Łochów, 23rd–27th September 2014

CHANNEL CAPACITY IN NOISY BIOCHEMICAL SIGNALLING NETWORKS

Agnieszka Dziekańska

Systems Biology Ireland, University College Dublin
Belfield, Dublin 4, Ireland,
[1]aggie.dziekanska@gmail.com

ABSTRACT

In this paper we consider a simple model of a biochemical signalling cascade as a communication channel. We apply the information theory approach to measure the input-output information relationship of such a channel, and to investigate how fluctuations in the level of network components affect its information processing capacity. The results indicate that information theory is a promising approach to analyse information processing of biochemical communication channels based on the amount of inputs they can distinguish in a certain level of intracellular fluctuations.

INTRODUCTION

Biochemical signalling network is computable representative of a complex biological system of intracellular protein-protein interactions. It coordinates cellular communication with its microenvironment. In this network, extracellular incoming signals are transduced from the cell membrane to the nucleus by a cascade of chemical reactions. Based on the information processed by the signalling network, cell can assume in response a different fate and modulate its phenotype.

Information theory (IT) is a branch of applied mathematics that focuses on the analysis of how communication takes place in an information channel [10]. The notion of Shannon's communication channel can be applied in a wide spectrum of communication forms, such as computer science, telecommunications [2, 6] as well as in molecular and systems biology [5, 9]. The information is quantified by bits, which is a basic unit in the field of computer science. A bit can have only two values of information, either 0 or 1. In a biological context, a single bit can be defined as the ability of a downstream process of a network to distinguish between two levels of incoming signals: high or low.

Biochemical signalling networks are noisy systems. In a noisy system the response to the input is relayed nondeterministically. Nonetheless, noisy information transduction can still be a reliable form of communication in biochemical systems [4, 8]. The source of noise in biochemical networks stems from chemically interacting molecules. Many important cellular processes, such as transcription, cellular respiration and apoptosis require collisions and bindings of discrete molecules. The distances between molecules within a cell may vary due to their location, and as a result the timing of their reaction varies as well. The magnitude of these fluctuations increases significantly when the copy number of the molecules is low. Low copy-number and diffusive effects result in biochemical reactions occurring randomly. Such noise may be sufficient to vary the gene expression within the cell, and to create switching between one cell fate and another [7]. Transduction capacity of a single signalling pathway is also very limited due to intracellular noise.

Individual signalling pathways are able to process the amount of information that allows only for binary cellular decisions [3].

In this paper we investigate the impact of variability of network components on the capability of signalling networks to process the information. Section 2 explains how to apply the IT approach on biochemical channel analysis, and Section 3 introduces the case study of it. Section 4 and Section 5 demonstrate our results and conclusions. The approach presented in this paper was inspired by the analysis of telecommunication channels and applies to biochemical channels as well.

INFORMATION THEORY IN BIOLOGICAL CONTEXT

Three terms are needed to fully specify the biochemical channel: input alphabet, output alphabet and a matrix of transduction probabilities. In other words one needs to define a set of possible discrete input and output concentrations as well as a scheme of probabilities of particular responses for individual input concentrations.

Scheme of probabilities of responses can be found by sampling a histogram of possible output concentrations for each input concentration in a channel. The marginal probabilities of each output distribution can be used to find the entropy of a channel. Entropy (H) defines the ability to process information in a theoretical, noiseless channel. In the case of a biochemical signalling network, it is a channel where no intracellular noise would ever appear. Entropy of such a channel is defined as $H(X) = -\sum_{k=1}^{n} p(X = x_i) \log_2 p(X = x_i)$. Here, $p(X = x_i)$ are marginal probabilities found for all n histograms of output concentrations. Mutual information (MI) is defined as $MI = H(Y) - H(Y|X)$. It describes the ability to process information in a channel affected by intracellular fluctuations. $H(Y|X)$ is the conditional entropy of the output (Y) when the input (X) is known [4]. MI of a noisy channel can be interpreted as the amount of information that theoretically could be processed correctly by the channel less the amount of information mistaken due to the noise. When increasing intracellular fluctuations we increase the part of information that is mistaken. In ideal conditions, where no noise ever appears no information would be mistaken. In such conditions MI would be maximised by entropy.

The amount of inputs distinguishable by the channel equals 2^C. Here, C is the channel capacity, a quantity that maximises mutual information when all inputs appear with equal probability.

CASE STUDY

As the example of a biochemical communication channel we selected a G-protein (guanosine nucleotide-binding protein) pathway. This pathway is a signalling cascade that governs many cellular functions, such as mitogenesis, cytoskeletal organisation and nuclear transport. Activated mutants of GTPases have been found in 10-20 % of all human tumours.

