Goldsmiths Research Online

*Goldsmiths Research Online (GRO)
is the institutional research repository for
Goldsmiths, University of London*

Citation

Persistent URL

Versions

Goldsmiths
UNIVERSITY OF LONDON

Information transfer through a signaling module with feedback: a perturbative approach

Gerardo Aquino[1], Martin Zapotocky[2]

[1] *Department of Life Sciences, Imperial College, SW7 2AZ, London, UK,*
[2] *Institute of Physiology of the Czech Academy of Sciences, Videnska 1083, 14220 Prague, Czech Republic.*

Abstract

Signal transduction in biological cells is effected by signaling pathways that typically include multiple feedback loops. Here we analyze information transfer through a prototypical signaling module with biochemical feedback. The module switches stochastically between an inactive and active state; the input to the module governs the activation rate while the output (i.e., the product concentration) perturbs the inactivation rate. Using a novel perturbative approach, we compute the rate with which information about the input is gained from observation of the output. We obtain an explicit analytical result valid to first order in feedback strength and to second order in the strength of input. The total information gained during an extended time interval is found to depend on the feedback strength only through the total number of activation/inactivation events.

Keywords: Signal transduction, Communication channel, Poisson process, Information theory, Feedback loop, Non-Markovian process

1. Introduction

Accurate sensing of the environment is crucial for the survival of biological organisms. Bacteria, as well as animal chemoreceptor cells, can sense certain chemicals in their chemical environment with high precision, in some cases near the single-molecule detection limit [1, 2]. The effect of the extracellular stimulus on the cell is mediated by the signal transduction pathway - a complex

biochemical reaction network. The pathway is based on a sequence of transduction steps, with each subsequent step being effected by the chemical product of the previous step. The first step typically consists in the activation of a receptor protein in the cell membrane by the external stimulus, which results in an ion influx or in the production of a second messenger chemical. This leads to the activation of subsequent steps within the cell. Through molecular feedback loops, the product of a given transduction step may regulate its own production, or influence earlier (upstream) steps of the pathway. As signal transduction pathways are inherently noisy [3], faithful transmission of information through the pathways requires amplification and/or adaptation. This is often accomplished through positive (amplification) and negative (adaptation) feedback built into the reaction network.

Recent years have brought the use of information-based measures to characterize the reliability of biological signal transduction [4, 5, 6, 7]. Such measures explicitly evaluate the amount of information about the stimulus that it is trasmitted through the signaling pathway. Mutual information between the stimulus and the pathway output (i.e., the product of the final transduction step) was evaluated in, e.g.[5, 6] , to assess the precision with which stationary stimuli of different strengths can be distinguished. In many signaling scenarios, however, it is important to faithfully transduce the temporal variations of the input stimulus, which encode biologically important information. Some recent studies have evaluated the transduction reliability for time-varying signals by computing the information transmission rate. In [7], prototypical signaling pathways with feedback were represented by coupled Langevin equations with additive Gaussian noise, and thefrequency-dependent gain-to-noise ratio was computed. In the Gaussian noise approximation, this gain (together with the input power spectrum) determines the the mutual information between the time courses of the stimulus and of the output. Some of the stochasticity within a signaling pathway, however, arises directly from the inherent stochastic dynamics of the transduction components, and cannot always betreated as additive noise; in such a case, the information calculation cannot be reduced to the evaluation

of the gain-to-noise ratio.

In this work, we carry out a perurbative computation of information transfer through a simple prototypical signaling module with biochemical feedback. The module switches stochastically between two states, with switching rates governed by the stimulus (i.e., input) and by the product (i.e., output). No additive external noise is assumed. Such autoregulated stochastic modules arise within various signal transduction and gene regulation pathways (see Sec. 2). In a previous investigation [8] negative feedback, in this module, was shown to decrease the signal-to-noise ratio (SNR) at the output, but at the same time to increase the spectral range of the response - thus yielding no obvious expectation on how feedback overall affectsinformation transmission through the module. Here, we address this question by directly quantifying the information that is gained about the external stimulus from following the module output. In order to achieve this we introduce a novel perturbative approach on a conveniently defined relative entropy for stochastic point processes [9]. This information gain is well-defined for a single stimulus trajectory (i.e., it requires no averaging over stimuli as in mutual-information-based measures) [9]. We obtain an explicit analytical result valid to first order in feedback strength and to second order in the strength of input. Surprisingly, the total information gained during a long time interval is found to be proportional to the total number of state-switching events, with no further dependence on the feedback strength or on the spectral distribution of the input. We compare this result to previous investigations of information transfer through some related information channels.

2. A two-state signaling module with feedback

We consider a simple signaling module based on a single protein that switches between two conformational states. These may correspond to the open and closed state of an ion channel, or to the active/inactive states of an enzyme within a larger signaling network. We introduce the module by referring to the example of a calcium ion channel that is autoregulated by calcium-mediated

3

feedback. When the channel is open, calcium ions flow from the extracellular space into the cell. This leads to a fast increase of the free calcium concentration in the immediate vicinity of the channel (the increase is localized as calcium buffering in the cytoplasm leads to the formation of a calcium microdomain [11]). For certain types of calcium channels (such as the voltage-activated L-type channels [12] or the cyclic-nucleotide-gated (CNG) channels [13, 14]), the cytoplasmic calcium can inactivate the channel when it binds to the channel/calmodulin complex. This implements an autoregulatory feedback loop that shortens the response to the external gating signal, and thus helps to faithfully transduce fast signal variations. Similar autoregulatory loops, in which the product of a particular step in the pathway downregulates its own production, arise in numerous signal transduction and gene regulation networks [15].

The stochastic switching of the channel state is governed by the opening rate (assumed to depend on the extracellular stimulus gating the channel) and the closing rate. The negative autoregulation may be effected through a calcium-dependent increase of the closing rate or decrease of the opening rate. For the CNG channels that motivated us in this study, electrophysiological data indicates that the binding of calcium to the channel/calmodulin complex increases the closing rate. We consequently make only the closing rate depend on the calcium concentration.

