling in the grey matter of the brain. *J. Cereb. Blood Flow Metab.* 21, 1133–1145. doi: 10.1097/00004647-200110000-00001
- Bair, W., Koch, C., Newsome, W., and Britten, K. (1994). Power spectrum analysis of bursting cells in area MT in the behaving monkey. *J. Neurosci.* 14, 2870–2892.
- Barlow, H. B. (1959). “Sensory mechanisms, the reduction of redundancy, and intelligence,” in *NPL Symposium on the Mechanization of Thought Process*, Vol. 10 (London: HM Stationery Office), 535–539.
- Barlow, H. B. (1961). “Possible principles underlying the transformation of sensory messages,” in *Sensory Communication*, ed W. A. Rosenblith (Cambridge, MA: MIT Press), 217–234.
- Berger, T., and Levy, W. B. (2010). A mathematical theory of energy efficient neural computation and communication. *IEEE Trans. Inf. Theory* 2, 852–874. doi: 10.1109/TIT.2009.2037089
- Bethge, M., Rotermund, D., and Pawelzik, K. (2003). Second order phase transition in neural rate coding: binary encoding is optimal for rapid signal transmission. *Phys. Rev. Lett.* 90:088104. doi: 10.1103/PhysRevLett.90.088104
- Borst, A., and Theunissen, F. E. (1999). Information theory and neural coding. *Nat. Neurosci.* 2, 947–957. doi: 10.1038/14731
- Day, M. L., Koka, K., and Delgutte, B. (2012). Neural encoding of sound source location in the presence of a concurrent, spatially separated source. *J. Neurophysiol.* 108, 2612–2628. doi: 10.1152/jn.00303.2012
- Dayan, P., and Abbott, L. (2001). *Theoretical Neuroscience-Computational and Mathematical Modeling of Neural Systems*. Cambridge, MA: MIT Press.
- Erecinska, M., Cherian, S., and Silver, I. A. (2004). Energy metabolism in mammalian brain during development. *Progr. Neurobiol.* 73, 397–445. doi: 10.1016/j.pneurobio.2004.06.003
- Gao, X., Grayden, D. B., and McDonnell, M. D. (2014). Stochastic information transfer from cochlear implant electrodes to auditory nerve fibers. *Phys. Rev. E* 90:022722. doi: 10.1103/PhysRevE.90.022722
- Goris, R. L. T., Movshon, J. A., and Simoncelli, E. P. (2014). Partitioning neuronal variability. *Nat. Neurosci.* 6, 858–865. doi: 10.1038/nn.3711
- Han, F., Wang, Z., Fan, H., and Sun, X. (2015). Optimum neural tuning curves for information efficiency with rate coding and finite-time window. *Front. Comput. Neurosci.* 9:67. doi: 10.3389/fncom.2015.00067
- Kostal, L. (2010). Information capacity in the weak-signal approximation. *Phys. Rev. E* 82:026115. doi: 10.1103/PhysRevE.82.026115
- Kostal, L., and Lansky, P. (2013). Information capacity and its approximations under metabolic cost in a simple homogeneous population of neurons. *BioSystems* 112, 265–275. doi: 10.1016/j.biosystems.2013.03.019
- Laughlin, S. (1981). A simple coding procedure enhances a neuron’s information capacity. *Z. Naturforsch.* 36C, 910–912.
- Levy, W. B., and Baxter, R. A. (1996). Energy efficient neural codes. *Neural Comput.* 8, 531–543. doi: 10.1162/neco.1996.8.3.531
- McDonnell, M. D., Ikeda, S., and Mantou, J. (2011). An introductory review of information theory in the context of computational neuroscience. *Biol. Cybern.* 105, 55–70. doi: 10.1007/s00422-011-0451-9
- Moujahid, A., d’Anjou, A., and Torrealdea, F. (2011). Energy and information in Hodgkin-Huxley neurons. *Phys. Rev. E* 83:031912. doi: 10.1103/PhysRevE.83.031912
- McDonnell, M. D., and Stocks, N. G. (2008). Maximally informative stimuli and tuning curves for sigmoidal rate-coding neurons and populations. *Phys. Rev. Lett.* 101:058103. doi: 10.1103/PhysRevLett.101.058103
- Nikitin, A., Stocks, N., Morse, R., and McDonnell, M. (2009). Neural population coding is optimized by discrete tuning curves. *Phys. Rev. Lett.* 103:138101. doi: 10.1103/PhysRevLett.103.138101
- Rolls, E., and Treves, A. (2011). The neuronal encoding of information in the brain. *Progr. Neurobiol.* 95, 448–490. doi: 10.1016/j.pneurobio.2011.08.002
- Sengupta, B., Laughlin, S. B., and Niven, J. E. (2014). Consequences of converting graded to action potentials upon neural information coding and energy efficiency. *PLoS Comput. Biol.* 10:e1003439. doi: 10.1371/journal.pcbi.1003439
- Sengupta, B., and Stemmler, M. B. (2014). Power consumption during neuronal computation. *Proc. IEEE* 102, 738–750. doi: 10.1109/JPROC.2014.2307755
- Torrealdea Francisco, J., Sarasola, C., and D’Anjou, A. (2009). Energy consumption and information transmission in model neurons. *Chaos Solitons Fractals* 40, 60–68. doi: 10.1016/j.chaos.2007.07.050
- Wang, R., and Zhang, Z. (2007). Energy coding in biological neural networks. *Cogn. Neurodyn.* 1, 203–212. doi: 10.1007/s11571-007-9015-z
- Wang, Z., Stocker, A. A., and Lee, D. D. (2013). “Optimal neural tuning curves for arbitrary stimulus distributions: discrimax, infomax and minimum Lp loss,” in *Part of: Advances in Neural Information Processing Systems 25, NIPS 2012* (Nevada), Vol. 25, 2177–2185.
- Wei, X. X., and Stocker, A. A. (2016). Mutual information, Fisher information, and efficient coding. *Neural Comput.* 28, 305–326. doi: 10.1162/NECO\_a\_00804
- Yaeli, S., and Meir, R. (2010). Error-based analysis of optimal tuning functions explains phenomena observed in sensory neurons. *Front. Comput. Neurosci.* 4:130. doi: 10.3389/fncom.2010.00130
- Yarrow, S., and Series, P. (2015). The influence of population size, noise strength and behavioral task on best-encoded stimulus for neurons with unimodal or monotonic tuning curves. *Front. Comput. Neurosci.* 9:18. doi: 10.3389/fncom.2015.00018
- Yu, L., and Liu, L. (2014). Optimal size of stochastic Hodgkin-Huxley neuronal systems for maximal energy efficiency in coding pulse signals. *Phys. Rev. E* 89:032725. doi: 10.1103/PhysRevE.89.032725

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Han, Wang and Fan. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.