Biological Cybernetics (2018) 112:523–538
https://doi.org/10.1007/s00422-018-0779-5

Biological Cybernetics

ORIGINAL ARTICLE

Improved lower bound for the mutual information between signal and neural spike count

Sergej O. Voronenko[1,2] · **Benjamin Lindner**[1,2]

Received: 24 March 2018 / Accepted: 20 August 2018 / Published online: 28 August 2018

Abstract
The mutual information between a stimulus signal and the spike count of a stochastic neuron is in many cases difficult to determine. Therefore, it is often approximated by a lower bound formula that involves linear correlations between input and output only. Here, we improve the linear lower bound for the mutual information by incorporating nonlinear correlations. For the special case of a Gaussian output variable with nonlinear signal dependencies of mean and variance we also derive an exact integral formula for the full mutual information. In our numerical analysis, we first compare the linear and nonlinear lower bounds and the exact integral formula for two different Gaussian models and show under which conditions the nonlinear lower bound provides a significant improvement to the linear approximation. We then inspect two neuron models, the leaky integrate-and-fire model with white Gaussian noise and the Na–K model with channel noise. We show that for certain firing regimes and for intermediate signal strengths the nonlinear lower bound can provide a substantial improvement compared to the linear lower bound. Our results demonstrate the importance of nonlinear input–output correlations for neural information transmission and provide a simple nonlinear approximation for the mutual information that can be applied to more complicated neuron models as well as to experimental data.

Keywords Sensory coding · Signal transmission · Information theory · Mutual information · Nonlinear lower bound · Nonlinear correlations

1 Introduction

One of the key features of neurons and neural systems is their property to transmit and process sensory signals, and it is therefore necessary to find means to reliably quantify the quality of signal transmission. Information theory provides a tool for the quantification of the quality of signal transmission without having to rely on specific decoding or encoding models (Rieke et al. 1996; Borst and Theunissen 1999). The key characteristics of information theory, the mutual information, is shaped not only by linear but also by nonlinear statistical dependencies between an input signal and an output signal. The mutual information and approximations of it have been widely applied in the research area of neural signal processing (see, e.g., Rieke et al. 1995; Chacron et al. 2003; Passaglia and Troy 2004; Sadeghi et al. 2007; Neiman and Russell 2011; Doose et al. 2016).

Some features that have been investigated with the help of the mutual information are for example the beneficial effects of noise on signal transmission (for specific examples, see, e.g., Droste et al. 2013; Voronenko et al. 2015, for general reviews, see Gammaitoni et al. 1998; McDonnell and Ward 2011), the maximal amount of information that the neural spike trains can carry about the stimulus (Strong et al. 1998; Stemmler and Koch 1999; Juusola and French 1997) and the ability of neurons to transmit more information about naturalistic sensory stimuli than about artificial stimuli (Rieke et al. 1995; Nemenman et al. 2008). The drawback of the mutual information is that it is often difficult to estimate. In particular, the neural responses to many repetitions of a stimulus have to be recorded and the neural system is required to remain in a stationary state. In biological experiments these conditions make it often difficult to obtain the necessary amount of data in order to obtain accurate and unbiased

Communicated by Peter J. Thomas.

✉ Sergej O. Voronenko
sergej@physik.hu-berlin.de

[1] Bernstein Center for Computational Neuroscience Berlin, Philippstr. 13, Haus 2, 10115 Berlin, Germany

[2] Physics Department, Humboldt University Berlin, Newtonstr. 15, 12489 Berlin, Germany

estimates of the mutual information (Strong et al. 1998; Borst and Theunissen 1999). Although a lot of improved estimates have been developed (see, e.g., Panzeri and Schultz 2001; Victor 2002; Kraskov et al. 2004; Victor 2006; Panzeri et al. 2007), the faithful estimation of the mutual information still remains difficult, in particular for large neuronal ensembles and complex stimuli.

One possible solution to the problems pertaining the direct estimation is to use a linear lower bound as an approximation for the mutual information (Bialek et al. 1991; Gabbiani 1996). In cases, where information can be decoded by a linear filter from the output, the linear lower bound has proven to be a very helpful tool and some examples have been reported for which the linear lower bound gives an accurate approximation of the information rate (Rieke et al. 1993; Borst and Theunissen 1999). However, intuitively one can expect that for an inherently nonlinear system as a neuron or a neural network it may not always be possible to linearly decode all of the available information (Theunissen and Miller 1995). Indeed, there are also examples of neural systems for which the linear approximation fails (Neiman and Russell 2011; Aldworth et al. 2011; Bernardi and Lindner 2015) and a detailed comparison has revealed that in many cases we may miss half of the information rate if the linear lower bound is used (Aldworth et al. 2011).

The linear lower bound is based on the linear correlations between input and output signals and their autocorrelations. In this paper we investigate how nonlinear correlation coefficients can be used to improve the estimation of the mutual information and come up with a *nonlinear lower bound* for the mutual information between stimulus and response. For simplicity, we consider the case of a static Gaussian input signal but hope that something can be learned from our exercise about how to treat the more broadly relevant case of a time-dependent input signal (a static input signal can be viewed as the limiting case of a very slow input signal which is encoded in the firing rate).

The formula that we derive is simple and is straightforwardly applicable to neural data. We consider a few models and different parameter regimes to facilitate the intuition about when the nonlinear lower bound can improve the linear approximation of the information transfer. In particular, we compare our nonlinear lower bound formula to the mutual information for a simple Gaussian model and two different neuron models (for the latter models, we also compare to an alternative approximation by Brunel and Nadal 1998). We find that for signals of intermediate strength the nonlinear estimate is close to the mutual information and performs better than the linear approximation. Some further results which are also derived in this paper are an integral formula for the mutual information between a Gaussian input signal and a Gaussian output variable and an upper bound for the mutual information which is valid in case of a Gaussian output noise.

Our paper is organized as follows. In Sect. 2, we recall some basic concepts of information theory and introduce the models to be inspected. In subsequent Sect. 3 we derive our main theoretical results. The latter are applied to two simple Gaussian models in Sect. 4 and to two neural models in Sect. 5. We conclude in Sect. 6 with a discussion of our theoretical and numerical results.

2 Measures and models

2.1 Mutual information

One established measure of neural information transmission that does not require undue assumptions about encoding or decoding mechanisms is the mutual information (Rieke et al. 1996). In this study we consider a noisy neuron which is subject to a repeated stimulation by a static signal and for which the observation time of the output is long enough such that transients can be ignored (stationary setup). By a static signal we mean a signal which is constant over the time course of a trial but assumes different values for different trials. Because in experiments researchers often employ Gaussian signals, we will here also choose a Gaussian distribution of the signal values on different trials. Because the signal is chosen to be static, the exact timing of the neural spikes will not play a role and only the spike count N within a time window T will carry information about the signal s.

In order to determine the information that the output variable N carries about the input variable s, we first have to define the total entropy of the output. This entropy is given by

$$H_N = -\sum_i P(N_i) \log_2 \left(P(N_i)\right) \ , \tag{1}$$

where $P(N_i)$ is the probability that the variable N will attain the state N_i and where the sum runs over all possible states. For simplicity, let us first consider the signal s in a discretized version, i.e., s attains different states s_j. Picking out only realizations for which the signal s was in the state s_j yields the conditional entropy of N

$$H_{N|s_j} = -\sum_i P(N_i|s_j) \log_2 \left(P(N_i|s_j)\right) \ . \tag{2}$$

The mutual information between the variables s and N is given by the difference of entropies and can be expressed in a symmetric fashion by the probabilities of input and output

$$\begin{aligned} \mathcal{M} &= H_N - \left\langle H_{N|s} \right\rangle_s \\ &= \sum_{i,j} P(N_i, s_j) \log_2 \left(\frac{P(N_i, s_j)}{P(N_i)P(s_j)} \right) \ . \end{aligned} \tag{3}$$

Here $\langle . \rangle_s$ indicates an ensemble average over different realizations of the signal. From Eq. (3) it can be seen that the mutual information is a nonlinear measure of statistical dependencies between s and N. For continuously distributed variables s and x (e.g., a Gaussian input signal and a Gaussian output variable) the sums in Eq. (3) have to be replaced by integrals

$$\mathcal{M} = \int\int \mathrm{d}s\,\mathrm{d}x \;\; p(x,s) \log_2 \left(\frac{p(x,s)}{p(x)p(s)} \right) . \tag{4}$$

In this case, the mutual information corresponds to the difference of *differential entropies*, i.e., Eqs. (1) and (2) in integral form (for a discussion of the properties of this kind of entropy see Cover and Thomas 1991).

2.2 Correlation coefficients

In our theory we will employ various correlation coefficients that for general stochastic variables a and b are defined by

$$\rho_{a,b}^2 = \frac{\langle (a - \langle a \rangle)\,(b - \langle b \rangle) \rangle^2}{\sigma_a^2 \sigma_b^2} , \tag{5}$$

where $\sigma_a^2 = \left\langle (a - \langle a \rangle)^2 \right\rangle$ and $\sigma_b^2 = \left\langle (b - \langle b \rangle)^2 \right\rangle$ are the variances. If a and b are linear functions of other stochastic variables (e.g., the input and output of a stochastic system) then $\rho_{a,b}$ is referred to as a linear correlation coefficient. However, if a and b are nonlinear functions of other stochastic variables, then $\rho_{a,b}$ is referred to as a nonlinear correlation coefficient.

2.3 Leaky integrate-and-fire neuron model

The popular stochastic leaky integrate-and-fire neuron model (see the comprehensive review by Burkitt 2006) is a point-neuron model in which the subthreshold voltage across the nerve membrane is determined by the following stochastic differential equation

$$\tau_m \dot{v} = -v + \mu + s + \xi(t) . \tag{6}$$

Here τ_m is the membrane time constant, μ is a constant input, s is a static input signal, and ξ is a white Gaussian noise with zero mean and the autocorrelation $\left\langle \xi(t)\xi(t+t') \right\rangle = 2D\delta(t')$; both μ and s are electric input currents but rescaled by the input resistance and thus given in units of a voltage. The stochastic equation for the subthreshold voltage, Eq. (6), is complemented by a fire-and-reset rule: Whenever the voltage hits the threshold, $v_\mathrm{T} = 1$, a spike is registered and the voltage is set to the value $v_\mathrm{R} = 0$ for a refractory time τ_ref after which it continues to evolve according to Eq. (6). In this paper, we will consider two different mean currents μ, corresponding to two different firing regimes of the neuron. In the suprathreshold firing regime ($\mu = 1.1$) the constant current leads to sustained firing even in the absence of noise (Vilela and Lindner 2009b). In the subthreshold firing regime ($\mu = 0.9$) there is no firing in the absence of noise and only fluctuations can lift the voltage over the threshold and induce firing. Throughout the paper we will use the noise value $D = 0.001$.