The signaling module is shown schematically in Fig. 1. When the channel is in the open state, ions flow into the microdomain at a fixed rate J. Once inside the microdomain, the ions are cleared out through ion pumps or exchangers in the membrane, as well as by diffusion within the cytoplasm; we assume a first-order clearance kinetics with rate constant λ (see Fig. 1). The dynamics of the concentration c of ions in the cell compartment is given by the following equation:

$$\frac{dc(t)}{dt} = \frac{J}{\Delta} S(t) - \lambda c(t), \tag{1}$$

where J is the flow of ions entering the cell through the open channel, Δ the

4

volume of the cell micro-domain and the two-valued function $S(t) = 1$ or 0 indicates the open or closed state of the channel. Adopting the dimensionless
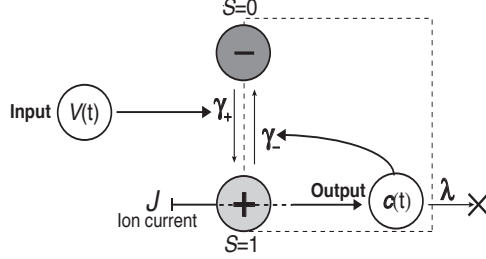

Figure 1: The ion-channel (gray) opens with rate γ_+ governed by the input $V(t)$ and closes with rate γ_- that depends on the ion concentration $c(t)$ (which is considered to be the output). Ions inside the micro-domain (dashed box) are cleared out at rate λ.

variable $\sigma(t) = c(t)\lambda\Delta/J$, Eq. (1) becomes

$$\frac{d\sigma}{dt} = \lambda(S(t) - \sigma(t)) \tag{2}$$

and $\sigma(t)$ is restricted to the range $[0, 1]$. Examples of trajectories of $\sigma(t)$ are shown in Fig. 2. The switching events of the channel between time 0 and t:

$$0 < t_1 < t_2 < \cdots < t_n < t \tag{3}$$

are the realization of a stochastic point process driving the dynamics of $\sigma(t)$. In fact, fixed a sequence $\vec{t}_n = (0, t_1, \cdots, t_n)$ of such events and the initial condition $S(0) = i$ and $\sigma(0) = \sigma_0$, an exact solution for $\sigma(t)$ follows directly from Eq. (2):

$$\sigma(t|\vec{t}_{2n}) = i + \left[\sigma_0 - i + (-1)^i \sum_{j=0}^{n-1} \left(e^{\lambda t_{2j}} - e^{\lambda t_{2j+1}}\right)\right] e^{-\lambda t}$$

$$\sigma(t|\vec{t}_{2n+1}) = 1 - i + \left[\sigma(t = t_{2n+1}|\vec{t}_{2n}) - 1 + i\right] e^{-\lambda(t - t_{2n+1})}. \tag{4}$$

Eqs. (4) determine $\sigma(t)$ after either an even or an odd sequence of switching events and show how this variable keeps a full record (memory) of these events.

We assume that the external input $V(t)$ gates the channel by perturbing only the opening rate, i.e.

$$\tilde{\gamma}_+ = \gamma_+(1 + \varepsilon_V V(t)), \tag{5}$$

while the internal ion concentration affects the closing rate

$$\tilde{\gamma}_- = \gamma_-(1 + \varepsilon_\sigma \sigma(t)). \tag{6}$$

When $\varepsilon_\sigma > 0$, an increase in calcium concentration leads to faster closing of the channel and hence a reduced calcium influx. Therefore $\varepsilon_\sigma > 0$ corresponds to negative autoregulation and $\varepsilon_\sigma < 0$ to positive feedback on the calcium dynamics (to guarantee that $\tilde{\gamma}_-$ remains positive in Eq. (6), $\varepsilon_\sigma > -1$ is required). Note that the opening rate in Eq. (6) depends, through Eq. (4), on the whole history of previous openings and closings of the channel. The feedback introduced through ε_σ therefore renders the channel state dynamics non-Markovian.

3. Signal-induced information gain

As quantification of the information transfer we adopt the information gained during the interval $[0, t]$ due to the stimulus V [9]. To introduce this measure of information transfer we start from the definition of entropy for a stochastic point process that generates switching events. Denoting by $\Sigma_n(t, t_n, \ldots, t_1, t_0)$ the probability density that a sequence of n events occurs at time $t_1, \ldots, t_n$ between time t_0 and t the entropy of the stochastic process during this time interval can be written as

$$S_{\Delta\tau}(t, t_0) = -\Sigma(t, t_0) \ln \Sigma(t, t_0) - \sum_{n=1}^{\infty} \int_{t_0}^{t} dt_1 \int_{t_1}^{t} dt_2$$
$$\ldots \int_{t_{n-1}}^{t} dt_n \Sigma_n(t, t_n, \ldots, t_0) \ln[\Sigma_n(t, t_n, \ldots, t_0)(\Delta\tau)^n] \tag{7}$$

Eq. (7) is the natural extension of the Shannon entropy to a continuous-time process assuming a finite resolution $\Delta\tau$ with which the continuous-time variable can be measured. This τ-entropy, as well as the difference of such entropies [9], depends on $\Delta\tau$ and is consequently ill-defined (not fully defined by the stochastic process). To reach a uniquely defined measure of information transfer, Goychuk and Hanggi [9], introduced the relative entropy (or information gain) $\mathcal{K}$, the deviation of the process entropy in presence of the stimulus from the entropy value

6

in the absence of stimulus. This information gain is defined using the Kullback-Leibler divergence [16] of the probability density $\overline{\Sigma}_n$ for events occurring in the presence of stimulus from the probability density Σ_n for events occuring in the stationary condition without stimulus. This information measure is independent of the resolution $\Delta\tau$ and is given by

$$
\mathcal{K}(t, t_0) = \Sigma(t, t_0) \ln \left(\frac{\Sigma(t, t_0)}{\overline{\Sigma}(t, t_0)} \right) + \tag{8}
$$
$$
\sum_{n=1}^{\infty} \int_{t_0}^{t} dt_1 \int_{t_1}^{t} dt_2 \cdots \int_{t_{n-1}}^{t} dt_n \Sigma_n(t, t_n, \ldots, t_0) \ln \left(\frac{\Sigma_n(t, t_n, \ldots, t_0)}{\overline{\Sigma}_n(t, t_n, \ldots, t_0)} \right).
$$

Our goal is to evaluate, for the signaling module introduced in Sec. 2, the information gained about the input stimulus $V(t)$ from the knowledge of the output $\sigma(t)$. In the module of Fig 1, given an initial concentration σ_0, to each realisa-
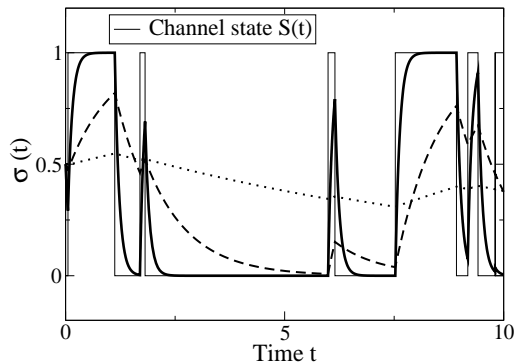