Fig. 1 presents the model that consists of two protein species, g_1 and g_2. The former is phosphorylated in the presence of a ligand, i.e. the input signal (g_{IN}). Phosphorylation of the latter is nonlinearly activated by g_{1p} – the active (phosphorylated) form of g_1. De–phosphorylation is catalysed by phosphatases, which are not explicitly included in the model. All reactions are modelled using Michaelis-Menten kinetics which is a basic approximation of enzyme-catalysed reactions. In accordance with this kinetics, v_i parameter represents the maximum rate of i-th reaction achieved by the system. The constant K_i is the substrate concentration at which the rate of reaction is half of v_i. The Michaelis-Menten kinetics is frequently used to simulate biochemical reactions involving a single substrate. The basic assumption of this model is that the concentration of the enzyme is much lower than the concentration of the substrate.

Two levels of a signalling cascade can be linked by the nonlinear activation, which takes the form:

$$\text{nonlinear activation} = \frac{k_{13} + A_{13}[\text{Substrate}]}{k_{13} + [\text{Substrate}]}. \tag{1}$$

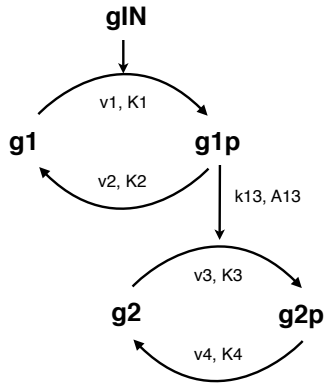

Figure 1. A model of a generic two-tier G-protein cascade without feedback regulation

Here, k_{13} is a reaction rate coefficient, and A_{13} is the activation coefficient. If A_{13} is greater than 0, the formula represents activation. Otherwise, the formula represents inhibition.

Additionally, both positive and negative feedback loops may be included in the model. This topological extension allows to generate various dynamic properties such as oscillations and bistability [11]. Therefore, despite its simplicity the model may also reflect more complex behaviours of large biochemical networks such as a MAPK (Mitogen-activated protein kinase) network.

The system presented in Fig. 1 can be described by the following set of ordinary differential equations (ODE):

$$\begin{aligned} \frac{\mathrm{d}\, g_{1p}}{\mathrm{d}\, t} &= v_1 \frac{g_1}{g_1 + K_1} g_{IN} - v_2 \frac{g_{1p}}{g_{1p} + K_2}, \\ \frac{\mathrm{d}\, g_{2p}}{\mathrm{d}\, t} &= v_3 \frac{g_2}{g_2 + K_3} \frac{k_{13} + A_{13} g_{1p}}{k_{13} + g1p} - v_4 \frac{g_{2p}}{g_{2p} + K_4}. \end{aligned} \tag{2}$$

Using the analytical steady-state solution, the dose response curve of the system presented in Eq.(2) was calculated. It represents the functional relationship between the single substrate concentration (input) and the steady-state concentration of response of the system (output). It can be both simulated numerically and measured experimentally by various techniques, such as Western blot [3].

The dose response curve of system modelled by Eq. (2) forms a sigmoidal curve. A good fit was found [1] utilising a Hill function (Fig. 2), which was introduced as a representative of the dose response curve for all investigations. The Hill function is defined as:

$$f(g_{IN}) = R_{max} \frac{g_{IN}^h}{x_{50}^h + g_{IN}^h} + \text{basal}, \tag{3}$$

where, R_{max} relates to the maximum response concentration, x_{50} is the argument for which the function assumes half of its maximum response concentration and h is a coefficient that specifies the steepness of the curve. Additionally the basal response level was introduced.

To represent variability in the GTPase system the total amount of components $g_{1\text{tot}}$ and $g_{2\text{tot}}$ was assumed to vary according to a lognormal distribution. The parameter values of the lognormal distributions were set to experimentally measured mammalian protein expression noise [1]. We