2.4 Izhikevich's Na–K model with channel noise

We also perform numerical simulations of Izhikevich's persistent-sodium-plus-potassium model (or Na–K model for short), which was originally proposed in Izhikevich (2007), is similar to the two-dimensional Morris–Lecar model (Morris and Lecar 1981), and has been studied in a stochastic version with discrete ion channel noise in Thomas and Lindner (2014) and Voronenko and Lindner (2017). The dynamics of the Na–K model is given by

$$C\dot{v} = I_0 + s - I_\mathrm{L}(v) - I_\mathrm{Na}(v) - I_\mathrm{K}(v,n) , \tag{7}$$

with a constant input current I_0, a static signal s, a passive leak current I_L and a deterministic "persistent sodium" current I_Na. The potassium-gated current I_K in Eq. (7) is gated by the number of open potassium channels $0 \le n(t) \le n_\mathrm{tot}$, and the dynamics of the number of open channels $n(t)$ is governed by the transition rates $R_o(v)$ for channel opening and $R_\mathrm{c}(v)$ for channel closing. The Na–K model exhibits excitability and generates spikes without an additional fire-and-reset rule as is required for the LIF.

Because the potassium channels open and close according to transition rates which are voltage dependent, the resulting noise for the voltage dynamics is effectively a multiplicative and temporally correlated noise in contrast to the white (uncorrelated) noise for the LIF. In this paper we consider a potassium current with a high threshold and standard parameters (Izhikevich 2007; Thomas and Lindner 2014), for which the passive leak current is given by $I_\mathrm{L} = g_\mathrm{L}(v - E_\mathrm{L})$ with $g_\mathrm{L} = 8\,\mathrm{mS/cm^2}$ and $E_\mathrm{L} = -80\,\mathrm{mV}$, the "persistent sodium" current is defined as $I_{\mathrm{Na},p} = g_\mathrm{Na} m_\infty(v)(v - E_\mathrm{Na})$ with $g_\mathrm{Na} = 20\,\mathrm{mS/cm^2}$, $E_\mathrm{Na} = 60\,\mathrm{mV}$, and a voltage-dependent activation $m_\infty(v) = 1/\{1 + \exp[(-20\,\mathrm{mV} - v)/15\,\mathrm{mV}]\}$; the potassium current is $I_\mathrm{K}(t) = g_k(n(t)/n_\mathrm{tot})(v - E_\mathrm{K})$ with $g_\mathrm{K} = 10\,\mathrm{mS/cm^2}$, $E_\mathrm{K} = -90\,\mathrm{mV}$, open channel number $0 \le n(t) \le n_\mathrm{tot}$ and the total number of channels $n_\mathrm{tot} = 100$. The per capita transition rate for channel opening is $R_o(v) = 1/\{1 + \exp[(-25\,\mathrm{mV} - v)/5\,\mathrm{mV}]\}$, and for channel closing, it is $R_\mathrm{c}(v) = 1 - R_o(v)$. The membrane capacitance is $C = 1\,\mu\mathrm{F/cm^2}$, and the constant input current is $I_0 = 6\,\mu\mathrm{A/cm^2}$.

3 Theory

3.1 Nonlinear lower bound: Gaussian input

In the following we introduce a quadratic reconstruction of a static Gaussian input variable s from a static output variable x. For this quadratic reconstruction we determine the optimal coefficients, for which the reconstruction extracts the maximally possible information from the output. We then derive the nonlinear lower bound, which corresponds to the minimal amount of information that is retrieved from the output by the introduced quadratic reconstruction. Furthermore, we demonstrate that the nonlinear lower bound is always smaller than or equal to the full mutual information and is always larger than or equal to the linear lower bound. The derivation which we present here is an extension of the derivation for the linear lower bound in Gabbiani (1996).

We first define the quadratic reconstruction, which takes into account second-order contributions from the output

$$s_{\mathrm{rec}} = h \cdot N + g \cdot N^2 \ . \tag{8}$$

The reconstruction error is given by the difference between the reconstructed and the original signal, $\eta = s - s_{\mathrm{rec}}$. The optimal coefficients h_{opt} and g_{opt} are determined by minimizing σ_η^2, the variance of the reconstruction error

$$\begin{aligned}\sigma_\eta^2 &= \left\langle \eta^2 \right\rangle - \langle \eta \rangle^2 \\ &= \sigma_s^2 + h^2\sigma_N^2 + g^2\sigma_{N^2}^2 + 2hg\left\langle \left(N - \langle N\rangle\right)\left(N^2 - \left\langle N^2\right\rangle\right)\right\rangle \\ &\quad -2h\left\langle \left(s - \langle s\rangle\right)\left(N - \langle N\rangle\right)\right\rangle - 2g\left\langle \left(s - \langle s\rangle\right)\left(N^2 - \left\langle N^2\right\rangle\right)\right\rangle .\end{aligned} \tag{9}$$

The coefficients h_{opt} and g_{opt} are determined by setting the partial derivatives of Eq. (9) with respect to h and g to zero; using the definition of ρ from Eq. (5), this yields

$$h_{\mathrm{opt}} = \frac{\sigma_s}{\sigma_N} \frac{1}{1-\rho_{N,N^2}^2} \left(\rho_{s,N} - \rho_{s,N^2}\rho_{N,N^2}\right) \ , \tag{10}$$

$$g_{\mathrm{opt}} = \frac{\sigma_s}{\sigma_{N^2}} \frac{1}{1-\rho_{N,N^2}^2} \left(\rho_{s,N^2} - \rho_{s,N}\rho_{N,N^2}\right) \ . \tag{11}$$

The minimal reconstruction error achieved in this way reads

$$\sigma_{\eta,nl}^2 = \sigma_s^2 - \sigma_s^2\rho_{s,N}^2 - \sigma_s^2 \frac{\left(\rho_{s,N^2} - \rho_{s,N}\rho_{N,N^2}\right)^2}{1-\rho_{N,N^2}^2} \ . \tag{12}$$

For a linear signal reconstruction $s_{\mathrm{rec}}^{\mathrm{lin}} = h_{\mathrm{lin}} N$ the minimal reconstruction error results in

$$\sigma_{\eta,\mathrm{lin}}^2 = \sigma_s^2 - \sigma_s^2\rho_{s,N}^2 \ . \tag{13}$$

Now, we determine the information that the quadratic reconstruction can recover from the output about the stimulus. This information is given by the mutual information between the input signal and the reconstructed signal, $\mathcal{M}_{s,s_{\mathrm{rec}}}$, and will in general be smaller than the mutual information between the input signal and the output N according to the *data processing inequality* (Cover and Thomas 1991)

$$\mathcal{M}_{s,s_{\mathrm{rec}}} \leq \mathcal{M}_{N,s} \ . \tag{14}$$

The mutual information between s and s_{rec} can be rewritten in terms of entropies as

$$\mathcal{M}_{s,s_{\mathrm{rec}}} = H_s - \left\langle H_{s|s_{\mathrm{rec}}} \right\rangle_{s_{\mathrm{rec}}} \ , \tag{15}$$

where H_s is the signal entropy and $\left\langle H_{s|s_{\mathrm{rec}}} \right\rangle$ is the noise entropy. For a fixed s_{rec}, the uncertainty about the input s is represented by the reconstruction noise η and we find

$$H_{s|s_{\mathrm{rec}}} = H_{\eta|s_{\mathrm{rec}}} \ . \tag{16}$$

The statistics of the reconstruction noise $\eta(s_{\mathrm{rec}})$ will in general depend on the specific realization of s_{rec} and may not be Gaussian. However, the fact that for fixed variance $\sigma_\eta^2(s_{\mathrm{rec}})$ the entropy is maximized by a Gaussian random variable η_G (Shannon 1948; Rieke et al. 1996) implies that

$$H_{\eta|s_{\mathrm{rec}}} \leq H_{\eta_G|s_{\mathrm{rec}}} \ . \tag{17}$$

Taking further into account that an averaged conditional entropy is always smaller than the unconditioned entropy (Rieke et al. 1996) leads to

$$\left\langle H_{\eta_G|s_{\mathrm{rec}}} \right\rangle_{s_{\mathrm{rec}}} \leq H_{\eta_G} \ . \tag{18}$$

In conclusion, using Eqs. (14–18) we obtain the lower bound for the mutual information between s and N ($\mathcal{M}_{lb} \leq \mathcal{M}_{s,N}$) to be

$$\mathcal{M}_{lb} = H_s - H_{\eta_G} = -\frac{1}{2}\log_2\left(\frac{\sigma_{\eta,nl}^2}{\sigma_s^2}\right) \ , \tag{19}$$

where H_s and H_{η_G} are both entropies of the Gaussian variables s and η_G with variances σ_s^2 and $\sigma_{\eta,nl}^2$, respectively. Combining Eqs. (12) and (19) we arrive at the final result for the nonlinear lower bound

$$\mathcal{M}_{lb}^{nl} = -\frac{1}{2}\log_2\left(1 - \rho_{N,s}^2 - \frac{\left(\rho_{s,N^2} - \rho_{s,N}\rho_{N,N^2}\right)^2}{1-\rho_{N,N^2}^2}\right) . \tag{20}$$

This may be compared to the linear lower bound, obtained by inserting Eq. (13) into Eq. (19)

$$\mathcal{M}_{lb}^{\text{lin}} = -\frac{1}{2}\log_2\left(1-\rho_{N,s}^2\right) . \tag{21}$$

Because the definition of $\rho_{a,b}$ in Eq. (5) implies that $\rho_{N,N^2}^2 \leq 1$, it holds that

$$\frac{\left(\rho_{s,N^2}-\rho_{s,N}\rho_{N,N^2}\right)^2}{1-\rho_{N,N^2}^2} \geq 0 , \tag{22}$$

which implies that the nonlinear lower bound is always higher than or equal to the linear lower bound

$$\mathcal{M}_{\text{lb}}^{\text{nl}} \geq \mathcal{M}_{lb}^{\text{lin}} = -\frac{1}{2}\log_2\left(1-\rho_{N,s}^2\right) , \tag{23}$$

as can be expected.

3.2 Upper bound: Gaussian input and Gaussian noise

It is possible to derive an upper bound formula for the mutual information in the case in which both the input signal and the intrinsic noise (for a frozen value of the input signal) have Gaussian statistics. This is less general than the case covered in the previous section but still comprises cases in which both the mean and variance of the output depend in a nonlinear fashion on the input signal.

In the following we consider a nonlinear system of the form

$$x = M(s) + \sqrt{V(s)}\,\xi , \tag{24}$$

where $M(s)$ and $V(s)$ are nonlinear functions of s and the noise ξ is a Gaussian noise variable with zero mean and unit variance. For the system in Eq. (24) we derive a generalized upper bound for the mutual information. Our derivation is a generalization of the derivation of the upper bound from Borst and Theunissen (1999), which only holds for constant $V(s) \equiv$const. The generalized upper bound can be used as a check for the error-prone numerical estimate of the full mutual information. Furthermore, whenever one of the lower bounds is close to the upper bound, it is possible to use either one of these measures as an estimate for the mutual information. This could be particularly helpful in experimental settings where the amount of available data is too limited to compute the mutual information directly.