Figure 2: Examples of time courses for the dimensionless calcium concentration $\sigma(t)$ following from Eq. (2) for a given sequence $\vec{t}_n$ of channel state switches. Thick, dashed and dotted lines correspond to $\lambda = 10.0, 1.0$ and 0.1 respectively.

tion, or trajectory, of $S(t)$ corresponds univocally one trajectory of $\sigma(t)$ (see Fig. 2). The trajectories of $\sigma(t)$ can be therefore grouped in four sets determined by the initial and final channel state. Let $\Sigma_{ij}(\vec{t}_n, t)$ be the probability of having a given realisation of $\sigma(t)$ in correspondence of n opening and closing events of the channel occurring between 0 and t (indicated with $\vec{t}_n$ for brevity), with $i, j = \pm$

indicating the initial and final channel state respectively (+ meaning open and
− closed). We introduce then the analogous probability for the unperturbed
channel dynamics $\Sigma_{ij}^0(\vec{t}_n, t)$, i.e. corresponding to have neither stimulus nor
feedback applied and we use $\overline{\Sigma}_{ij}(\vec{t}_n, t)$ for the dynamics including feedback but
not the stimulus. For the unperturbed process the waiting times distributions
(WTDs) in each of the two states are Poissonian, i.e. $\psi_{\pm}^0(\tau) = \gamma_{\pm} e^{-\gamma_{\pm}\tau}$ and
the corresponding survival probabilities $\Psi_{\pm}^0 = \int_t^{\infty} \psi_{\pm}^0(\tau)d\tau = e^{-\gamma_{\pm}\tau}$. It follows
that the probability $\Sigma_{ij}^0(\vec{t}_n, t)$ of a time realisation of $\sigma(t)$ with a sequence $\vec{t}_n$ of
channel switches beginning in state i of the channel and ending in state j is (in
absence of stimulus and feedback):

$$\Sigma_{ij}^0(\vec{t}_n, t) = P_i(\sigma_0)\psi_{-i}^0(\tau_1)\psi_i^0(\tau_2)\cdots\psi_{-j}^0(\tau_n)\Psi_j^0(t - t_n)$$

where $\tau_k = t_k - t_{k-1}$ and $P_i(\sigma_0)$ is the probability that at time $t = 0$ the gate
is in state i and the initial concentration is σ_0.

Eq. (8) applied to our module of Fig. 1 with the notation introduced above,
becomes

$$\mathcal{K}[\sigma|V] = \sum_{i,j=\pm} \left[\delta_{ij}\Sigma_{ii}(t) \ln\left(\frac{\Sigma_{ii}(t)}{\overline{\Sigma}_{ii}(t)}\right) + \right. \tag{9}$$
$$\left. \sum_{n=1}^{\infty} \int_{t_0}^t dt_1 \int_{t_1}^t dt_2 \cdots \int_{t_{n-1}}^t dt_n \Sigma_{ij}(\vec{t}_n, t) \ln\left(\frac{\Sigma_{ij}(\vec{t}_n, t)}{\overline{\Sigma}_{ij}(\vec{t}_n, t)}\right) \right],$$

which, since there is a one-to-one relation between the output trajectories $\sigma(t)$
and the sequences of channel switches, expresses the difference in uncertainty
about the calcium fluctuations in the absence and presence of the input. Note
that Eq. (9) does not include any averaging over inputs; rather, it is the infor-
mation gain for a specific input realization $V(t)$.

4. Evaluation of the information gain

In this section, we carry out a perturbative calculation of the signal-induced
information gain in the presence of feedback. For a switching process without
feedback, the information gain was derived previously in [9]. To evaluate Eq. (9)

8

in our case, the probabilities $\Sigma_{ij}(\vec{t}_n, t)$ and $\overline{\Sigma}_{ij}(\vec{t}_n, t)$ must include the feedback contributions to the desired order in ε_σ.

Perturbation due to either the stimulus or feedback leads to time-dependent rates and consequently modified survival probabilities $\Psi_\pm(t, t_k) = \exp[-\int_{t_k}^t \tilde{\gamma}_\pm dt']$ in each of the two channel states. The ensuing processes are non-Poissonian, since the perturbed rates $\tilde{\gamma}_\pm$ depend on time, with WTDs $\psi_\pm(t, t_k) = -d\Psi_\pm(t, t_k)/dt$ which for small perturbation can be expanded as:

$$\psi_+(\tau, t_k) = \psi_+^0(\tau)\left[1 + \varepsilon_V \psi^V(\tau, t_k) + \varepsilon_V^2 \psi^{V^2}(\tau, t_k) \cdots\right] \qquad (10a)$$

$$\psi_-(\tau, \vec{t}_k) = \psi_-^0(\tau)\left[1 + \varepsilon_\sigma \psi^\sigma(\tau, \vec{t}_k) + \varepsilon_\sigma^2 \psi^{\sigma^2}(\tau, \vec{t}_k) \cdots\right], \qquad (10b)$$

where $\tau = t - t_k$. The first order corrections are:

$$\psi^V(\tau, t_k) = V(t_k + \tau) - \gamma_+ \int_{t_k}^{t_k + \tau} dx V(x) \qquad (11a)$$

$$\psi^\sigma(\tau, \vec{t}_k) = \sigma(t_k + \tau | \vec{t}_k) - \gamma_- \int_{t_k}^{t_k + \tau} dx \sigma(x | \vec{t}_k) \qquad (11b)$$

and the corresponding expressions for higher order corrections can be analogously derived. The appearance of $\vec{t}_k$ in (11b) remarks the dependence of this correction on all the sequence of events up to time t_k. Due to the form of Eqs. (11) and of the analogous higher order corrections, the WTDs in Eqs. (10) remain normalized at each order (the perturbative corrections integrate to zero as can be checked by applying integration by parts on the first order corrections while using Eqs. (11)). As the integral of $\sigma(t)$ in Eq. (11b) can diverge with τ, the validity of the small perturbation expansion for ψ_- has to be examined. The terms of the expansion of the information gain ensuing from the adoption of Eqs. (11) can be proved to be well behaved when averaged over sequences $\vec{t}_n$ (see Appendix, last paragraph). Carrying out the expansion to second order according to the Eqs. (10) and (11) leads to:

$$\Sigma_{ij}(\vec{t}_n, t) \simeq \Sigma_{ij}^0 + \varepsilon_V \Sigma_{ij}^V + \varepsilon_\sigma \Sigma_{ij}^\sigma + \varepsilon_V^2\left(\Sigma_{ij}^{VV} + \Sigma_{ij}^{V^2}\right) +$$

$$+ \varepsilon_\sigma \varepsilon_V \Sigma_{ij}^{\sigma V} + \varepsilon_\sigma^2\left(\Sigma_{ij}^{\sigma\sigma} + \Sigma_{ij}^{\sigma^2}\right) \qquad (12)$$

where to lighten notation we have dropped the dependence on $(\vec{t}_n, t)$ in the terms on the right-hand side. In Eq. (12) the first order contributions to the

expansion for small $\varepsilon_V, \varepsilon_\sigma$(indicated by superscripts V and σ) are obtained by adding to $\Sigma_{ij}^0(\vec{t}_n, t)$ the corrections obtained drawing only one interval in the sequence $\vec{t}_n$ with a WTD perturbed up to the first order correction in Eqs. (10) while the remaining intervals are generated with unperturbed WTDs $\psi_\pm^0(\tau)$. The second order corrections to $\Sigma_{ij}(\vec{t}_n, t)$ are obtained either by drawing two intervals in the sequence $\vec{t}_n$ with WTDs corrected to first order (superscripts $VV, \sigma V, \sigma\sigma$) or by drawing only one interval in $\vec{t}_n$, but with WTDs corrected up to second order (superscripts V^2, σ^2), with the remaining intervals drawn with unperturbed WTDs $\psi_\pm^0(\tau)$.

Considering the probability $\Sigma_{ij}(\vec{t}_n, t) \equiv Q_{ij}(\varepsilon_\sigma, \varepsilon_V)$ as function of the small parameters $\varepsilon_\sigma, \varepsilon_V$, due to the structure of Eq. (7) the information gain is a sum of terms with the following functional form:

$$G_{ij}(\varepsilon_\sigma, \varepsilon_V) = Q_{ij}(\varepsilon_\sigma, \varepsilon_V) \ln \frac{Q_{ij}(\varepsilon_\sigma, \varepsilon_V)}{Q_{ij}(\varepsilon_\sigma, 0)}. \tag{13}$$

A Taylor expansion of $G_{ij}(\varepsilon_\sigma, \varepsilon_V)$ to second order around $\varepsilon_\sigma = \varepsilon_V = 0$ leads to:

$$G_{ij}(\varepsilon_\sigma, \varepsilon_V) = \varepsilon_V \frac{\partial Q_{ij}(\varepsilon_\sigma, \varepsilon_V)}{\partial \varepsilon_V}|_{\varepsilon_\sigma = \varepsilon_V = 0} + \varepsilon_\sigma \varepsilon_V \left[\frac{\partial^2 Q_{ij}(\varepsilon_\sigma, \varepsilon_V)}{\partial \varepsilon_\sigma \partial \varepsilon_V} + \tag{14} \right.$$

$$\left. \frac{\frac{\partial Q_{ij}(\varepsilon_\sigma, \varepsilon_V)}{\partial \varepsilon_\sigma} \frac{\partial Q_{ij}(\varepsilon_\sigma, \varepsilon_V)}{\partial \varepsilon_V}}{Q_{ij}(\varepsilon_\sigma, \varepsilon_V)} \right]_{\varepsilon_\sigma = \varepsilon_V = 0} + \frac{\varepsilon_V^2}{2} \left[\frac{\partial^2 Q_{ij}(\varepsilon_\sigma, \varepsilon_V)}{\partial \varepsilon_V^2} + \frac{\left(\frac{\partial Q_{ij}(\varepsilon_\sigma, \varepsilon_V)}{\partial \varepsilon_V} \right)^2}{Q_{ij}(\varepsilon_\sigma, \varepsilon_V)} \right]_{\varepsilon_\sigma = \varepsilon_V = 0}$$

with all missing second-order terms having coefficient equal to zero when evaluated at $\varepsilon_\sigma = \varepsilon_V = 0$ (for the full formal expansion see Appendix). Replacing the expression of Q_{ij} in terms of Σ_{ij} using Eq. (12), one obtains for the second order expansion of G_{ij} the following terms:

$$G_{ij}(\varepsilon_\sigma, \varepsilon_V) \simeq \varepsilon_V \Sigma_{ij}^V(\vec{t}_n, t) + \varepsilon_\sigma \varepsilon_V \Sigma_{ij}^{\sigma V}(\vec{t}_n, t) +$$

$$+ \varepsilon_V^2 \left[\Sigma_{ij}^{V^2}(\vec{t}_n, t) + \Sigma_{ij}^{VV}(\vec{t}_n, t) + \frac{\Sigma_{ij}^V(\vec{t}_n, t)^2}{2\Sigma_{ij}^0(\vec{t}_n, t)} \right] \tag{15}$$

with the contribution of order ε_σ^2 being exactly zero. As already mentioned, Eqs. (10) imply that the corrections to the unperturbed WTDs integrate to exactly zero over time, so that the WTDs remain normalized. As a consequence $\Sigma_{ij}^0(\vec{t}_n, t)$, when averaged over all possible sequences $\vec{t}_n$ and initial and final

10

channel state, adds up to one, while all the perturbative corrections of each order give a zero overall contribution. It follows therefore that integrating and summing Eq. (15) over all possible paths $\vec{t}_n$ and initial and final state i, j the linear term Σ_{ij}^V, and the second-order terms $\Sigma_{ij}^{\sigma V}$, $\Sigma_{ij}^{V^2}$ and Σ_{ij}^{VV} give zero contribution.

Let us then analyse the remaining 2nd-order term: $A_{ij}^{(2)} \equiv \frac{(\Sigma_{ij}^V)^2}{2\Sigma_{ij}^0}$.