[1] Fitting curve was obtained using *Mathematica*® software

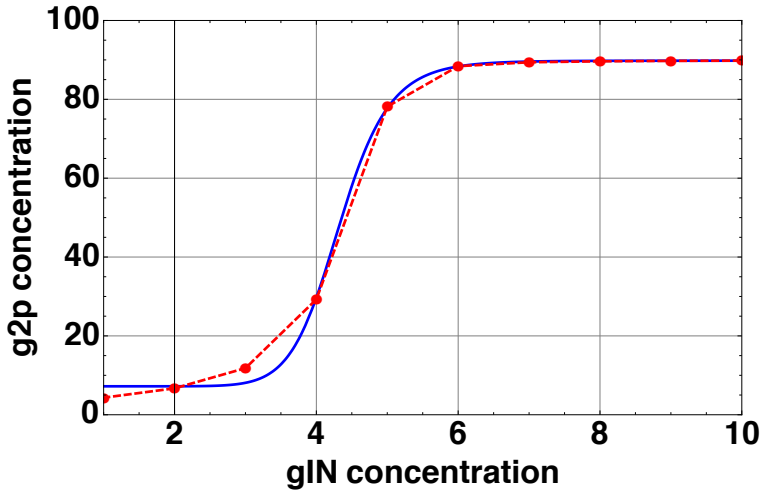

Figure 2. The fitting Hill function to the dose response function. Dashed red line: The dose response function of GTPase system (Fig. 1) evaluated at 10 discrete gIN values with the parameter set $A_{13} = 5$, $k_{13} = 10$, $v_1 = 1.8$, $v_2 = 9.5$, $v_3 = 0.8$, $v_4 = 2.6$, $K_1 = 9$, $K_2 = 3.7$, $K_3 = 5.4$, $K_4 = 7.2$. Solid blue line: fitted Hill function (Eq.(3)) evaluated with the fitting parameter set $h = 12.3$, $x_{50} = 4.3$, $R_{\max} = 82.6$, basal = 7.2.

chose the mean values of the lognormal distributions to be 1, which may correspond for example to 1 nM. We also chose the arbitrary standard deviation (std) of each lognormal distribution by finding the distribution parameters μ and σ that result in a distribution with a desired standard deviation.

We estimated the total response variability by finding its one–standard–deviation boundries within a population. Three dose response functions were computed assuming $g_{1\text{tot}}$ and $g_{2\text{tot}}$ as $g_{1,2tot}$=lognormal mean value, then $g_{1,2tot}$ = lognormal mean value + std and $g_{1,2tot}$=lognormal mean value − std respectively.

In order to represent different cases of variability by Hill fitting function, the standard deviation of Hill parameter x_{50} was estimated. A very crude approximation assumes that the distribution of x_{50} follows the same distribution as g_{1tot} and g_{2tot}. As a consequence, one standard deviation of $g_{1,2tot}$ corresponds to one standard deviation of x_{50}. In other words, the value of x_{50} parameter of the response function with $g_{1,2tot} \pm \text{std}g_{1,2tot}$ corresponds to $x_{50} \pm \text{std}x_{50}$.

Standard deviations of x_{50} Hill parameters were estimated for different examples of system's response variability. If the ergodic hypothesis was assumed, these results can be applied not only to model the protein variability of a population of cells, but also the protein variability of a single cell over a long period of time.

RESULTS

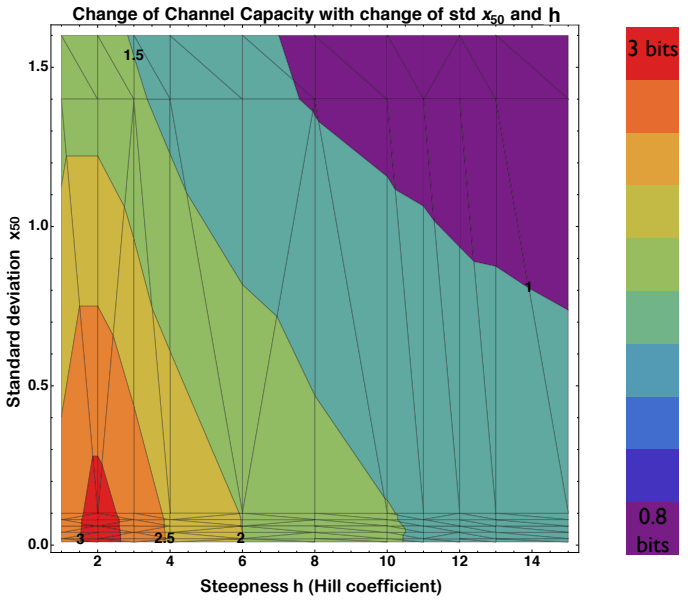

Figure 3. Heat map of values of channel capacity of a generic two-tier G-protein cascade with change of standard deviation of x_{50} parameter and Hill coefficient.

Fig. 3 illustrates the results of computations of the channel capacity of a G-protein pathway (Fig. 1) as a heat map. The values on the x-axis correspond to the values of Hill coefficient (h) of the Hill function (Eq. (3)). The changes of the Hill coefficient correspond to changes of steepness of the dose response function. The higher the value of a Hill coefficient, the steeper the dose response function becomes. Looking from the left to the right side of the heat-map one observes the changes of the channel's capacity with the changes of the dynamics of the dose response function. The left side of the heat map corresponds to channels with nearly linear response, the right side corresponds to channels with switch response.