The mutual information between a signal s and the output of the nonlinear Gaussian model, Eq. (24), reads

$$\mathcal{M}_{x,s} = H_x - \left\langle H_{x|s}\right\rangle_s . \tag{25}$$

For a fixed signal, all the variability of x comes from the noise ξ, which is multiplied with $\sqrt{V(s)}$. Hence, we find that

$$p(x|s) = \frac{1}{\sqrt{2\pi V(s)}}\mathrm{e}^{-\frac{(x-M(s))^2}{2V(s)}} , \tag{26}$$

which leads to the differential entropy

$$\left\langle H_{x|s}\right\rangle_s = \frac{1}{2}\left\langle \log_2\left(2\pi e V(s)\right)\right\rangle_s , \tag{27}$$

where e is the Euler constant. In general, the output x will be non-Gaussian, such that the entropy H_x will be smaller than or equal to the entropy H_{x_G} of a Gaussian variable x_G with the same variance as x. The entropy of a Gaussian variable with the density $p_G(x) = \exp[-x^2/(2\sigma_x^2)]/\sqrt{2\pi\sigma_x^2}$ is given by

$$\begin{aligned} H_{x_G} &= -\int \mathrm{d}s\; p(s) \int \mathrm{d}x\; p_G(x)\log_2(p_G(x)) \\ &= \frac{1}{2}\log_2\left(2\pi e\sigma_x^2\right) , \end{aligned} \tag{28}$$

which leads to

$$\mathcal{M}_{x,s} = H_x - \left\langle H_{x|s}\right\rangle_s \leq H_{x_G} - \left\langle H_{x|s}\right\rangle_s = \mathcal{M}_{\text{gub}} , \tag{29}$$

where $\mathcal{M}_{\text{gub}}$ is the generalized upper bound for the mutual information. Inserting Eqs. (27) and (28) into Eq. (29) leads to the final result for the upper bound for the mutual information of a signal processing system with a Gaussian intrinsic noise:

$$\mathcal{M}_{\text{gub}} = \frac{1}{2}\left\langle \log_2\left(\frac{\sigma_x^2}{V(s)}\right)\right\rangle_s \tag{30}$$

The relation in Eq. (29) holds if the output can be written in the form of Eq. (24) and if the noise ξ is Gaussian for a fixed realization of the signal. If the unconditional distribution of x is also Gaussian, then the equality in Eq. (29) holds and the upper bound is equal to the mutual information. However, when the distribution x deviates from the Gaussian distribution, e.g., due to a nonlinearity of $M(s)$ and $V(s)$ in Eq. (24) or due to a non-Gaussian distribution of the signal, then the upper bound overestimates the mutual information. Note that Eq. (29) breaks down if the noise ξ is non-Gaussian and that in this case there is no clear relation between $\mathcal{M}_{\text{gub}}$ and the mutual information. This is a restriction which is not shared by the linear or the nonlinear lower bounds.

4 Mutual information and its bounds for a nonlinear Gaussian model

We consider the Gaussian model, Eq. (24). Then the mutual information between s and x is given by

$$\mathcal{M} = \iint \mathrm{d}s\,\mathrm{d}x\; p_s(s)\,p_x(x|s)\,\log_2\left[\frac{p_x(x|s)}{\int \mathrm{d}s'\; p_s(s')\,p_x(x|s')}\right], \tag{31}$$

where

$$p_s(s) = \frac{1}{\sqrt{2\pi\sigma_s^2}}\mathrm{e}^{-\frac{s^2}{2\sigma_s^2}} \tag{32}$$

and

$$p_x(x|s) = \frac{1}{\sqrt{2\pi V(s)}}\mathrm{e}^{-\frac{(x-M(s))^2}{2V(s)}}\,. \tag{33}$$

Using Eqs. (5) and (24) we can compute the linear and nonlinear correlation coefficients between s and x and find

$$\begin{aligned}
\rho^2_{s,x} &= \frac{\langle s\cdot M(s)\rangle^2}{\sigma_s^2\sigma_x^2}\,,\\
\rho^2_{s,x^2} &= \frac{\left(\langle s\cdot M(s)^2\rangle + \langle s\cdot V(s)\rangle^2\right)^2}{\sigma_s^2\sigma_{x^2}^2}\,,\\
\rho^2_{x,x^2} &= \frac{1}{\sigma_x^2\sigma_{x^2}^2}\Big(\langle M^3(s)\rangle - \langle M(s)\rangle\langle M^2(s)\rangle\\
&\quad + 3\langle M(s)V(s)\rangle - \langle M(s)\rangle\langle V(s)\rangle\Big)^2,
\end{aligned} \tag{34}$$

where the variance of the output is given by

$$\sigma_x^2 = \left\langle M^2(s)\right\rangle - \langle M(s)\rangle^2 + \langle V(s)\rangle \tag{35}$$

and the variance of the squared output reads

$$\begin{aligned}
\sigma_{x^2}^2 &= \left\langle M^4(s)\right\rangle - \left\langle M^2(s)\right\rangle^2 + 6\left\langle M^2(s)V(s)\right\rangle\\
&\quad - \left\langle M^2(s)\right\rangle\langle V(s)\rangle + 3\left\langle V^2(s)\right\rangle - \langle V(s)\rangle^2\,.
\end{aligned} \tag{36}$$

Nonlinear signal dependence of the mean

We now consider a Gaussian model with the functions

$$M(s) = s + \alpha s^2 \quad \text{and} \quad V(s) = 1\,, \tag{37}$$

where α is a constant coefficient. The functions $M(s)$ and $V(s)$ are plotted in Fig. 1a for $\alpha = 0.001$. Note that for the displayed range of signal values the function $M(s)$ only slightly deviates from a linear function (gray-dotted line). The third panel in Fig. 1a shows the Gaussian signal distributions for $\sigma_s = 20, 50, 100$, which were renormalized such that the maximum is always equal to one. The renormalized distribution illustrates which signal values are most likely assumed by the random signal. For example, for $\sigma_s = 20$ 99.7% of the signal realizations fall within the interval $[-60, 60]$ over which the function $M(s)$ is hardly distinguishable from the gray-dashed line. The colored shaded regions indicate signal values which lie within one standard deviation away from the mean (yellow for $\sigma_s = 20$, blue for $\sigma_s = 50$, red for $\sigma_s = 100$). For the correlation coefficients we find

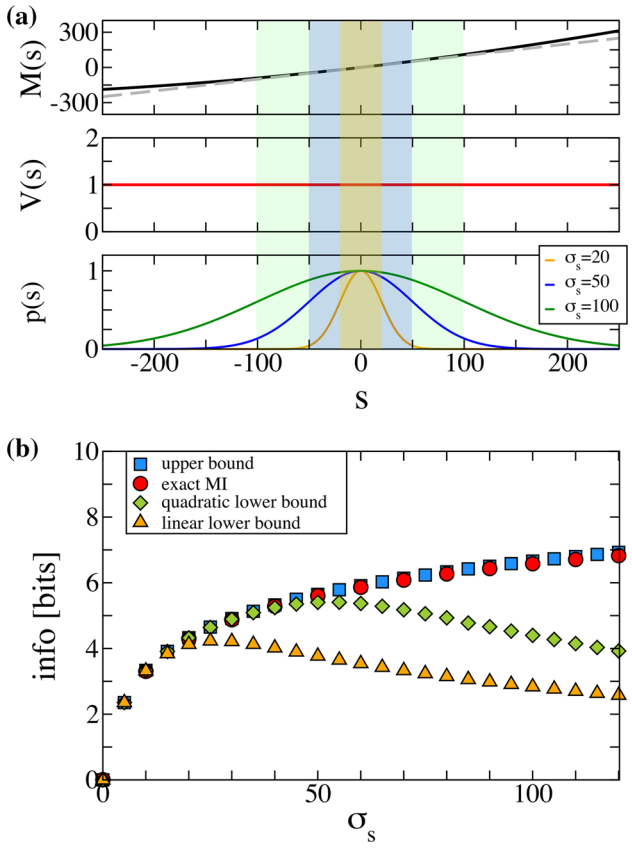

Fig. 1 Gaussian model with a nonlinear signal dependence of the mean. **a** Function $M(s)$ vs signal value; function $V(s)$ vs signal value; and rescaled signal distributions for three different values of σ_s. **b** Mutual information (red circles), linear lower bound (orange triangles), nonlinear lower bound (green diamonds) and upper bound (blue squares) vs σ_s (color figure online)

$$\begin{aligned}
\rho^2_{s,x} &= \frac{\sigma_s^2}{2\alpha\sigma_s^4 + \sigma_s^2 + 1}\,,\\
\rho^2_{s,x^2} &= \frac{18\alpha^2\sigma_s^6}{48\alpha^4\sigma_s^8 + 42\alpha^2\sigma_s^6 + \sigma_s^4(1+6\alpha^2) + 2\sigma_s^2 + 1}\,,
\end{aligned}$$

$$\rho^2_{x,x^2} = \frac{2\alpha^2\sigma_s^4\left(1+4\sigma_s^2+6\alpha^2\sigma_s^4\right)^2}{1+\sigma_s^2+2\alpha\sigma_s^4}\Big[1+\sigma_s^4 + 42\alpha^2\sigma_s^6+48\alpha^4\sigma_s^8+2(\sigma_s^2+3\alpha^2\sigma_s^4)\Big]^{-1}. \quad (38)$$

The nonlinear and linear lower bounds are computed by inserting the nonlinear correlation coefficients from Eq. (38) into Eq. (20) and Eq. (21), respectively. The upper bound for the Gaussian model, Eq. (30), with the functions $M(s)$ and $V(s)$ from Eq. (37) is given by

$$\mathcal{M}_{\text{gub}} = \frac{1}{2}\log_2(2\alpha^2\sigma_s^4+\sigma_s^2+1)\,. \quad (39)$$

The mutual information is computed by numerically evaluating the integral in Eq. (31).

The resulting mutual information and its various bounds are plotted in Fig. 1b. The mutual information (red circles) increases monotonically with increasing σ_s and is matched closely by the upper bound (blue squares). As we discussed in the introduction, the linear lower bound can often be used as a simple estimate for the mutual information for weak signal strengths. Indeed, we find that the linear lower bound (yellow triangles) is a faithful approximation of the mutual information for $\sigma_s \leq 20$. For larger signal strengths, however, the linear lower bound decays with increasing signal strength after reaching a maximum at $\sigma_s \approx 25$. The nonlinear lower bound (green diamonds) closely matches the mutual information for $\sigma_s \leq 40$, a range which is twice as large than for the linear lower bound. For larger signal strengths the nonlinear lower bound also decays with increasing signal strength in a similar way as the linear lower bound, after reaching a maximum at $\sigma_s \approx 55$. The plot also illustrates that the nonlinear lower bound always is equal to or larger than the linear lower bound.

