Multiplexing Biochemical Signals

This thesis was reviewed by:
Prof. dr. F. MacKintosh
Prof. dr. P. Swain
Prof. dr. G. Tkačik
dr. G. Stephens
dr. F. Bruggeman

The research described in this thesis was performed at the *FOM Institute AMOLF*, Amsterdam, The Netherlands. This work is part of the research programme of the *Foundation for Fundamental Research on Matter* (FOM), which is financially supported by the *Netherlands Organization for Fundamental Research* (NWO).

© W. H. de Ronde, 2012
Cover illustration: "Chaotisch patroon V - Lukas van de Vrande, 2011"

ISBN: 978-90-77209-81-3
Nederlandse titel: Het multiplexen van biochemische signalen
A digital version of this thesis is available at

www.ubvu.nl/dissertations and www.amolf.nl
Printed copies can be obtained by request via library@amolf.nl.

Printed by Ponsen & Looijen BV, The Netherlands.

VRIJE UNIVERSITEIT

Multiplexing Biochemical Signals

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad Doctor aan
de Vrije Universiteit Amsterdam,
op gezag van de rector magnificus
prof.dr. L.M. Bouter,
in het openbaar te verdedigen
ten overstaan van de promotiecommissie
van de faculteit der Exacte Wetenschappen
op maandag 7 mei 2012 om 13.45 uur
in de aula van de universiteit,
De Boelelaan 1105

door

Wiet Hendrik de Ronde

geboren te Singapore, Singapore

promotor: prof.dr. P.R. ten Wolde

C e a s el e s s l y the water flows
and yet
the water is never the same

(vrij naar) Kamo no Chōmei

"I almost wish I hadn't gone down the rabbit hole
— and yet — and yet
it's rather curious, you know, this sort of life"

Alice

Contents

1 Introduction 1
- 1.1 Cellular communication . . . 1
  - 1.1.1 Signaling cascades . . . 2
- 1.2 Information theory . . . 4
  - 1.2.1 Mutual information . . . 6
- 1.3 Linear-Noise Approximation . . . 17
- 1.4 Numerical optimization . . . 21
  - 1.4.1 Simulated Annealing . . . 22
  - 1.4.2 Numerical Evolution . . . 22
- 1.5 Scope of this thesis . . . 23

2 Effect of feedback on the fidelity of information transmission of time-varying signals 25
- 2.1 Introduction . . . 26
- 2.2 Methods . . . 28
- 2.3 Results . . . 31
  - 2.3.1 The simple cascade . . . 32
  - 2.3.2 Autoregulation . . . 34
  - 2.3.3 Feedback . . . 38
- 2.4 Discussion . . . 44
- 2.5 Acknowledgements . . . 48
- 2.A Supplementary Information . . . 49
  - 2.A.1 Gillespie Simulations . . . 49
  - 2.A.2 Linear cascades . . . 50
  - 2.A.3 Autoregulation . . . 52
  - 2.A.4 Autoregulation by V . . . 52
  - 2.A.5 Feedback . . . 54
  - 2.A.6 Comments on Fig. 2.7c . . . 58

3 Information transmission in networks with feed-forward loops or diamond motifs 59
- 3.1 Introduction . . . 60
- 3.2 Methods . . . 61
- 3.3 Results . . . 63

3.3.1 Simple Cascades . . . 63
3.3.2 The feed-forward motif . . . 63
3.3.3 Multimerization . . . 73
3.3.4 Diamond motif . . . 75
3.4 Discussion . . . 77
3.5 Acknowledgements . . . 80
3.A Supplementary Information . . . 81
3.A.1 Simple cascades . . . 81
3.A.2 OR coherent feed-forward . . . 82
3.A.3 AND coherent feed-forward . . . 85
3.A.4 Comparison of the coherent feed-forward AND and OR motifs . . . 88
3.A.5 Incoherent feed-forward motif . . . 90
3.A.6 Multimerization . . . 93
3.A.7 Numerical validation . . . 96
3.A.8 Influence of the phase . . . 97