Considering for example the case of initial and final closed state $i = j = -$, we obtain

$$A_{--}^{(2)}(\vec{t}_{n|k}, t) = A_{--}^{(2)}(t, t_n, \cdots t_k \cdots t_1, t_0) = \frac{\Sigma_{--}^V(\vec{t}_{n|k}, t)^2}{2\Sigma_{--}^0(\vec{t}_{n,k}, t)} =$$

$$= \frac{P_-(\sigma_0)}{2} \psi_+^0(\tau_1) \psi_-^0(\tau_2) \cdots \psi^V(\tau_{k+1}, t_k)^2 \cdots \psi_-^0(\tau_n) \Psi_+^0(t - t_n), \quad (16)$$

where $\Psi_{\pm}^0$ are unperturbed survival probabilities for the channel opening/closing events. Further subscript k in $\vec{t}_{n|k}$ indicates the perturbed time interval within the set of n time intervals, with k and n in (16) being even integers for $i = j = -$.

For the distribution of initial conditions $P_\pm(\sigma_0)$, it is natural to assume a stationary solution for channel dynamics with feedback but no stimulus. In this case the initial probabilities are $P_\pm(\sigma_0) = \gamma_\pm/(\gamma_+ + \gamma_-) \pm \Sigma_\sigma$, where a linear correction $\Sigma_\sigma \propto \varepsilon_\sigma$ is added to the stationary solution for the unperturbed dynamics without feedback. It follows that the correction due to feedback in $P_\pm(\sigma_0)$ produces in Eq. (13) a contribution of order $\varepsilon_V^2 \varepsilon_\sigma$, and can therefore be neglected to second order. Moreover, a direct calculation shows that this correction is $\propto \varepsilon_V^2 \varepsilon_\sigma \gamma_+ e^{-t(\gamma_+ + \gamma_-)} V^2(t)$ and therefore vanishes on the channel state switching time-scale. It can consequently be neglected also to third order when evaluating information gain over long time intervals. Averaging over all possible paths, the second order contribution A_{ij} to the information gain is:

$$\mathcal{K}^{(2)}(t) = \sum_{ij,nk} \int d\vec{t}_{n|k} A_{ij}^{(2)}(\vec{t}_{n|k}, t) = \quad (17)$$

$$= \varepsilon_V^2 \int d\omega d\omega' \hat{V}(\omega) \hat{V}(\omega') \frac{\iota \gamma_+ \gamma_- [1 - e^{\iota(\omega + \omega')t}]}{2(\omega + \omega')(\gamma_- + \gamma_+)}$$

where the sum over n runs from 1 to $+\infty$ ($k < n$) and for convenience of

calculation we introduced the Fourier transform $\hat{V}(\omega)$

$$\hat{V}(\omega) = \int_{-\infty}^{\infty} dt e^{-\imath \omega t} V(t) \tag{18}$$

of the stimulus $V(t)$, adopting the character $\imath$ for the imaginary unit. Taking the time derivative and integrating over frequencies gives the rate of information gain:

$$\frac{d\mathcal{K}}{dt} = \frac{\varepsilon_V^2}{2} V^2(t) \frac{\gamma_+ \gamma_-}{\gamma_+ + \gamma_-} \tag{19}$$

which coincides with the result obtained in [9] for a similar signaling module without feedback.

In order to see the contribution of the feedback one then has to expand Eq. (15) to third order. This leads to the additional terms shown in Eq. (A2). After summation, integration (for details see Appendix) we obtain

$$\mathcal{K}^{(3)}(t) = \frac{1}{2} \sum_{ij,nkl} \int d\vec{t}_{n|kl} A_{ij}^{(3)}(\vec{t}_{n|kl}, t) = -\frac{\imath \varepsilon_V^2 \varepsilon_\sigma}{2} \frac{\gamma_+^2 \gamma_-}{(\gamma_+ + \gamma_-)^2} \tag{20}$$

$$\int d\omega d\omega' \frac{\hat{V}(\omega)\hat{V}(\omega')}{\omega + \omega'} \frac{\gamma_+ + \lambda}{\gamma_- + \gamma_+ + \lambda} e^{\imath(\omega+\omega')t}.$$

We differentiate (20) with respect to time, integrate over frequencies and add up to (19) so as to obtain the information gain rate including the feedback contribution:

$$\frac{d\mathcal{K}}{dt} = \frac{\varepsilon_V^2}{2} V^2(t) \frac{\gamma_+ \gamma_-}{\gamma_+ + \gamma_-} \left[1 + \varepsilon_\sigma \frac{\gamma_+}{\gamma_+ + \gamma_-} \frac{\gamma_+ + \lambda}{\gamma_+ + \gamma_- + \lambda} \right], \tag{21}$$

where terms vanishing on time scale $t \gg \max(\lambda^{-1}, \gamma_\pm^{-1})$, which include those carrying the dependence on the initial concentration σ_0, have been neglected. Our calculation was carried out for the case of the input and the feedback acting on the opening and closing rate, respectively. Our approach can obviously be extended to the case of interchanged action of feedback and input on the rates, which would lead to the same result as in Eq. (21) with the simple interchange $\gamma_+ \leftrightarrow \gamma_-$. In the case of both feedback and input affecting the same rate, the leading feedback correction would again be of order $\varepsilon_\sigma \varepsilon_V^2$, but the coefficient would be different. The terms of order e.g. $\varepsilon_\sigma \varepsilon_V$ in the derivation can in this

case emerge from perturbing to second order the same switching event, while in the case we analyzed they can only emerge from perturbing two switching events to first order. Derivation of the correction for this case, as well as of higher-order corrections, is beyond the scope of this paper.

5. Discussion

The information gain rate Eq. (21), when integrated over a time interval $[0, T]$ with $T \gg \max(\lambda^{-1}, \gamma_+^{-1})$, gives the total information obtained about the stimulus by observing the output. It is seen that (to first order in feedback strength and second order in stimulus strength) the information gain depends only on the total power $\int_0^T dt V^2(t)$ of the stimulus, rather than on the spectral distribution of this power. This implies that the information gain rate cannot be optimized by matching the temporal structure of the stimulus with the time scale of the feedback dynamics.

The factor $\frac{\gamma_+ \gamma_-}{\gamma_+ + \gamma_-}$ in Eq. (21) expresses the rate of 'double flip events' (i.e., openings and successive closings of the channel) in the absence of stimulus and feedback. In the case $\varepsilon_\sigma = 0$, the information gain rate is therefore simply proportional to the unperturbed rate of double flip events. [1]To examine if an analogous relation holds generally, we evaluate the mean rate of double flip events in absence of stimulus but in presence of feedback. The mean closing rate is given by $\gamma_-(1 + \varepsilon_\sigma \overline{\sigma}_+)$, where $\overline{\sigma}_+$ is the average concentration when the channel is in the open state. To obtain a result valid to first order in ε_σ, it is sufficient to express $\overline{\sigma}_+$ to 0th order; this conditional average was computed in Ref. [8] and equals

$$\overline{\sigma}_+ = \frac{\gamma_- + \lambda}{\gamma_- + \gamma_+ + \lambda}. \tag{22}$$

---

[1]The true double flip rate in the presence of stimulus differs from $\frac{\gamma_+ \gamma_-}{\gamma_+ + \gamma_-}$ by a correction that is first order in ε_V. In Eq. (21), such a correction would result in a higher-order term $o(\varepsilon_\sigma^3)$, which is beyond the order in which we carried out the expansion.