Why do the lower bounds *decrease* at large σ_s with growing signal strength? This asymptotic behavior can be understood by considering the linear and nonlinear correlation coefficients, Eq. (38), in the limit of a strong signal. Because of the nonlinear term in the model, the variance for very strong amplitude grows more strongly than the co-variance between signal and output and thus the corresponding correlation coefficients vanish in the limit:

$$\lim_{\sigma_s\to\infty}\rho^2_{s,x}=0\,,\ \lim_{\sigma_s\to\infty}\rho^2_{s,x^2}=0\,,\ \lim_{\sigma_s\to\infty}\rho^2_{x,x^2}=\frac{3}{4}\,. \quad (40)$$

Because the lower bounds are entirely based on these correlations, their respective limits vanish as well

$$\lim_{\sigma_s\to\infty}\mathcal{M}^{\text{lin}}_{lb}=0 \quad \text{and} \quad \lim_{\sigma_s\to\infty}\mathcal{M}^{nl}_{lb}=0\,. \quad (41)$$

In marked contrast to the lower bounds, the upper bound, Eq. (39), grows for increasing signal strengths. Although we cannot determine the asymptotics of the mutual information analytically, we expect that stronger signals will imply a stronger statistical dependence between x and s resulting in an unlimited growth of the mutual information with increasing σ_s. The simple example illustrates that the lower bounds do not allow general conclusions about the qualitative behavior of the mutual information at arbitrary signal amplitudes.

Despite the strong discrepancy between mutual information and its lower bounds at large signal amplitude, the difference between the two lower bounds for the intermediate amplitude range deserves attention. In our simple example, the nonlinear lower bound provides a simple estimate of the mutual information for a much larger range of signal strengths than the linear lower bound ($0 \leq \sigma_s \leq 40$ vs $0 \leq \sigma_s \leq 20$, respectively). Hence, the nonlinear lower bound constitutes, at least for this simple example, a significant improvement compared to the linear lower bound.

Nonlinear signal dependence of the output variance

Now we consider a Gaussian model for which only the variance exhibits a signal dependence:

$$M(s)=0 \quad \text{and} \quad V(s)=(1+\beta s)^2\,, \quad (42)$$

where β is a constant parameter. The functions $M(s)$ and $V(s)$ with $\beta = 0.01$ are shown in Fig. 2a. The third panel in Fig. 1a shows the renormalized Gaussian signal distributions for $\sigma_s = 20, 50, 100$. For $\sigma_s = 20$ more than 99% of all signal realizations are within the interval $[-60, 60]$ over which the function $V(s)$ is invertible. For $\sigma_s = 100$, however, the values of the signal upon different realizations are drawn from a wider interval, over which the function $V(s)$ is not invertible anymore. For our choice of $M(s)$ and $V(s)$ in Eq. (42), the linear and nonlinear correlation coefficients, Eq. (34), can be calculated and read

$$\rho^2_{s,x}=0\,,\quad \rho^2_{s,x^2}=\frac{2\beta^2\sigma_s^2}{1+8\beta^2\sigma_s^2+4\beta^4\sigma_s^4}\,,\quad \rho^2_{x,x^2}=0\,, \quad (43)$$

which for the linear and nonlinear lower bounds leads to

$$\mathcal{M}^{\text{lin}}_{lb}=0\,, \quad (44)$$

$$\mathcal{M}^{nl}_{lb}=-\frac{1}{2}\log_2\left(1-\frac{2\beta^2\sigma_s^2}{1+8\beta^2\sigma_s^2+4\beta^4\sigma_s^4}\right)\,. \quad (45)$$

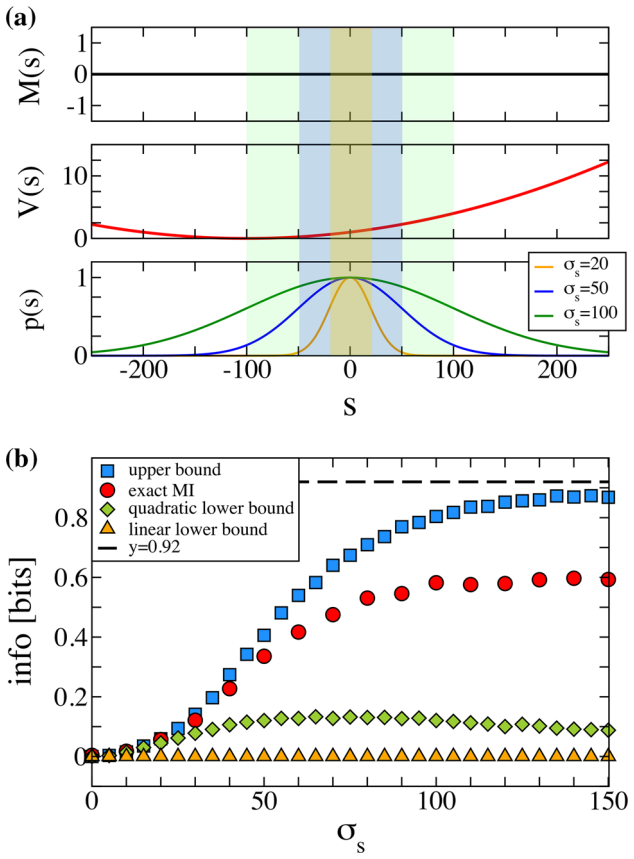

Fig. 2 Gaussian model with a nonlinear signal dependence of the output variance. **a** Function $M(s)$ vs signal value; function $V(s)$ vs signal value; and rescaled signal distributions for three different values of σ_s. **b** Mutual information (red circles), linear lower bound (orange triangles), nonlinear lower bound (green diamonds) and upper bound (blue squares) vs σ_s (color figure online)

The upper bound is given by

$$\mathcal{M}_{\text{gub}} = \frac{1}{2}\left\langle \log_2\left(\frac{1+\beta^2\sigma_s^2}{(1+\beta s)^2}\right)\right\rangle_s = \frac{1}{2}\int_{-\infty}^{\infty} \mathrm{d}s\, \frac{\mathrm{e}^{-\frac{s^2}{2\sigma_s^2}}}{\sqrt{2\pi\sigma_s^2}} \log_2\left(\frac{1+\beta^2\sigma_s^2}{(1+\beta s)^2}\right). \tag{46}$$

and still requires the numerical evaluation of one integral. The mutual information is computed via a numerical integration of Eq. (31).

We can see in Fig. 2b that the mutual information (red circles) increases monotonically for increasing signal strengths. The linear lower bound (orange triangles) is always zero irrespective of the strength of the signal. The nonlinear lower bound (green diamonds) matches closely the mutual information for $\sigma_s \leq 20$, reaches a maximum and then decreases for increasing signal strength. From Eq. (45) we can find that the maximum is attained at $\sigma_s = 1/(\sqrt{2}\cdot|\beta|)$ and that the maximal value of the nonlinear lower bound is given by

$$\mathcal{M}^{nl}_{lb,\max} = -0.5\left(\log_2(5) - \log_2(6)\right) \approx 0.13. \tag{47}$$

Interestingly, the maximal value of the nonlinear lower bound is independent of the choice of the coefficient β. In the limit of large signal strengths, the nonlinear correlation coefficients between s and x^4 are given by $\lim_{\sigma_s\to\infty}\rho^2_{s,x^2} = 0$, which implies that the nonlinear lower bound decays to zero for large signal strengths similar to what we already observed for the Gaussian model with the nonlinear signal dependence of the mean. For small signal strengths, the upper bound (blue squares in Fig. 2b) increases monotonically with increasing σ_s. In order to determine the behavior of the upper bound for large signal strength, we recast the integral expression in Eq. (46) by a simple variable transformation in the following form

$$\mathcal{M}_{\text{gub}} = -\frac{1}{\sqrt{8\pi}}\int_{-\infty}^{\infty} \mathrm{d}s\, \mathrm{e}^{-\frac{s^2}{2}} \log_2\left(\frac{\frac{1}{\sigma_s^2} + \frac{2\beta s}{\sigma_s} + \beta^2 s^2}{\frac{1}{\sigma_s^2} + \beta^2}\right). \tag{48}$$

In the limit $\sigma_s \to \infty$, the above equation can be expressed in terms of the Euler–Mascheroni constant $\gamma \approx 0.58$ (Ryzhik and Gradshtein 1963; Lagarias 2013)

$$\lim_{\sigma_s\to\infty}\mathcal{M}_{\text{gub}} = -\frac{1}{\sqrt{8\pi}}\int_{-\infty}^{\infty} \mathrm{d}s\, \mathrm{e}^{-\frac{s^2}{2}} \log_2\left(s^2\right) = \frac{1}{2}\left(1 + \frac{\gamma}{\ln(2)}\right) \approx 0.92\,. \tag{49}$$

We thus find that for large signal strength the upper bound saturates at a finite value (dashed line in Fig. 2b). This implies that also the mutual information remains finite even in the limit of strong signal.

In conclusion, we find for the Gaussian model with a signal-dependent output variance: (i) The linear lower bound cannot provide a nontrivial approximation for the mutual information because the linear correlation coefficient between input and output is always zero, irrespective of the signal strength. (ii) The nonlinear lower bound provides a good estimate of the mutual information for weak signal strengths ($\sigma_s \leq 20$ in our case). (iii) The mutual information does not grow without bounds but saturates at a finite value.

5 Nonlinear transmission of information about a Gaussian signal by the neural spike count

We now investigate the information transmission for a Gaussian signal for two neuron models, the LIF model and the

Na–K model. In particular, we are interested in the information which the neural spike count carries about a static Gaussian input signal. In addition to the numerical estimate of the mutual information, we also determine the upper bound and the linear and the nonlinear lower bounds for the mutual information and study under which conditions these bounds can be used as approximations for the mutual information. We consider a Gaussian input signal with zero mean and standard deviation σ_s.