4 The Berg-Purcell limit revisited 101
4.1 Introduction . . . 102
4.2 Theory . . . 104
4.3 Numerical Results . . . 108
4.4 Validity assumption under biological conditions . . . 112
4.5 A simple coarse-grained model . . . 114
4.6 Discussion . . . 115
4.A Supplementary Information . . . 117
4.A.1 The correlation function . . . 117
4.A.2 Derivation of Eq. 4.13 . . . 119

5 Protein Logic 123
5.1 Introduction . . . 124
5.2 Methods . . . 125
5.3 Results . . . 126
5.3.1 Functions accessible by parameter variation . . . 127
5.3.2 Functions accessible by recombination . . . 130
5.4 Discussion . . . 133
5.5 Acknowledgements . . . 135
5.A Supplementary Information . . . 136
5.A.1 Optimization details . . . 136
5.A.2 Formal proof for a XOR-gate for receptor Q_{UW} . . . 138
5.A.3 Parameter sensitivity . . . 140
5.A.4 Figure 5.3b is exhaustive . . . 141

6 Reliability of frequency- and amplitude-decoding in gene regulation 145
6.1 Introduction . . . 146
6.2 Model . . . 147
6.3 Results . . . 148
6.4 Acknowledgements . . . 153

7 Amplitude multiplexing of biochemical signals 155
7.1 Introduction . . . 156
7.2 Model . . . 156
7.3 Multiplexing . . . 159
7.4 Acknowledgements . . . 162

8 Multiplexing oscillatory biochemical signals 163
8.1 Introduction . . . 164
8.2 The model . . . 166
8.2.1 Encoding . . . 167
8.2.2 Decoding V^P to X_1, X_2 . . . 167
8.3 Multiplexing . . . 171
8.4 Experimental observations . . . 174
8.5 Discussion . . . 176
8.A Supplementary Information . . . 179
8.A.1 Encoding . . . 179
8.A.2 Linear Approximation . . . 180
8.A.3 Decoding . . . 182
8.A.4 The conversion of W^P to X_1 . . . 188
8.A.5 Numerical optimization . . . 191

References 193

Summary 209

Samenvatting 213

x

Chapter 1

Introduction

1.1 Cellular communication

The basic building block of all living organisms is the cell. Indeed, a single cell can be a living organism by itself, like the unicellular bacteria from the kingdom of the prokaryotes. Next to these unicellular organisms, multicellular organisms exist. In these organisms multiple cells group, act and communicate in order for the organism to survive. The realization that cells are the fundamental units dates back to the German biologists Theodor Schwann and Matthias Schleiden, who in 1839, formulated a theory which posits that the cell, both in animals and plants, is the fundamental unit of life [1].

Since then more than a 170 years have passed in which our knowledge of these cells has increased dramatically. One important realization is that cells, although being coined the building block of life, themselves consist of many other smaller structures. The main difference between these structures — e.g. membranes, the nucleus, DNA, motors — and the cell is that a cell is capable of reproduction, while the individual building blocks themselves are not. A cell therefore is alive, while its building blocks are lifeless — a fascinating observation on its own. The building blocks themselves consist of molecules like proteins, lipids, receptors and amino-acids, which are the essential lego-parts of all cellular structures and thus the cell.

These smaller building cellular structures are essential for any cell to survive. If we focus on a single cell, for example a bacterium, we observe that it has to overcome many challenges in its struggle for survival. Only looking at the process of cell-division, it is directly clear that the cells faces challenges such as: When does the cell "know" that division should take place? How does the cell spatially divide itself in equal halves, and as important, ensure that essential proteins and organelles are equally distributed over each half? Where and why do structures form that were not present before, like the FtsZ-ring in bacterial cell-division?

Not only for cell division these kinds of questions arise. Bacteria (e.g. *E. coli*) are capable of detecting sugar concentrations in their environment and even move to locations with a larger abundance of sugar, a process called chemotaxis [2]. This is a remarkable achievement,

recognizing the limited size of the bacterium. The same bacterium is capable of surviving on many different sugar sources [3], including glucose and lactose, but in general only the enzyme for the conversion of glucose to energy (ATP) is present. This means that if only lactose is available, the bacterium has to change its internal metabolism path, e.g. produce new enzymes that can act on lactose. Another example of sensing is provided by yeast (*S. cerevisiae*) a-cells, which secretes small pheromone molecules (a-factor), to mark their presence to yeast α-cells, which eventually leads to mating of these two cells [4].