The mean rate of double flip events in the presence of feedback is then given by

$$R = \gamma_+ \gamma_- (1 + \varepsilon_\sigma \overline{\sigma}_+)/(\gamma_+ + \gamma_- (1 + \varepsilon_\sigma \overline{\sigma}_+)) \tag{23}$$

$$= \frac{\gamma_- \gamma_+}{\gamma_- + \gamma_+} \left[1 + \varepsilon_\sigma \frac{\gamma_+}{\gamma_- + \gamma_+} \frac{\gamma_+ + \lambda}{\gamma_- + \gamma_+ + \lambda} + o(\varepsilon_\sigma^2) \right]$$

and Eq. (21) becomes

$$\frac{dK}{dt} = \frac{\varepsilon_V^2}{2} V^2(t) R. \tag{24}$$

The information gained per double flip event is therefore given only by the power of the input and does not depend on the feedback strength ε_σ or on the kinetic parameters $\gamma_+, \gamma_-, \lambda$.

Before relating this result to findings from the previous literature, we first point out that the definition of feedback in communication / information theory is more restrictive than the definition used in the literature on biochemical signaling. In the biological literature, feedback arises when a product of a signal transduction step influences the upstream components in the pathway. This fits with the general definition of feedback in the early cybernetics literature: "When (this) circularity of action exists between the parts of a dynamic system, feedback may be said to be present" [17]. In communication theory, however, feedback is typically required to act in such a way that it effectively modifies the input of the system. The signaling module analyzed in this work contains in fact a feedback loop in the former sense; functionally, such type of feedback permits to e.g. achieve sensory adaptation to repeated stimuli [14] and can improve the temporal resolution of signaling [8]. In the latter sense (viewed as a communication channel), however, our module cannot be said to have feedback, as the input $V(t)$ is not combined with the output. Rather, the back-coupling implements an autoregulatory loop with the communication channel. The effect of the autoregulatory loop is to give memory to the channel state-switching dynamics: the closing rate depends (through the instantaneous calcium concentration) on the full history of previous channel opening and closings. Our signaling module can therefore be viewed as a non-Markovian point process channel without feedback.

14

Following this clarification of terminology, to put the results of Eqs.(21) and (24) into perspective, we recall a known result from information theory It was proved in [18, 19] (see also [20]), that for any Markov point process the channel capacity per event cannot exceed the capacity of the Poisson process. I.e., the capacity per event is not improved by memory. While we have calculated the information gain rate for one realization of the stimulus (rather than the channel capacity), our finding may be viewed as somewhat analogous. In our case, the channel is non-Markovian, but as the memory decays exponentially in time (akin to a Hawkes process), the system is 'near-Markov'. It is possible, however, that a different result (i.e., information gain affected by memory) would be obtained if we carried out the perturbation expansion to higher orders in ε_σ. For discrete-time (rather than point-process) channels, the influence of memory on channel capacity was recently analyzed in [21, 22].

In [7], an analysis of several prototypical signaling pathways with feedback was carried out under the additive Gaussian noise approximation. Under this restriction (see also Sec. 1), the authors were able to compute the mutual information rate for modules in which nonlinear feedback affected the activation of an upstream component. They concluded that when the feedback was mediated by the final output of the pathway, no improvement of the information transmission was obtained. This is reminiscent of our main finding. In [7], an enhancement of information transmission was obtained only when the feedback was mediated by an intermediate product in the pathway, and not the final output. The module we analyzed in this paper (Fig. 1) falls outside of this class.

In conclusion, we presented a novel perturbative approach that permits to analytically evaluate information transfer through non-Markovian point process channels. We applied this approach to a prototypical signaling module with biochemical feedback and showed that to first order in feedback strength the information gain rate is increased by negative feedback (and decresead by positive feedback). However, this change in information gain rate is fully accounted for by the feedback-induced change in the rate of signaling events (channel

15

opening/closings). To first order in feedback strength, the information gain per signaling event is not affected by feedback.

6. Acknowledgments

Appendix: Expansion of the Information gain

The Taylor expansion of Eq. (13) to second order in $\varepsilon_\sigma, \varepsilon_V$ gives

$$
G_{ij}(\varepsilon_\sigma, \varepsilon_V) = \varepsilon_\sigma \left[\frac{\partial Q_{ij}(\varepsilon_\sigma, \varepsilon_V)}{\partial \varepsilon_\sigma} - \frac{Q_{ij}(\varepsilon_\sigma, \varepsilon_V)}{Q_{ij}(\varepsilon_\sigma, 0)} \frac{\partial Q_{ij}(\varepsilon_\sigma, 0)}{\partial \varepsilon_\sigma} \right]_{\varepsilon_\sigma, \varepsilon_V = 0} \tag{A1}
$$

$$
+ \varepsilon_V \frac{\partial Q_{ij}(\varepsilon_\sigma, \varepsilon_V)}{\partial \varepsilon_V} \Big|_{\varepsilon_\sigma, \varepsilon_V = 0} + \frac{\varepsilon_\sigma^2}{2} \left[\frac{\partial^2 Q_{ij}(\varepsilon_\sigma, \varepsilon_V)}{\partial \varepsilon_\sigma^2} + \left(\frac{\partial Q_{ij}(\varepsilon_\sigma, \varepsilon_V)}{\partial \varepsilon_\sigma} \right)^2 \frac{1}{Q_{ij}(\varepsilon_\sigma, \varepsilon_V)} \right.
$$

$$
+ \frac{Q_{ij}(\varepsilon_\sigma, \varepsilon_V)}{Q_{ij}^2(\varepsilon_\sigma, 0)} \left(\frac{\partial Q_{ij}(\varepsilon_\sigma, 0)}{\partial \varepsilon_\sigma} \right)^2 - \frac{2}{Q_{ij}(\varepsilon_\sigma, 0)} \frac{\partial Q_{ij}(\varepsilon_\sigma, \varepsilon_V)}{\partial \varepsilon_\sigma} \frac{\partial Q_{ij}(\varepsilon_\sigma, 0)}{\partial \varepsilon_\sigma}
$$