5.1 Gaussian model for the LIF spike count for large *T*

The LIF model if driven by a white Gaussian noise and a constant input current, generates a renewal point process—the interspike intervals are statistically independent because the reset of the voltage erases any memory that could be carried by the voltage variable and the driving noise is uncorrelated by definition. For large observation times, the statistics of the neural spike count for a renewal process approaches a Gaussian distribution (Cox 1962). This allows us to approximate the neural spike count by a Gaussian variable x_N which we define by the stochastic equation

$$x_N = M_N(s) + \sqrt{V_N(s)}\,\xi\ , \tag{50}$$

where ξ is a Gaussian noise with unit variance. The nonlinear functions $M_N(s)$ and $V_N(s)$ correspond to the mean and variance of the spike count for a fixed signal realization, respectively. Numerically, these functions can be determined by repeatedly stimulating a neuron with a signal s and by computing mean spike count and the spike count variance, respectively. By repeating this procedure for different values of s it is possible to obtain the full function $M_N(s)$. For the LIF model, the functions $M_N(s)$ (blue circles) and $V_N(s)$ (red squares) are shown in Figs. 3a and 4a for the suprathreshold and subthreshold firing regimes, respectively. Note that the only parameter that changes from the suprathreshold regime to the subthreshold regimes is the mean current μ which amounts to a simple shift of the signal's mean value. Consequently, the functions $M_N(s)$ and $V_N(s)$ exhibit identical shapes in the suprathreshold and subthreshold regimes but are shifted along the x-axis (note the different ranges of the x-axis in Figs. 3a and 4a). We find that the function $M_N(s)$ increases monotonically with s. The function $V_N(s)$ first increases with s and then decreases again after attaining a maximum around $s = -0.15$ in the suprathreshold regime and $s = 0.05$ in the subthreshold regime[1].

[1] There is a strong similarity of the observed maximum of the spike count vs mean signal to a standard problem in statistical physics: the giant acceleration of diffusion in a tilted periodic potential (Reimann et al. 2001; Lindner and Sokolov 2016). A Brownian particle in an inclined washboard potential attains a pronounced maximum in its positional variance [quantified by the diffusion coefficient, $D = \lim_{T\to\infty}\langle (x-\langle x\rangle)^2\rangle/(2T)$] for a certain intermediate value of the bias force [equivalent to the mean signal in the LIF model]. The maximum is attained for a force (mean signal strength) close to the bifurcation value that determines the transition from a noise-induced transport (or firing) regime to a regime, in which movement (spiking) is possible without noise.

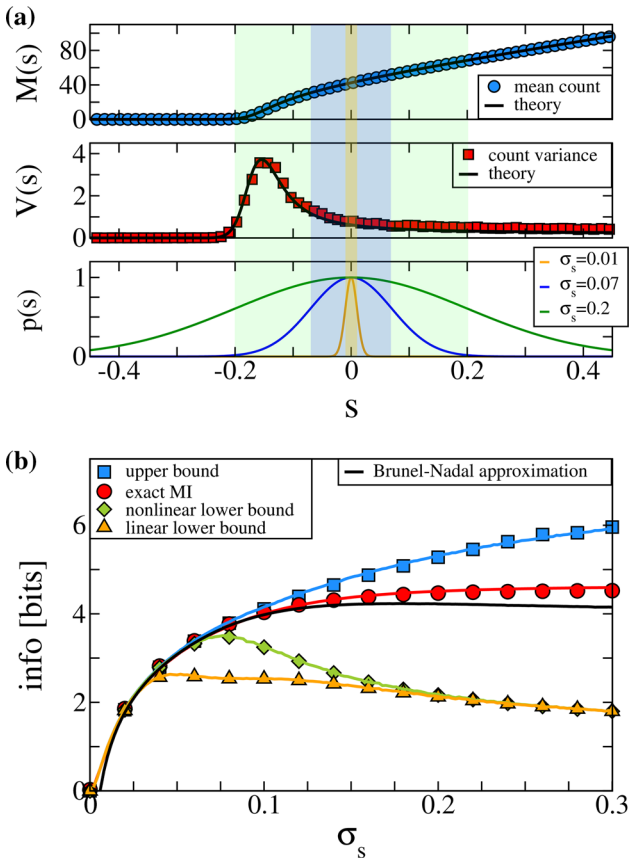

Fig. 3 LIF model with $T = 100$ **and** $\mu = 1.1$. **a** Signal-dependent mean count $M_N(s)$, vs signal value; signal-dependent count variance $V_N(s)$ vs signal value; and rescaled signal distributions for three different values of σ_s. **b** Mutual information (red circles), linear lower bound (orange triangles), nonlinear lower bound (green diamonds) and upper bound (blue squares) vs σ_s. The colored solid lines correspond to the mutual information (red), the linear lower bound (orange), the nonlinear lower bound (green) and the upper bound for the Gaussian model with the same functions $M_N(s)$ and $V_N(s)$ as in **a**. The black solid line shows the Brunel–Nadal approximation for the mutual information (color figure online)

For the LIF model, the function $M_N(s)$ can be computed analytically by employing the fact that the mean spike count can be expressed as a product between the observation time T and the signal-dependent firing rate:

$$M_N(s) = Tr(s)\ . \tag{51}$$

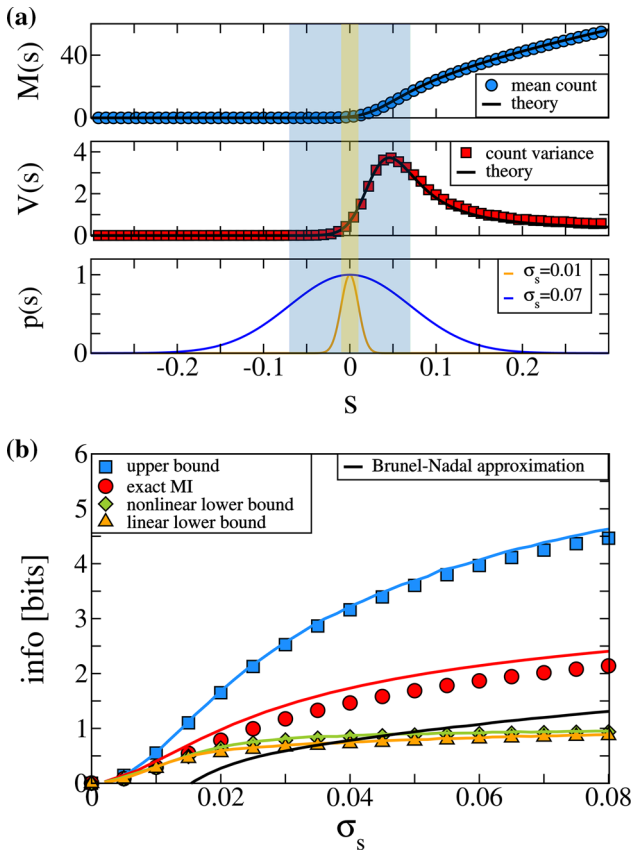

Fig. 4 LIF model with $T = 100$ **and** $\mu = 0.9$. **a** Signal-dependent mean count $M_N(s)$, vs signal value; signal-dependent count variance $V_N(s)$ vs signal value; and rescaled signal distributions for two different values of σ_s. **b** Mutual information (red circles), linear lower bound (orange triangles), nonlinear lower bound (green diamonds) and upper bound (blue squares) vs σ_s. The colored solid lines correspond to the mutual information (red), the linear lower bound (orange), the nonlinear lower bound (green) and the upper bound for the Gaussian model with the same functions $M_N(s)$ and $V_N(s)$ as in **a**. The black solid line shows the Brunel–Nadal approximation for the mutual information (color figure online)

Using the well-known equation for the LIF firing rate (Siegert 1951; Ricciardi 1977)

$$r(s) = \left[\tau_{\text{ref}} + \sqrt{\pi} \int_{(\mu+s-v_R)/\sqrt{2D}}^{(\mu+s-v_T)/\sqrt{2D}} \mathrm{d}x \; \mathrm{e}^{x^2} \, \mathrm{erfc}(x) \right]^{-1} \tag{52}$$

we obtain the analytical result for the function $M_N(s)$ (black line) in Figs. 3a and 4a. For both firing regimes of the LIF model, we find a very good agreement between the numerical estimation and the analytical prediction. For a more detailed illustration of the function $r(s)$ for varying parameters μ and D of the LIF model please refer to Vilela and Lindner (2009a).

In order to find an analytical expression for the function $V_N(s)$ we first express the spike count variance in terms of the Fano factor, $F = \langle \Delta N^2 \rangle / \langle N \rangle$, and the firing rate, $r(s)$, which leads to

$$V_N(s) = Tr(s)F(s) \; . \tag{53}$$

In the limit of large time windows the above equation transforms into

$$\lim_{T\to\infty} V_N(s) = Tr(s)C_v(s)^2 = T \lim_{\omega\to 0} S(\omega) \; , \tag{54}$$

where C_v is the coefficient of variation of the interspike intervals and $\lim_{\omega\to 0} S(\omega)$ is the low-frequency limit of the spike train power spectrum. For the LIF model the coefficient of variation and the spike train power spectrum are known analytically (Lindner et al. 2002). Here we used the low-frequency limit of the analytical equation for the power spectrum for an analytical prediction of the function $V_N(s)$ (black line) in Figs. 3a and 4a. For $s > -0.1$ in the suprathreshold regime and for $s > 0.1$ in the subthreshold regime the analytical theory underestimates the function $V_N(s)$ because of the limit $T \to \infty$ in the analytical formula (here $T = 100$ in both firing regimes). For smaller s we find a good agreement between the numerical estimates and the theoretical predictions.

The Gaussian approximation for the neural spike count, Eq. (50), allows us to use Eqs. (31–33) for the estimation of the mutual information and Eq. (34) together with Eq. (20) and Eq. (21) for the estimation of the linear and nonlinear lower bounds for the LIF model. We also compute the upper bound, Eq. (30). As we discussed above, in the limit of large T the functions $M_N(s)$ and $V_N(s)$ of the LIF model could in principle be computed analytically. For the subsequent analysis of the mutual information and the lower and upper bounds, however, we employ the numerically measured $M_N(s)$ and $V_N(s)$. This allows a better comparison between the LIF spike count and the Gaussian approximation in Eq. (50), without the confounding influence of the discrepancies between the numerically measured functions M_N and V_N and their analytical predictions due to a finite T.

We show the mutual information (red line), the upper bound (blue line), the linear lower bound (orange line) and the nonlinear lower bound (green solid line) for the Gaussian model of the LIF spike count in the suprathreshold firing regime in Fig. 3b. For a vanishing signal strength the mutual information (red line) starts at zero and increases monotonically with increasing σ_s. The upper bound (blue line) closely matches the mutual information for $\sigma_s < 0.1$. The linear lower bound (yellow line) also increases monotonically with σ_s and matches closely the mutual information for $\sigma_s \leq 0.03$. For $\sigma_s > 0.03$, however, the linear lower bound exhibits a qualitatively different behavior from the mutual information and decreases with increasing σ_s. The nonlinear lower bound (green line) follows the mutual information closely for $\sigma_s \leq 0.07$. For stronger signals the nonlinear lower bound

decays with increasing σ_s in a similar way as the linear lower bound but never falls below the latter as predicted by Eq. (23). Both lower bounds provide accurate estimates of the mutual information for weak signals, but the nonlinear lower bound does so for a broader range of signal strengths than the linear lower bound.