Clearly many processes continuously take place in cells, a small list based upon the above examples could include: the sensing of the environment by membrane proteins (receptors), initiation or termination of gene expression by proteins (transcription factors), activation of motors (myosin, kinesin), or the changes in the cellular structure by protein (actin, dynein). All of these processes are vital for the cell in its order to survive.

Many, if not all, of these processes require some form of communication or signal transfer between the cell and its environment, or between processes entirely within the cell. Considering the above examples, here sugar levels lead to cell motility, sensing a or α factor leads to mating, and the start of division leads to an internal re-organization of cellular components. Cells have developed special networks that essentially have this task: transmitting signals from the extracellular to the intracellular environment or transmitting signals within the cell. These networks are referred to as signaling cascades (or signal transduction cascades) and these cascades are the main topic of this thesis.

1.1.1 Signaling cascades

The goal of a signaling cascade is to transmit a signal, through multiple intermediate steps to a response. The concept of passing a signal around through multiple steps leading to a response is very well mimicked in a very famous children's game: *the telephone game*.

In the telephone game a group of children sit in half a circle. At one end, a child gets a message — in general a complicated or hilarious sentence — and has to whisper this message to his or her neighbor, but such that the other children can not hear it. This process continues until the last child has to inform everyone what the original message was. Of course, the message is corrupted (one could argue that this is the goal of the game), since at every step in the cascade, there is a probability that the message is altered. However, beforehand, it is unclear what the final message precisely will be. Probably the final message will have a resemblance to the original input, but it will have changed. As importantly, repeating the same game with different children will lead to a different final message, but this final message will also be different from the final message of the first group of children. This variability in the final message, originating in the random changes of the message at each step of the cascade, is intrinsically present in every signaling cascade in the cell.

In living cells there are many of these signaling cascades. And, as important, many processes in cells are stochastic. The diffusion of molecules both extracellular and intracellular is a random process, but also the binding, unbinding, activation or deactivation dynamics of proteins are random. These cellular signaling cascades therefore suffer from the same problem as the telephone game; in every step there is a probability that the message is corrupted. The simple example of the children's game is elegantly paraphrased by the data-processing

inequality [5]. This inequality loosely states that

> *"the information between input and output in a sequence of steps can only go down, but never go up."*

Therefore, signal processing in cells is a complicated process, where, due to the stochasticity of the networks, it is difficult to maintain the original signal.

This problem becomes even more pronounced if we take into account the low-copy numbers which are typical for cellular signaling cascades. The central limit theorem states that fluctuations around the average response in any random process become insignificant with respect to the average response, if the average response becomes very large. However, the actual output of a signaling cascades can be very small, as small as only 10 copies of a specific molecule. Indeed, rhodopsin, the receptor for photons (light) in the eye can accurately measure the difference between 1 and 2 photons [6], eukaryotic cells can measure a concentration difference of order nanomolar (nM) between the front and back side [7], while gene transcription is sensitive to concentrations of nM, which in bacteria like *E. coli* corresponds to only a handful of molecules [8, 9].

The stochasticity leading to the inevitable uncertainty in the response of signaling cascades is a major challenge for cells. And yet, for the survival of a cell, an accurate response to incoming messages is important. An incorrect transfer of the measurement of a sugar concentration, could result in moving away from abundances of sugar, while the misinterpretation of the α-factor signal in yeast, prevents mating. To overcome these problems, signaling cascades have specific features.