The values on the y-axis correspond to the standard deviation values of the parameter x_{50} of the Hill function. Standard deviations of x_{50} were sampled from the lognormal distribution with a fixed mean value. The standard deviation values change from std $x_{50} = 0.02$ up to std $x_{50} = 1.5$. Looking from bottom to the top side of the heat map one observes the changes in the channel's variability. The bottom side of the heat map corresponds to the channels with very low amount of fluctuations, the top side of the heat map corresponds to channels with large amount of fluctuations.

Colours on the heat map relate to values of channel capacity obtained for a range of channels with different properties explained above. A channel with $h = 2$ and std $x_{50} = 1.5$ achieves the same channel capacity as a channel with $h = 10$ and std $x_{50} = 0.02$. In other words, the channel with mild response function and large value of intracellular variability can process the same amount of information as a switch channel with a significantly lower value of intracellular variability. Consequently, the channel with mild response can process more information than the switch channel when they contain equal levels of intracellular variability.

CONCLUSIONS

Information transduction capacity is limited to the value of channel capacity $C \sim 1$bit [3]. One bit of information transduction capacity means the output can distinguish only between $2^1 = 2$ concentrations of input. An individual communication channel processes a low amount of information about the signal intensity, however from the point of view of decision making this amount is optimal to make binary decisions. To optimise the decision making process a system should only distinguish between two input signal levels: such as high and low level. Our analysis indicates that many signalling channels achieve this level of channel capacity, however the channel characteristics depends on the level of systemic fluctuations in order to keep channel capacity at the level of $C = 1$ bit. The channel with a switch-like response function and a small threshold variability has the same channel capacity as the channel with a nearly linear response and large threshold variability.

One of the main challenges in the application of the IT framework is to simulate channel capacity of real biochemical networks. Information transduction pathway can be identified for any biochemical network. By choosing two arbitrary nodes of the network one can assign the input and the output. The shape of the input-output relationship between these two nodes determines the information transduction capacity of such a channel. When a node of the network that constitutes the channel is perturbed (by means of chemical inhibition, for example) the input-output characteristics of the channel changes and so does the capacity as measured by IT. This procedure is equivalent to performing sensitivity analysis of channel capacity. Nodes whose perturbations do not translate to significant changes in C do not play an important role in information transduction across the network. The advantage of indices provided by IT over classic sensitivity analysis is that the former take into account intracellular fluctuations that affect the channel. However, such a detailed analysis can be performed only in the presence of high quality mathematical models of signalling networks. Lack of such models poses an obstacle in current applications of IT.

ACKNOWLEDGEMENTS

I wish to express my gratitude to Systems Biology Ireland and Science Foundation Ireland for providing the research environment and the funding which allowed me to complete this research.

REFERENCES

[1] M. R. Birtwistle, A. von Kriegsheim, M. Dobrzyński, B. N. Kholodenko, and W. Kolch: *Mammalian protein expression noise: scaling principles and the implications for knockdown experiments.*, Mol. Biosyst. **8** (2012), 3068–76.

[2] C. Burke: *History of information science.*, vol. 41, Wiley Subscription Services, Inc., A Wiley Company, 2007.

[3] R. Cheong, A. Rhee, I. Nemenman, and A. Levchenko: *Information Transduction Capacity of Noisy Biochemical Signaling Networks.*, Science **334** (2011), 354–358.

[4] T. M. Cover and J. A. Thomas: *Elements of Information Theory.*, 2nd, John Wiley and Sons, 2006.

[5] A. M. Honegger, T. J. Dull, S. Felder, E. van Obberghen, F. Bellot, and D. Szapary et al.: *Point mutation at the ATP binding site of EGF receptor abolishes protein-tyrosine kinase activity and alters cellular routing.*, Cell **51** (1987), 199–209.

[6] S. Lloyd and J. E. Slotine: *Information theoretic tools for stable adaptation and learning.*, International Journal of Adaptive Control and Signal Processing **10** (1996), 499–530.

[7] J. M. Raser and E. K. O'Shea: *Noise in Gene Expression: Origins, Consequences, and Control.*, Science **309** (2005), 2010–2013.

[8] T. D. Schneider: *Claude Shannon: Biologist.*, IEEE Engineering in Medicine and Biology Magazine **25** (2006), 30–33.

[9] R. K. Schultzaberger, L. R. Roberts, I. G. Lyakhov, I. A. Sidorov, A. G. Stephen, R. J. Fisher, and T. D. Schneider: *Correlation between binding rate constants and individual information of E. coli Fis binding sites.*, Nucleic Acids Research **35** (2007), 5275–5283.

[10] C. E. Shannon: *A mathematical theory of communication.*, vol. 27, 1948.

[11] M. A. Tsyganov, W. Kolch, and B. N. Kholodenko: *The topology design principles that determine the spatiotemporal dynamics of G-protein cascades.*, Molecular BioSystems **8** (2012), 730–743.