$$
\left. - \frac{\partial^2 Q_{ij}(\varepsilon_\sigma, 0)}{\partial \varepsilon_\sigma^2} \frac{Q_{ij}(\varepsilon_\sigma, \varepsilon_V)}{Q_{ij}(\varepsilon_\sigma, 0)} \right]_{\varepsilon_\sigma, \varepsilon_V = 0} + \frac{\varepsilon_V^2}{2} \left[\frac{\partial^2 Q_{ij}(\varepsilon_\sigma, \varepsilon_V)}{\partial \varepsilon_V^2} + \frac{1}{Q_{ij}(\varepsilon_\sigma, \varepsilon_V)} \left(\frac{\partial Q_{ij}(\varepsilon_\sigma, \varepsilon_V)}{\partial \varepsilon_V} \right)^2 \right]_{\varepsilon_\sigma = \varepsilon_V = 0}
$$

$$
+ \varepsilon_\sigma \varepsilon_V \left[\frac{\partial^2 Q_{ij}(\varepsilon_\sigma, \varepsilon_V)}{\partial \varepsilon_\sigma \partial \varepsilon_V} + \frac{1}{Q_{ij}(\varepsilon_\sigma, \varepsilon_V)} \frac{\partial Q_{ij}(\varepsilon_\sigma, \varepsilon_V)}{\partial \varepsilon_\sigma} \frac{\partial Q_{ij}(\varepsilon_\sigma, \varepsilon_V)}{\partial \varepsilon_V} \right]_{\varepsilon_\sigma = \varepsilon_V = 0},
$$

from which it is readily seen that the coefficient of the term $\propto \varepsilon_\sigma$ and $\propto \varepsilon_\sigma^2$ vanish when evaluated at $\varepsilon_\sigma = \varepsilon_V = 0$.

Expanding Eq. (13) to third order, following the same procedure as in Eq. (A1) and converting back Q_{ij} to Σ_{ij} leads to three additional terms:

$$
\frac{\varepsilon_\sigma^2 \varepsilon_V}{3} \left[\Sigma_{ij}^{\sigma\sigma V} + \Sigma_{ij}^{\sigma^2 V} \right] + \frac{\varepsilon_V^2 \varepsilon_\sigma}{6} \left[4 \frac{\Sigma_{ij}^V \Sigma_{ij}^{\sigma V}}{\Sigma_{ij}^0} - \frac{(\Sigma_{ij}^V)^2 \Sigma_{ij}^\sigma}{(\Sigma_{ij}^0)^2} \right] \tag{A2}
$$

$$
+ \frac{\varepsilon_V^3}{6} \left[-\frac{(\Sigma_{ij}^V)^3}{(\Sigma_{ij}^0)^2} + 6\Sigma_{ij}^V \frac{\Sigma_{ij}^{VV} + \Sigma_{ij}^{V^2}}{\Sigma_{ij}^0} \right].
$$

where we have used the next order expansion of (12) i.e.

$$\Sigma_{ij}(\vec{t}_n, t) \simeq \Sigma_{ij}^0 + \varepsilon_V \Sigma_{ij}^V + \varepsilon_\sigma \Sigma_{ij}^\sigma + \varepsilon_V^2 \left(\Sigma_{ij}^{VV} + \Sigma_{ij}^{V^2} \right) + \varepsilon_\sigma \varepsilon_V \Sigma_{ij}^{\sigma V} \quad \text{(A3)}$$

$$+ \varepsilon_\sigma^2 \left(\Sigma_{ij}^{\sigma\sigma} + \Sigma_{ij}^{\sigma^2} \right) + \varepsilon_\sigma^3 \left(\Sigma_{ij}^{\sigma\sigma\sigma} + \Sigma_{ij}^{\sigma\sigma^2} + \Sigma_{ij}^{\sigma^3} \right) + \varepsilon_\sigma^2 \varepsilon_V \left(\Sigma_{ij}^{\sigma\sigma V} + \Sigma_{ij}^{\sigma^2 V} \right)$$

$$\varepsilon_\sigma \varepsilon_V^2 \left(\Sigma_{ij}^{\sigma VV} + \Sigma_{ij}^{\sigma V^2} \right) + \varepsilon_V^3 \left(\Sigma_{ij}^{VVV} + \Sigma\Sigma_{ij}^{VV^2} + \Sigma_{ij}^{V^3} \right)$$

The third term in Eq. (A2) depends only on the stimulus while the first, proportional to $\varepsilon_\sigma^2 \varepsilon_V$, is exactly the same coefficient that would appear as the third order in the expansion in Eq. (12) and therefore gives zero contribution after integration over $\vec{t}_n$ and summation, since perturbative corrections do not affect the normalization of the unperturbed part. Therefore the contribution with the feedback is given only by the term proportional to $\varepsilon_V^2 \varepsilon_\sigma$ in Eq. (A2), whose first element can be written as:

$$\frac{\Sigma_{ij}^V \Sigma_{ij}^{\sigma V}}{\Sigma_{ij}^0} = \Sigma_{ij}^{\sigma VV} + \Sigma_{ij}^\sigma \frac{\left(\Sigma_{ij}^V \right)^2}{\Sigma_{ij}^0} = \left[\Sigma_{ij}^{\sigma VV} + \Sigma_{ij}^{\sigma V^2} \right] + \Sigma_{ij}^\sigma \frac{\left(\Sigma_{ij}^V \right)^2}{\Sigma_{ij}^0} - \Sigma_{ij}^{\sigma V^2} \quad \text{(A4)}$$

The term between brackets has been so rearranged in order to show that after summation and integration, it gives zero contribution (since it corresponds to the term of order $\varepsilon_V^2 \varepsilon_\sigma$ in the the expansion of Eq. (12), i.e. first term in last row in Eq. (A3)). The last term in (A4) gives a contribution that vanishes on the time scale of the channel opening/closing dynamics. The remaining term combines with the second element in the feedback contribution in (A2) taking an overall $1/2$ coefficient. The expression for this term, evaluated between initial and final closed states, is:

$$A_{--}^{(3)}(\vec{t}_{n|kl}, t) = \Sigma_{--}^\sigma \left(\frac{\Sigma_{--}^V}{\Sigma_{--}^0} \right)^2 = \quad \text{(A5)}$$

$$= P_-(\sigma_0) \psi_+^0(\tau_1) \psi_-^0(\tau_2) \cdots \psi^V(\tau_{k+1}, t_k)^2 \cdots \psi^\sigma(\tau_{l+1}|\vec{t}_l) \cdots \psi_-^0(\tau_n) \Psi_+^0(t - t_n)$$

where further subscripts k and l in $\vec{t}_{n|k,l}$ indicate that the closing and the opening events affected by the stimulus and the feedback are respectively the kth and the lth of the n event occurring between 0 and t. $\psi^\sigma(\tau_{l+1}|\vec{t}_l)$ is the first order term given by Eq. (11) and carries the dependence on the history of the process

prior to t_l (i.e. the all sequence $\vec{t_l}$ of events up to t_l), but, due to the exponential form of the WTDs, does not affect the convolution structure in (A5).