Indeed, studies of complete protein (and gene) interaction networks have shown that specific architectures of coupled interactions are over-represented, and these are called network motifs [10, 11, 12]. The over-representation of these motifs has driven many studies, since this suggest that these motifs have an actual function, or at least an evolutionary advantage for cells [13, 14, 15, 16]. Let us consider what the evolutionary advantages of signaling cascades could be. Two advantages, or criteria, naturally arise: *reliability* and *cost*. Of course, many different advantages can be thought of, like transmission speed [17], memory-capability, both short-term [18] or long-term [19], robustness [20], change of temporal structure [21, 22] or timing [23, 24], but we argue that for a signaling cascade reliability and cost are essential.

Let us start with the second criterion: cost. As every process — e.g. gene expression, active transportation or phosphorylation — in the cells consumes energy, which is limited, it seems natural for a cell to try to reduce the energy consumption while transmitting signals through the cells. There are essentially two ways of reducing cost: 1) making less components and 2) share components between different tasks. Clearly, signaling cascades that rely on a small number of protein copies may have an evolutionary benefit, since they require a smaller production cost of making molecules. The other way of reducing the cost of maintaining a signaling cascade is to re-use the same proteins in multiple signaling cascades. Different signals then share a common cascade to transduce information. This principle, commonly referred to as multiplexing, is also at the heart of our modern macroscopic communication systems: indeed, many telephone calls are transmitted through the same wire

and the European and American internet is connected through the main transatlantic glass fiber backbone. Both ways of reducing cost share an important drawback, the possible loss of correspondence between signal and response. In the case of the shared pathway there is the possibility of interference between the different signals that travel through this pathway, while in the case of the low copy number the correspondence can be lost due to the large impact of fluctuations. This directly brings us back to the first criterion: reliability.

Arguably, the most important property for a signaling cascade is to create a reliable connection between signal and response. Let's think back of the telephone game one more time. To increase the likelihood of a perfect passing around of the message, instead of whispering, children could tell or, even more extreme, shout the message from one child to the next. This increase in amplitude of the voice in this game is similar to our definition of the gain of a signaling cascade. The gain is the amplification of the signal throughout the cascade and a large gain increases the reliability of the cascade as a whole. The counterpart of the gain is the noise: the uncertainty that is present in every biochemical process, be it protein production, activation, deactivation or degradation [25, 26, 27, 28, 29]. A large noise reduces the fidelity in the final message. In forthcoming sections, we will discuss the gain and the noise in much more detail.

The over-representation of these motifs suggests benefits for the cell, and since the gain and the noise are important characteristics of signaling cascades, indeed many studies have been devoted to the gain and noise properties of simple motifs [30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44]. However, as we stress here, to understand the reliability of signal transmission the gain and the noise should be studied simultaneously, since it is the relation, in fact their ratio, between the two that ultimately describes the information transmission capacity. Curiously, we have now introduced a new quantity, which might make sense to every reader intuitively, but, at first sight, seems hard to define quantitatively, namely *information*. Yet, it is precisely this quantity, information that quantifies the reliability of a signaling cascade. Therefore we now, in the next section, first discuss this in detail.

1.2 Information theory

In 1948 Claude Shannon published a paper which would shape the face of communication science for the rest of the century.In fact, one could state that that paper created a whole new scientific field: *information theory* [45]. An important part of information theory studies the reliability of communication systems. The challenge for every communication system is formulated by Shannon as

> *"The fundamental problem of communication is that of reproducing at one point either exactly or approximately a message selected at another point."*

Shannon was able to study this problem in a quantitative manner, which was an important achievement. To do this, he introduced and, importantly, quantified two important concepts, entropy and information.

Assume I have a very special dice, which has 8-sides, listing the numbers $1, 2, 3, \ldots, 8$. I would challenge you for a guessing game, in which I roll this special dice, but I hide the outcome from you. How many closed (yes/no) questions do you need on average, to know the