We also consider the mutual information and its upper and lower bounds for the Gaussian model of the LIF spike count in the subthreshold regime ($\mu = 0.9$) in Fig. 4b. The mutual information (red line) increases monotonically for increasing σ_s but does not rise as fast as in the suprathreshold regime. At $\sigma_s = 0.1$ the mutual information attains a value of approximately 2 bits in the subthreshold regime as compared to approximately 4 bits in the suprathreshold regime. The upper bound (blue line) increases monotonically with σ_s, but for stronger signals the difference between the upper bound and the mutual information is much stronger in the subthreshold regime than in the suprathreshold regime. In contrast to the suprathreshold regime for which the linear and the nonlinear lower bounds (yellow and green lines, respectively) exhibited a peak with respect to σ_s, in the subthreshold regime they increase monotonically with σ_s. Both lower bounds coincide with the mutual information for $\sigma_s < 0.01$ but significantly underestimate the mutual information for larger signal strengths. In contrast to the suprathreshold regime, the nonlinear lower bound does not provide any significant improvement for the estimation of the mutual information as compared to the linear lower bound.

5.2 Numerical estimation of the mutual information and its upper and lower bounds for the LIF model

We now turn to the numerical estimation of the mutual information between the Gaussian signal and the spike count of the LIF neuron model. In particular, we repeatedly draw realizations of a Gaussian signal, which then drive the LIF model, Eq. (6). For each specific signal realization, we record the number of spikes which occurred within a time window T. From the resulting list of pairs of signal realizations and spike counts we then estimate the probabilities $P(N, s)$, $P(s)$ and $P(N)$ and determine the mutual information by means of Eq. (3) [please refer to Voronenko (2018) for a more detailed discussion of the numerical estimate]. For every measurement of the spike count, the LIF model is subject to independent intrinsic noise. From the list of signal and spike count pairs we also numerically determine the linear and nonlinear correlation coefficients and compute the linear and nonlinear lower bounds by employing Eqs. (21) and (20). Finally, we perform a binning in s of the signal and spike count pairs, determine the signal-dependent output variance, $V(s)$, and the total output variance, σ_x^2, and obtain the upper bound by means of Eq. (30).

The numerical results for the mutual information, and the linear and nonlinear lower bounds for the LIF model in the suprathreshold regime ($\mu = 1.1$) are shown as functions of σ_s in Fig. 3b. We find a very good agreement between the mutual information which we estimated for the LIF spike count (colored symbols) and the mutual information which we computed for the Gaussian model of the spike count (colored lines). This good agreement allows us to use the functions $M_N(s)$ and $V_N(s)$ for the interpretation of the behavior of the nonlinear lower bound. First, we consider a weak signal with $\sigma_s = 0.01$, the distribution of which is illustrated (orange line) in Fig. 3a. The range of values which are predominantly sampled by the signal is indicated by the orange shading. In the relevant signal range, the functions $M_N(s)$ and $V_N(s)$ are effectively linear and this leads to the good agreement of the linear lower bound, the nonlinear lower bound and the mutual information for $\sigma_s < 0.01$ in Fig. 3b. A strong signal with $\sigma_s = 0.2$ has a much broader distribution (green solid line in Fig. 3a) and samples a much larger range of signal values (green shading). In this range the functions $M_N(s)$ and $V_N(s)$ are strongly nonlinear, which is the reason for the large discrepancy between the mutual information and its lower bounds for $\sigma_s = 0.2$. For a signal of intermediate strength with $\sigma_s = 0.07$ (distribution indicated by the blue solid line in Fig. 3a) the range of values which are predominantly sampled by the signal (blue shading) is such that in this range the functions $M_N(s)$ and $V_N(s)$ exhibit only weak nonlinearities, i.e., nonlinearities which could be captured by the first terms of a Taylor expansion. As we have already seen in the analysis of the Gaussian model with a nonlinear signal dependence of the mean in the previous chapter, for this type of nonlinearities the nonlinear lower bound comprises a much better approximation of the mutual information than the linear lower bound. Indeed, we find that for $\sigma_s = 0.07$ the nonlinear lower bound is much closer to the mutual information than the linear lower bound (Fig. 3b).

In the subthreshold regime in Fig. 4b, the agreement between the mutual information for the LIF spike count (red circles) and the mutual information for the Gaussian model (red line) is worse than in the suprathreshold regime but still reasonably good. The reason for the disagreement between the LIF spike count and the Gaussian model can be explained by the low mean spike count in the subthreshold regime. For our choice of the observation time ($T = 100$), the discrete nature of the spike count influences the behavior of the mutual information. For the lower bounds for the mutual information for the LIF spike count and the Gaussian model the agreement is remarkably good, despite the discussed limitations of the Gaussian model. As before we can therefore use the Gaussian model for the interpretation of the behavior of the nonlinear lower bound for different signal strengths. For weak signals with $\sigma_s = 0.01$ the signal distribution is narrow (orange line in Fig. 4a) and the signal predominantly samples a range of

signal values (indicated by the orange shading) in which the functions M_N and V_N are approximately linear. This leads to the good agreement between the lower bounds and the mutual information for $\sigma_s = 0.01$ in Fig. 4b. For a relatively strong signal with $\sigma_s = 0.07$, the signal distribution is wider (blue line in Fig. 4a) and the signal predominantly samples a range of values (blue shading) over which the functions M_N and V_N are strongly nonlinear, i.e., they cannot be expressed by the first terms of a Taylor expansion. Consequently, the mutual information is strongly underestimated by the lower bounds for $\sigma_s = 0.07$ in Fig. 4b. Interestingly, in the subthreshold firing regime of the LIF there are no values of the signal strength for which the nonlinear lower bound significantly exceeds the linear lower bound and provides a good agreement with the mutual information.

An alternative approximation for the mutual information between a static signal and the neural spike count was proposed by Brunel and Nadal (1998). The authors model the neural spike count by the Gaussian model, Eq. (50). In the limiting case in which the modulation of the mean output by the signal is much stronger than the modulation of the noise variance, the authors derive an approximate expression for the mutual information. For a Gaussian signal, their result reduces to a single integral

$$\mathcal{M}_{BN} = \frac{1}{\sqrt{8\pi}} \int_{-\infty}^{\infty} \mathrm{d}z \; \mathrm{e}^{-\frac{z^2}{2}} \log_2 \left(\frac{\sigma_s^2 M_N'^2(z\sigma_s)}{V_N(z\sigma_s)} \right) , \tag{55}$$

where the prime denotes the derivative of the mean with respect to s. For general functions $M_N(s)$ and $V_N(s)$, the above equation cannot be simplified further and the integral has to be evaluated numerically. However, in contrast to the mutual information of our Gaussian model in Eq. (31) with two integrals, Eq. (55) requires a numerical integration over only one variable. We compare the approximation from Eq. (55) with the mutual information for the LIF spike count and with the mutual information of the Gaussian model in Figs. 3b and 4b. In the suprathreshold regime, the approximation from Eq. (55) (black line) is in good agreement with the mutual information (red circles) for $\sigma_s < 0.15$ and exhibits a slight disagreement for stronger signals. In the subthreshold regime, however, the approximation from Eq. (55) (black line in Fig. 4b) consistently underestimates the mutual information irrespective of the signal strength. For weak signals ($\sigma_s < 0.015$) the approximation by Brunel and Nadal predicts negative values for the mutual information because the argument of the logarithm in Eq. (55) becomes smaller than one for a significant range of signal values. This behavior is a consequence of the assumption of Brunel and Nadal about the smallness of $V_N(s)$. Note that although in our examples the Brunel–Nadal approximation is always equal to or smaller than the actual mutual information, it is in general not a lower bound.

5.3 Dependence of the linear and nonlinear correlation coefficients on the observation time *T*

The Gaussian approximation for the spike count allows us to investigate the dependence of the linear and nonlinear correlation coefficients on T. First, we use Eqs. (51) and (53) and express the linear and nonlinear correlation coefficients, Eq. (34), in terms of the signal-dependent firing rate $r(s)$ and the signal-dependent Fano factor $F(s)$. For the linear correlation coefficient between signal and output we find

$$\begin{aligned} \rho^2_{s,x_N}(T) &= \frac{\langle (s-\langle s\rangle)(x_N-\langle x_N\rangle)\rangle^2}{\langle (s-\langle s\rangle)^2\rangle \langle (x_N-\langle x_N\rangle)^2\rangle} \\ &= \frac{\rho^2_{s,r}}{1+\frac{1}{T}\frac{\langle r(s)F(s)\rangle}{\langle (r(s)-\langle r(s)\rangle)^2\rangle}} . \end{aligned} \tag{56}$$

For the nonlinear correlation coefficient between the signal and the squared output we find

$$\begin{aligned} \rho^2_{s,x_N^2} &= \frac{\langle (s-\langle s\rangle)(x_N^2-\langle x_N^2\rangle)\rangle^2}{\langle (s-\langle s\rangle)^2\rangle \langle (x_N^2-\langle x_N^2\rangle)^2\rangle} \\ &= \frac{\rho^2_{s,r^2}}{1+\frac{4}{T}\frac{\langle r(s)^3F(s)\rangle}{\langle (r(s)^2-\langle r(s)^2\rangle)^2\rangle}+\frac{2}{T^2}\frac{\langle r(s)^2F(s)^2\rangle}{\langle (r(s)^2-\langle r(s)^2\rangle)^2\rangle}} . \end{aligned} \tag{57}$$

Finally we also obtain the nonlinear correlation coefficient between the output and the squared output,

$$\begin{aligned} \rho^2_{x_N,x_N^2} &= \frac{\langle (x_N-\langle x_N\rangle)(x_N^2-\langle x_N^2\rangle)\rangle^2}{\langle (x_N-\langle x_N\rangle)^2\rangle \langle (x_N^2-\langle x_N^2\rangle)^2\rangle} \\ &= \frac{\rho^2_{r,r^2}\left(1+\frac{2}{T}\frac{\langle r(s)^2F(s)\rangle}{\langle (r(s)-\langle r(s)\rangle)(r(s)^2-\langle r(s)^2\rangle)\rangle}\right)^2}{\left(1+\frac{4}{T}\frac{\langle r(s)^3F(s)\rangle}{\langle (r(s)^2-\langle r(s)^2\rangle)^2\rangle}+\frac{2}{T^2}\frac{\langle r(s)^2F(s)^2\rangle}{\langle (r(s)^2-\langle r(s)^2\rangle)^2\rangle}\right)} \\ &\quad \times \left[1+\frac{1}{T}\frac{\langle r(s)F(s)\rangle}{\langle (r(s)-\langle r(s)\rangle)^2\rangle}\right]^{-1} . \end{aligned} \tag{58}$$