Averaging over all possible paths gives:

$$\mathcal{K}(t) = \frac{1}{2} \sum_{ij,nkl} \int d\vec{t}_{n|kl} A_{ij}^{(3)}(\vec{t}_{n|kl}, t) = -\frac{\imath \varepsilon_V^2 \varepsilon_\sigma}{2} \frac{\gamma_+^2 \gamma_-}{(\gamma_+ + \gamma_-)^2} \qquad (A6)$$

$$\int d\omega d\omega' \frac{\hat{V}(\omega)\hat{V}(\omega')}{\omega + \omega'} \frac{\gamma_+ + \lambda}{\gamma_- + \gamma_+ + \lambda} e^{\imath(\omega+\omega')t},$$

plus terms that either do not depend on t or vanish for time $t \ll \max(\lambda^{-1}, \gamma_+^{-1})$ exceeding the channel state switching time scales.

Finally a remark on the validity of the perturbative approach we introduced. This approach is based on a direct perturbation of the switching rates and therefore of the WTDs, due to both input and feedback. The only delicate point in following this procedure is the integral of $\sigma(x|\vec{t})$ in Eq. (11b), which can diverge with τ. It can be verified, however that this correction is well-behaved after summation over all stochastic paths and integration. In fact we can replace the contribution of this integral to the first order correction in Eq. (11b) with the term $-\gamma_- \psi_-^0(\tau)\tau$, which is the upper limit to the correction due to the integral between 0 and τ according to our perturbative prescription. The Laplace transform of this correction is

$$\mathcal{L}[-\gamma_-^2 t e^{-\gamma_- t}] = -\int_0^\infty dt \gamma_-^2 t e^{-\gamma_- t} e^{-st} = \left(\frac{\gamma_-}{\gamma_- + s}\right)^2 = \hat{\psi}_-^0(s)^2 \qquad (A6)$$

where

$$\hat{\psi}_-^0(s) = \mathcal{L}[\gamma_- e^{-\gamma_- t}] = \int_0^\infty dt \gamma_- e^{-\gamma_- t} e^{-st} = \frac{\gamma_-}{\gamma_- + s}. \qquad (A7)$$

We can now replace Eq. (A6) in the transition amplitude. Due to the convolution structure of $\Sigma_{ij}(\tau, t)$ in terms of WTDs, this replacement adds just a multiplicative factor in Laplace space. The final result is (for $i = j = -$)

$$\Sigma_{--}(s) = \Sigma_{--}^0(s)(1 + \varepsilon_\sigma \hat{\psi}_-^0(s)), \qquad (A8)$$

where we have included only the contribution of the integral and not the first term in Eq. (11b) since the latter is trivially well-behaved. Due to Tauberian

theorem, one can deduce the time asymptotic limit by analyising the $s \to 0$ behavior in Eq. (A8). Since $\hat{\psi}_-(s) \to 1$ for $s \to 0$, such correction obtained after averaging over trajectories is always limited, non-divergent and therefore well-behaved for $t \to \infty$.

References

References

[1] H.C. Berg, E. M. Purcell, *Biophys. J.* **20**, 193 (1977).

[2] H. Mao, P. S. Cremer, M. D. Manson, Proc. Natl. Acad. Sci. U.S.A. **100**, 5449 (2003).

[3] J. E. Ladbury and S. T. Arold, Trends Biochem. Sci. **37**, 172 (2012).

[4] F. Tostevin and P. R. ten Wolde, Phys Rev. Lett. **102** ,21801 (2009).

[5] R. Cheong , A. Rhee, J.W. Chiaochun, I. Nemenman and A. Levchenko, Science **334**, 354-358 (2011).

[6] G. Tkačik, C.G. Callan, W. Bialek, Physical Review E **78**, 011910 (2008).

[7] W.H. de Ronde, F. Tostevin, and P. R. ten Wolde Phys. Rev. E **82**, 031914.

[8] M. Gopalakrishan, P. Borowski, F. Julicher and M. Zapotocky, Phys. Rev. E **76**, 021904 (2007).

[9] I. Goychuk and P. Hanggi, Phys. Rev. E **61**, 4272 (2000).

[10] Matthews, Hugh R., and Reisert, J. Current Opinion in Neurobiology **13**, 469-475 (2003).

[11] A. B. Parekh, The Journal of physiology, **586**, 3043 (2008).

[12] B.Z. Peterson, C.D. DeMaria and D. Yue, Neuron **22**, 549 (1999).

[13] J. Bradley, J. Reisert,and S. Frings, Current opinion in neurobiology **15**, 343 (2005).

[14] J. Reidl, P. Borowski, A. Sensse, J. Starke, M. Zapotocky, M. Eiswirth: Biophysical Journal **90**, 1147-1155 (2006)

[15] N. Rosenfeld, MB Elowitz, and U. Alon, Journal of molecular biology **23**, 785 (2002).

[16] S. Kullback and R.A. Leibler, Ann. Math. Statist. **22**, 79-86 (1951).

[17] W.R. Ashby, *An introduction to cybernetics*, Chapman & Hail Ltd, London (1956).

[18] M. H. A. Davis, IEEE Trans. Info. Th., **IT-26**, 710-715 (1980).

[19] Y.M. Kabanov, Theory of Probability and Its Applications, **23**, 143-147 (1978).

[20] D. H. Johnson, Journal of Comp. Neurosci. **3**, 275-299 (1999).

[21] L. Kostal and P. Lansky, Phys. Rev. E **81**, 050901(R) (2010).

[22] L. Kostal, Phys. Rev. E **82**, 026115 (2010).