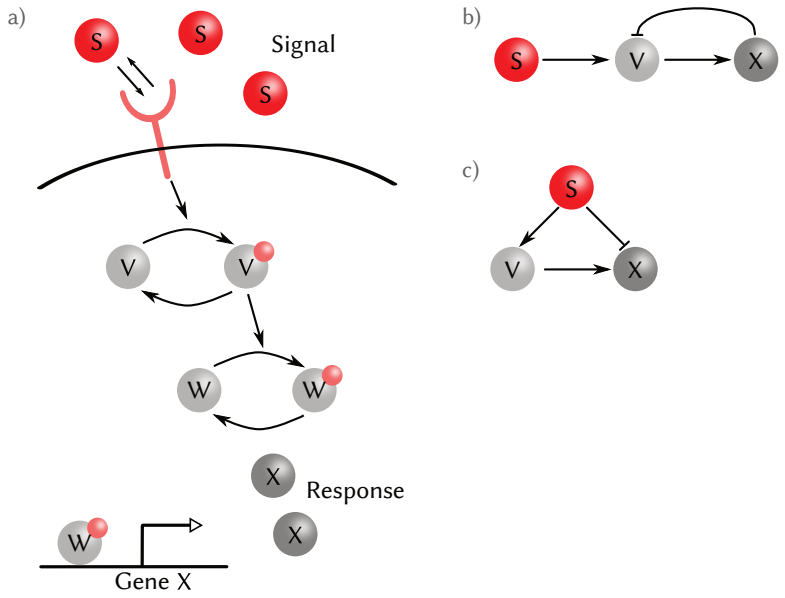


Figure 1.1: a) Example of a signaling cascade, **b,c)** typical motifs. S is the input signal and X is the response, V,W are an intermediate components **b)** feed-back motif, **c)** incoherent feed-forward motif. The motifs in panel **b,c** are studied in more detail in Chapter 2 and Chapters 3, 7, 8 respectively.

outcome of the roll with absolute certainty? It is the answer to this question, that captures the quantification of information as defined by Shannon.

The correct answer is three questions. Assume I have rolled a 7. You start by asking if the outcome is larger than 4, thereby dividing the number of possible outcomes in two. The outcome is either $1, 2, 3, 4$ or $5, 6, 7, 8$. If my answer is yes, you apply the same trick by asking if the outcome is larger than 6, again dividing the set of possible outcomes in two equal halves. After your next question, you know, with absolute certainty that the outcome is 7. What is the relation between this process of "asking questions" and information?

In a mathematical framework, the outcome of the roll is a stochastic variable. Assuming I play fair and the dice is fair, all outcomes are as probable and the variable has a uniform probability distribution,

$$p(X = x) = \frac{1}{8} \tag{1.1}$$

where X is the stochastic variable and x is one of the eight different outcomes. The entropy of the stochastic variable X, or more precisely, the entropy of the distribution $p(X)$

is defined as

$$H(X) \equiv -\sum_{x=1}^{x=N} p(X=x) \log_2 p(X=x) \text{ bits} \tag{1.2}$$

The entropy is often presented in units of bits, which is reflected by the use of $\log_2$. The entropy can be defined for any probability distribution and is bounded by $0 \geq H(X) \geq \log_2 N$. The entropy is zero if there is only a single possible outcome with probability $p(X=x)=1$. The entropy is maximal for a uniform distribution of the outcomes, such that all outcomes have equal probability, or equivalently all outcomes are as likely. The entropy is thus a measure for the uncertainty I have about the stochastic variable (Fig. 1.2).

For a continuous variable the entropy is ill-defined. From a mathematical perspective, the value of a continuous real variable has infinite precision. No matter how accurate your derivation is, or how many questions you ask, you will never discover the full precision of the number. For all practical purposes, continuous variables are discretized into bins and the entropy for the resulting discrete distribution is derived.

The entropy for our stochastic variable X, the outcome of the dice roll,

$$H(X) = -\sum_{x=1}^{x=8} p(X=x) \log_2 p(X=x) = 3 \text{ bits} \tag{1.3}$$

Strikingly, or maybe not, this exactly corresponds to the three questions you had to ask to obtain the answer. Phrased differently, and more precisely, this is the minimum number of questions required to obtain with absolute certainty the value of any stochastic number, *on average*. Of course, you could have guessed, and maybe even have guessed right, but on average guessing would have taken you more questions, than the bi-sectioning procedure described before [46].