Taking the limit of $T \to \infty$ in Eqs. (56–58) we finally arrive at

$$\begin{aligned} &\lim_{T\to\infty} \rho^2_{s,x_N} = \rho^2_{s,r} , \quad \lim_{T\to\infty} \rho^2_{s,x_N^2} = \rho^2_{s,r^2} , \\ &\lim_{T\to\infty} \rho^2_{x_N,x_N^2} = \rho^2_{r,r^2} . \end{aligned} \tag{59}$$

From the above equations it can be seen that in the limit $T \to \infty$ the linear and nonlinear correlation coefficients

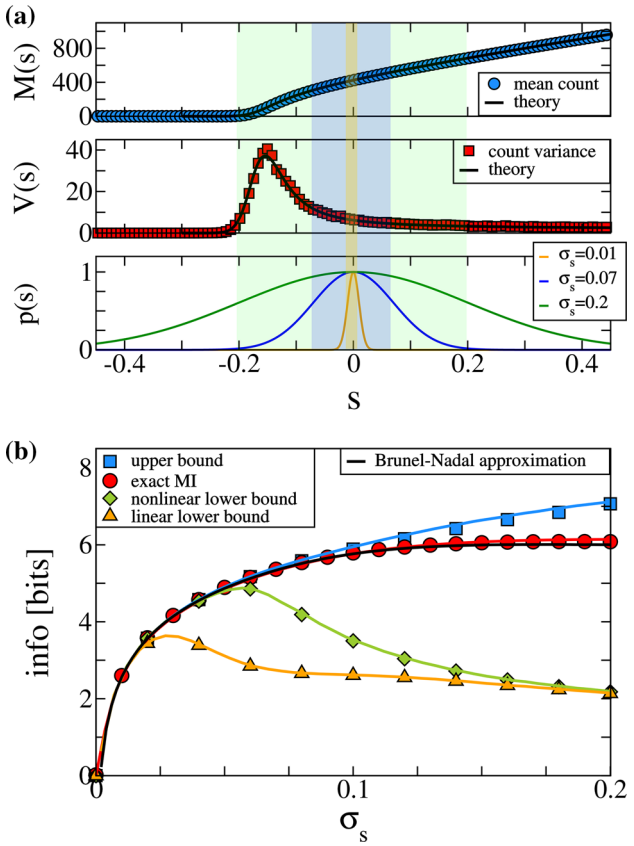

Fig. 5 LIF model with $T = 1000$ **and** $\mu = 1.1$. **a** Signal-dependent mean count $M_N(s)$, vs signal value; signal-dependent count variance $V_N(s)$ vs signal value; and rescaled signal distributions for three different values of σ_s. **b** Mutual information (red circles), linear lower bound (orange triangles), nonlinear lower bound (green diamonds) and upper bound (blue squares) vs σ_s. The colored solid lines correspond to the mutual information (red), the linear lower bound (orange), the nonlinear lower bound (green) and the upper bound for the Gaussian model with the same functions $M_N(s)$ and $V_N(s)$ as in **a**. The black solid line corresponds to the Brunel–Nadal approximation for the mutual information (color figure online)

converge to finite values, which do not depend on $F(s)$ but are expressed in terms of correlation coefficients between the signal and the function $r(s)$. This is consistent with the fact that for long time windows the noise is effectively averaged out and only the signal dependence of the mean contributes to the estimation of the mutual information.

In Fig. 5a we show the numerically measured functions $M_N(s)$ and $V_N(s)$ for $T = 1000$ and compare the numerical results to theoretical predictions from Eqs. (51) and (54) (black lines). In comparison with the suprathreshold firing regime for $T = 100$ (Fig. 3a) we can see that both functions $M_N(s)$ and $V_N(s)$ increased by a factor of 10 and that the theoretical prediction of $V_N(s)$ now exhibits a better match with the numerical measurements. Because in the stochastic equation for the Gaussian variable x_N, Eq. (50), the function $V_N(s)$ enters as the argument of a square root, the increase of T from 100 to 1000 effectively diminishes the strength of the noise in comparison with the noise independent term $M_N(s)$. Consequently, the output x_N is less noisy and we expect a better transmission of information about the signal. Indeed, Fig. 5b shows that the mutual information (red circles) increases in comparison with the setup with $T = 100$ for all values of σ_s (for $\sigma_s = 0.2$ it increases from 4.5 bits to approximately 6 bits).

For the linear and the nonlinear lower bounds (orange triangles and green diamonds, respectively) we find a qualitatively different behavior. For strong signals (e.g., $\sigma_s = 0.2$) both lower bounds attain the value of 2 bits independent of the choice of T. For intermediate signals however, both lower bounds attain larger values for $T = 1000$ than for $T = 100$. Furthermore, the nonlinear lower bound exhibits a stronger increase than the linear lower bound. For $\sigma_s = 0.6$ the nonlinear lower bound increases from approximately 3.5 bits for $T = 100$ to approximately 5 bits for $T = 1000$. From our discussion of the nonlinear correlation coefficients in the limit of large T in Eq. (59) we can infer that the relative increase of the nonlinear lower bound with respect to the linear lower bound can be attributed to the decreased effect of the noise on the output x_N and on the nonlinear signal dependence of the function $r(s)$ which leads to nonvanishing nonlinear correlation coefficients ρ_{s,r^2} and ρ_{r,r^2}. The mutual information and its upper and lower bounds for the Gaussian model (red, blue, orange and green lines) are again in good agreement with the numerical results for the LIF model (colored symbols). Finally note that due to the output variability of the spike count for $T = 1000$ the approximation by Brunel and Nadal (black line) shows a good agreement with the mutual information even for large signal strengths.

5.4 Nonlinear information transmission for the Na–K model

In order to demonstrate that our results are not specific to the LIF model but are of a more general nature, we apply our theory to the Na–K model with ion channel noise. First we numerically estimate the mutual information between the signal and the neural spike count (red circles in Fig. 6b). Furthermore, we numerically measure the linear and nonlinear correlation coefficients between the signal and the neural spike count and compute the linear lower bound (orange triangles) and the nonlinear lower bound (green diamonds) for the mutual information. Finally, we numerically determine the signal-dependent spike count variance and compute the upper bound (blue squares). We find that the MI show a very similar behavior as in the suprathreshold firing regime of the LIF model. For weak signals ($\sigma_s < 0.3$) the lower and upper bounds coincide with the mutual information. For strong signals ($\sigma_s > 0.7$) the mutual information is overestimated by the upper bound and significantly underestimated

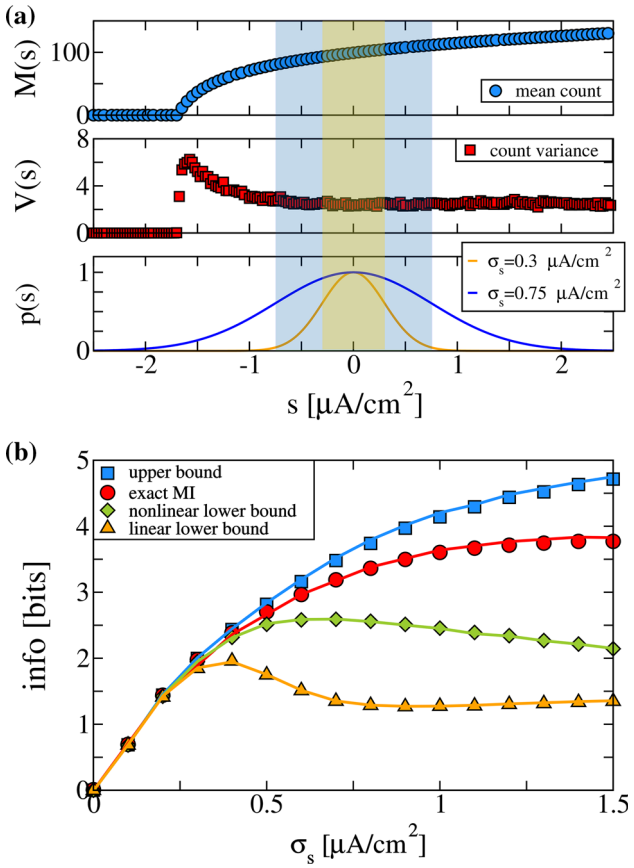

Fig. 6 NaK model with $T = 1000$ms **and** $I_0 = 6\mu$A/cm^2. **a** Signal-dependent mean count $M_N(s)$, vs signal value; signal-dependent count variance $V_N(s)$ vs signal value; and rescaled signal distributions for two different values of σ_s. **b** Mutual information (red circles), linear lower bound (orange triangles), nonlinear lower bound (green diamonds) and upper bound (blue squares) vs σ_s. The colored solid lines correspond to the mutual information (red), the linear lower bound (orange), the nonlinear lower bound (green) and the upper bound for the Gaussian model with the same functions $M_N(s)$ and $V_N(s)$ as in **a** (color figure online)

by the linear and the nonlinear lower bounds. For signals of intermediate strength ($0.3 < \sigma < 0.7$) the nonlinear lower bound provides a significantly better approximation for the mutual information than the linear lower bound.

We also determine the mean spike count as a function of the value of the input signal, $M_N(s)$, and the signal-dependent spike count variance, $V_N(s)$, which are shown in Fig. 6a. As for the LIF model, we use the functions $M_N(s)$ and $V_N(s)$ for a Gaussian model of the neural spike count. For this model we then compute the mutual information (red line Fig. 6b), the upper bound (blue line), the linear lower bound (orange line) and the nonlinear lower bound (green line). We find that the results for the Gaussian model are in good agreement with the mutual information between the signal and the spike count and its upper and lower bounds.

For two values of the signal standard deviation, $\sigma_s = 0.3$ and $\sigma_s = 0.75$, we plot a rescaled version of the signal distributions (orange and blue lines, respectively) in the third panel of Fig. 6a. For $\sigma_s = 0.3$ the signal predominantly samples a narrow range of signal values (orange -shaded region); for $\sigma_s = 0.75$ the sampled signal range is considerably broader (blue-shaded region). The widths of the two sampling regions in turn determine which parts of the functions $M_N(s)$ and $V_N(s)$ are sampled by the signal. If the functions $M_N(s)$ and $V_N(s)$ are sampled in a region where they are essentially linear, then the linear lower bound provides a reliable estimate of the mutual information (e.g., for $\sigma_s = 0.3$ in Fig. 6b). However, for signals of intermediate strengths, which sample the functions $M_N(s)$ and $V_N(s)$ in a region where they are weakly nonlinear, the nonlinear lower bound provides a significant improvement for the approximation of the mutual information as compared to the linear lower bound (e.g., for $\sigma_s = 0.75$ in Fig. 6b).