1.2.1 Mutual information

The entropy is an important quantity in information theory. Instead of defining the entropy for X alone, we can also define the conditional entropy $H(X|Y)$

$$H(X|Y) = -\sum_{y} p(Y) \sum_{x} p(X|Y) \log_2 p(X|Y) \tag{1.4}$$

with $p(X) = p(X=x)$ and $p(Y) = p(Y=y)$. The inner sum reflects the uncertainty in X for a specific value of Y. If X and Y are independent, $p(X,Y) = p(X)\,p(Y)$ and $p(X|Y) = p(X)$, and as a result, the conditional entropy $H(X|Y) = H(X)$. This indicates that Y is not informative on X. However, if X and Y are not independent, $H(X|Y) < H(X)$, and on average the uncertainty about X is reduced by knowledge of Y. This insight provides the idea for the definition of a new quantity, the mutual information. The mutual information is a unique measure of interdependency between X and

Y [45]. Its definition is

$$I(X;Y) \equiv H(X) - H(X|Y) \tag{1.5}$$
$$= H(X) + H(Y) - H(X,Y) \tag{1.6}$$
$$= -\sum_y \sum_x p(X,Y) \log_2 p(X) + \sum_y \sum_x p(X,Y) \log_2 \frac{p(X,Y)}{p(Y)} \tag{1.7}$$
$$= \sum_y \sum_x p(X,Y) \log_2 \frac{p(X,Y)}{p(X)\,p(Y)}, \tag{1.8}$$

The mutual information is the difference of the entropy in X, $H(X)$, and the conditional entropy $H(X|Y)$, which is the uncertainty in X given a specific Y averaged over the probability distribution $p(Y)$. The mutual information reflects the decrease in the uncertainty of the variable X upon knowledge of the value of the variable Y. More informally said, the mutual information characterizes the information that is shared between Y and X.

Tkačik lists some important well-known properties of the mutual information [46]. First, the mutual information is symmetric in X and Y, as is readily observed from Eq. 1.6 by interchanging X and Y. The reduction in uncertainty in X due to the knowledge of Y is thus equivalent to the reduction in uncertainty of Y upon knowledge of X. This is an intriguing property of the mutual information. Second, the mutual information can be defined for both discrete and continuous variables. Although the entropy for a continuous variable is ill-defined, the mutual information is not. Third, the mutual information obeys the data-processing inequality. For example for a cascade of three stochastic variables, where X regulates Y and Y regulates Z, $X \to Y \to Z$, then $I(X;Z) \leq I(X;Y)$.

More in depth information on the mutual information can be found in the original paper by Shannon [45] or the books by Cover and Clover [5] or MacKay [47]. The mutual information as defined in Eq. 1.5 is used in Chapters 7 and 8.

In the neuroscience community information theory has been used for a few decades (see for a review [48]). But Ziv *et al* [49] were, arguably, the first to use information theory for a biochemical network, in this case a gene regulatory network. Since then many more studies have been performed, all using information transmission as the main quantity to describe the signaling properties of the network [50, 51, 52, 53, 54, 55, 56, 57, 58].

Noisy channel theorem

Let us take a look at the following process. We consider a system where Arthur J. (A) sends a single signal $\mathcal{S}$ through a channel, where at the other end of the channel G. Bomans (B) receives a response $\mathcal{X}$. A selects his signal from a probability distribution $p(S)$. If the channel is perfect, or in other words deterministic or noiseless, the response $\mathcal{X}$ and signal $\mathcal{S}$ are uniquely related and B would know precisely what $\mathcal{S}$ is, if B has received $\mathcal{X}$.

Most natural systems (unfortunately[1]) are not noiseless. As a result, the unique relation between $\mathcal{S}$ and $\mathcal{X}$ disappears, and B can only estimate what $\mathcal{S}$ is using the response

[1]Many studies have been devoted to the advantages of noise in natural systems, some of which are the introduction of heterogeneity in populations [59], stochastic amplification [60], stochastic resonance [61], increase in growth rate [62] or excitable systems [63]

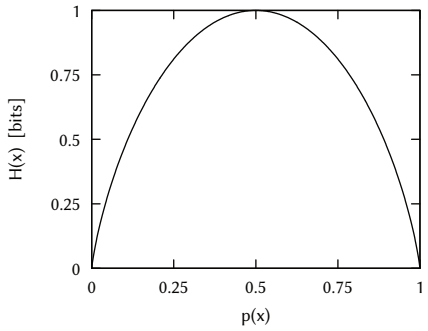


Figure 1.2: The entropy of a binary variable p(X=0)=1−p(X=1). For p(X=1)=1 or p(X=1)=0, the entropy is zero, since there is only a single possible outcome. The entropy is maximal for p(X=1)=p(X=0)=1/2 since then both outcomes are as likely.