6 Discussion

In this paper we derived a nonlinear lower bound for the mutual information of a nonlinear system which is driven by a static Gaussian signal. This bound incorporates nonlinear correlation coefficients between input and output. We measured this nonlinear bound for two simple models with Gaussian noise and for two spiking neuron models (a stochastic LIF neuron with white Gaussian noise and Izhikevich's Na–K model endowed with channel noise) and compared the nonlinear bound to the mutual information and to previously suggested approximations of it.

As demonstrated in several examples, our formula provides a significant improvement of the standard linear lower bound approximation if the system is governed by a weak nonlinearity and the signal amplitude is in an intermediate range. In biological systems the weak nonlinearity often arises for neurons which are subject to sufficient background noise. Furthermore, we can generally expect that the most relevant signals for sensory neurons are the ones that are neither barely detectable nor overwhelmingly strong but in an intermediate range (for a recent discussion of noise and signal amplitudes in a biological model system, the electro-sensory system of weakly electric fish, see Grewe et al. 2017).

We emphasize that the calculation of our formula hardly requires much more efforts than for the computation of the linear lower bound: We have to compute two more correlation coefficients, which can be quickly done both in neural (experimental) data and in neuron models that are explored by numerical simulations. Hence, our results offer a significant improvement for both theoretical and experimental studies of neural information transmission.

Our work also provides a link between nonlinear input–output correlations and the mutual information. Although nonlinear signal reconstructions have been successfully

applied to biological systems in the past (e.g., Marmarelis and Naka 1972), many experimental studies have argued that the nonlinear terms of the signal reconstruction do not contribute to a better estimation of the mutual information (Bialek et al. 1991, 1993). Here, we demonstrated that for an intermediate stimulus range the nonlinear correlation coefficients contribute to a significant improvement of the lower bound. Furthermore, the deviation of the mutual information from the linear lower bound can be completely understood by the nonlinearity of transmission and in particular by a quadratic nonlinearity. We believe that our result will be helpful in further assessing the role of nonlinear input–output correlations for neural signal transmission. We also think that for models for which the nonlinear correlation coefficients can be computed analytically (e.g., for the IF models with white background noise) our work can contribute to new analytic studies of neural signal transmission.

In our approach we used a specific (quadratic) nonlinearity. One obtains very similar formulas if other nonlinear functions are used. In particular, if the term proportional to N^2 in Eq. (8) is replaced by some nonlinear function $f(N)$, then the nonlinear lower bound is given by the same expression as in Eq. (20) but where ρ_{N,N^2} and ρ_{s,N^2} are replaced by $\rho_{N,f(N)}$ and $\rho_{s,f(N)}$, respectively. One possible direction for further research is to investigate whether additional terms of a higher order than N^2 could further improve the nonlinear lower bound. Another interesting approach would be to try to find a nonlinear function which gives the strongest difference from the linear lower bound (i.e., the most significant improvement). If it is possible to uniquely determine such an optimal function, then it would be interesting to know how it is related to the nonlinearity of the system.

In this paper we improved the linear lower bound by incorporating nonlinear terms into the signal reconstruction. We then used the mean squared error of the signal reconstruction to obtain a lower bound for the mutual information. A different approach to improve the linear lower bound (and also further improve our nonlinear lower bound) would be to find a relation between the mutual information and the higher-order moments of the signal reconstruction (not only the mean squared error but also for example the skewness and kurtosis). We believe that this approach would further improve the lower bounds for strong signals or for neural systems in strongly nonlinear firing regimes.

Another possible extension of our work concerns the case of multiple neurons transmitting information about the same signal (but being subject to different but possibly correlated noise sources). Finally, the case of a time-dependent stimulus seems to be the most involved generalization of our approach: Here we expect that terms involving higher-order coherence functions (Nikias and Petropulu 1993), reflecting in particular the synergetic transmission of information about signal components from multiple frequency bands (Bernardi and Lindner 2015), will replace the correlation functions of the simple static system that we have investigated.

Acknowledgements This work was supported by the BMBF (FKZ: 01GQ1001A) and the DFG (Research Training Group GRK1589/2).

References

Aldworth ZN, Dimitrov AG, Cummins GI, Gedeon T, Miller JP (2011) Temporal encoding in a nervous system. PLOS Comput Biol 7(5):e1002041

Bernardi D, Lindner B (2015) A frequency-resolved mutual information rate and its application to neural systems. J Neurophysiol 113(5):1342–1357

Bialek W, Rieke F, Vansteveninck RRD, Warland D (1991) Reading a neural code. Science 252:1854

Bialek W, Deweese M, Rieke F, Warland D (1993) Bits and brains—information-flow in the nervous system. Physica A 200:581

Borst A, Theunissen F (1999) Information theory and neural coding. Nat Neurosci 2:947

Brunel N, Nadal JP (1998) Mutual information, fisher information, and population coding. Neural Comput 10:1731

Burkitt AN (2006) A review of the integrate-and-fire neuron model: I. Homogeneous synaptic input. Biol Cybern 95:1

Chacron MJ, Doiron B, Maler L, Longtin A, Bastian J (2003) Non-classical receptive field mediates switch in a sensory neuron's frequency tuning. Nature 423:77

Cover T, Thomas J (1991) Elements of information theory. Wiley, New York

Cox DR (1962) Renewal theory. Methuen, London

Doose J, Doron G, Brecht M, Lindner B (2016) Noisy juxtacellular stimulation in vivo leads to reliable spiking and reveals high-frequency coding in single neurons. J Neurosci 36(43):11120–11132

Droste F, Schwalger T, Lindner B (2013) Interplay of two signals in a neuron with heterogeneous short-term synaptic plasticity. Front Comput Neurosci 7:1

Gabbiani F (1996) Coding of time-varying signals in spike trains of linear and half-wave rectifying neurons. Netw Comput Neural Syst 7:61

Gammaitoni L, Hänggi P, Jung P, Marchesoni F (1998) Stochastic resonance. Rev Mod Phys 70:223

Grewe J, Kruscha A, Lindner B, Benda J (2017) Synchronous spikes are necessary but not sufficient for a synchrony code in populations of spiking neurons. PNAS 114(10):1977–1985

Izhikevich EM (2007) Dynamical systems in neuroscience: the geometry of excitability and bursting. The MIT Press, Cambridge

Juusola M, French AS (1997) The efficiency of sensory information coding by mechanoreceptor neurons. Neuron 18(6):959–968

Kraskov A, Stögbauer H, Grassberger P (2004) Estimating mutual information. Phys Rev E 69(6):066138

Lagarias J (2013) Euler's constant: Euler's work and modern developments. Bull Am Math Soc 50(4):527–628

Lindner B, Sokolov IM (2016) Giant diffusion of underdamped particles in a biased periodic potential. Phys Rev E 93(4):042106

Lindner B, Schimansky-Geier L, Longtin A (2002) Maximizing spike train coherence or incoherence in the leaky integrate-and-fire model. Phys Rev E 66:031916

Marmarelis PZ, Naka K (1972) White-noise analysis of a neuron chain: an application of the Wiener theory. Science 175(4027):1276–1278

McDonnell MD, Ward LM (2011) The benefits of noise in neural systems: bridging theory and experiment. Nat Rev Neurosci 12:415

Morris C, Lecar H (1981) Voltage oscillations in the barnacle giant muscle fiber. Biophys J 35:193
Neiman AB, Russell DF (2011) Sensory coding in oscillatory electroreceptors of paddlefish. Chaos 21:047505
Nemenman I, Lewen GD, Bialek W, de Ruyter van Steveninck RR (2008) Neural coding of natural stimuli: information at submillisecond resolution. PLOS Comput Biol 4:e1000025
Nikias CL, Petropulu AP (1993) Higher-order spectral analysis. PTR Prentice Hall, Upper Saddle River
Panzeri S, Schultz SR (2001) A unified approach to the study of temporal, correlational, and rate coding. Neural Comput 13(6):1311–1349
Panzeri S, Senatore R, Montemurro MA, Petersen RS (2007) Correcting for the sampling bias problem in spike train information measures. J Neurophys 98(3):1064–1072
Passaglia CL, Troy JB (2004) Information transmission rates of cat retinal ganglion cells. J Neurophysiol 91(3):1217–1229
Reimann P, Van den Broeck C, Linke H, Hänggi P, Rubi M, Perez-Madrid A (2001) Giant acceleration of free diffusion by use of tilted periodic potentials. Phys Rev Lett 87:010602
Ricciardi LM (1977) Diffusion processes and related topics on biology. Springer, Berlin
Rieke F, Warland D, Bialek W (1993) Coding efficiency and information rates in sensory neurons. Europhys Lett 22:151
Rieke F, Bodnar D, Bialek W (1995) Naturalistic stimuli increase the rate and efficiency of information transmission by primary auditory afferents. Proc Biol Sci 262:259
Rieke F, Warland D, de Ruyter van Steveninck R, Bialek W (1996) Spikes: exploring the neural code. MIT Press, Cambridge
Ryzhik IM, Gradshtein IS (1963) Tables of series, products, and integrals. VEB Deutscher Verlag der Wissenschaften, Berlin
Sadeghi SG, Chacron MJ, Taylor MC, Cullen KE (2007) Neural variability, detection thresholds, and information transmission in the vestibular system. J Neurosci 27(4):771–781
Shannon R (1948) The mathematical theory of communication. Bell Syst Tech J 27:379
Siegert AJF (1951) On the first passage time problem. Phys Rev 81:617
Stemmler M, Koch C (1999) How voltage-dependent conductances can adapt to maximize the information encoded by neuronal firing rate. Nat Neurosci 2:521
Strong SP, Koberle R, van Steveninck RRD, Bialek W (1998) Entropy and information in neural spike trains. Phys Rev Lett 80:197
Theunissen F, Miller JP (1995) Temporal encoding in nervous systems: a rigorous definition. J Comput Neurosci 2(2):149–162
Thomas PJ, Lindner B (2014) Asymptotic phase for stochastic oscillators. Phys Rev Lett 113(25):254101–5
Victor JD (2002) Binless strategies for estimation of information from neural data. Phys Rev E 66(5):051903
Victor JD (2006) Approaches to information-theoretic analysis of neural activity. Biol Theory 1(3):302–316
Vilela RD, Lindner B (2009a) Are the input parameters of white-noise-driven integrate and fire neurons uniquely determined by rate and CV? J Theor Biol 257:90
Vilela RD, Lindner B (2009b) A comparative study of three different integrate-and-fire neurons: spontaneous activity, dynamical response, and stimulus-induced correlation. Phys Rev E 80:031909
Voronenko SO (2018) Nonlinear signal processing by noisy spiking neurons. Ph.D. thesis, Humboldt-Universität zu Berlin, Berlin
Voronenko SO, Lindner B (2017) Weakly nonlinear response of noisy neurons. New J Phys 19(3):033–038
Voronenko SO, Stannat W, Lindner B (2015) Shifting spike times or adding and deleting spikes—how different types of noise shape signal transmission in neural populations. JMN 5(1):1–35