$\mathcal{X}$. For these channels, what is the maximum amount of information that can be transmitted? Thinking back of the telephone game, the information is transmitted more reliably if the participants would not whisper, but talk or even shout. In reality however, there is a physical limit to shouting, since it drains energy from anyone. Using this analogy loosely, any channel can transmit infinite information, but considering the cost of using energy, I constrain the average amplitude of the signal $\langle A_S^2 \rangle \leq v$. With this knowledge alone, we can not determine the maximal information through the channel; for this we also require the noise characteristics of the channel. I assume that the noise η in the channel has the three following properties: 1) η has a Gaussian distributed amplitude ($\eta \sim \mathcal{N}(0, \sigma_\eta^2)$)

$$\mathcal{N}\left(y; \mu, \sigma^2\right) = p\left(y|\mu, \sigma^2\right) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{(y-\mu)^2}{2\sigma^2}}, \tag{1.9}$$

2) η is spectrally white and 3) η is independent of the signal $\mathcal{S}$. This type of channel is called an additive Gaussian channel.

The maximal mutual information, defined as the channel capacity C, that can be transmitted through an additive Gaussian channel is

$$C = \max I\left(S, X\right) = \frac{1}{2} \ln\left(1 + \frac{A_S}{\sigma^2}\right) \text{ nats}. \tag{1.10}$$

The channel capacity is a function of the ratio of the signal amplitude and the noise strength, which is commonly referred to as *signal-to-noise-ratio* (*SNR*). The channel capacity is obtained if the input distribution $p\left(S\right)$ is Gaussian distributed; $\mathcal{N}\left(0, A_S\right)$ [47, 64]. This is an important observation, which I will rephrase: given a specific variance of the input distribution $p(S)$ and a Gaussian channel, the mutual information between $\mathcal{S}$ and $\mathcal{X}$ is maximized if $\mathcal{S}$ has a Gaussian distribution. Very related to this observation is the following statement: for a channel with a Gaussian distribution for the input signal the mutual information is minimal if the conditional distribution between signal and response $p(x|s)$ is Gaussian [64].

The origin of this lower bound is that, for a given variance, a Gaussian distribution has the maximum entropy. Therefore, the channel with a Gaussian conditional distribution between signal and response "creates" the largest possible uncertainty between signal and response and therefore a lower bound on the mutual information.

Mutual information for time-varying signals

As we have seen the mutual information is a measure for the statistical dependence between two single univariate stochastic variables X and Y. But a straightforward extension to this analysis provides the mutual information between two multivariate stochastic variables $\boldsymbol{V} = (V_1, V_2, \ldots, V_{N_v})^T, \boldsymbol{W} = (W_1, W_2, \ldots, W_{N_w})^T$ [61, 64, 65, 66, 67]. This extension is the connection we require to obtain the mutual information between time-varying signals.

Assume we have a stochastic process which provides us with a continuous time trace $V(t)$. This trace has Gaussian fluctuations such that multiple realizations of the process are different, and we can write $V(t) = \langle V(t) \rangle + v(t)$, with $\langle V(t) \rangle$ the time-dependent average and $v(t)$ the fluctuations around the average.

We sample the signal $v(t) = V(t) - \langle V(t) \rangle$ at N evenly spaced time points to obtain the vector $\boldsymbol{v} = (v(t_0), v(t_1), v(t_2), \ldots, v(t_N))^T$ (we thus assume the fluctuations to be the signal). The correlation matrix $\mathbf{C}$ for $\boldsymbol{v}$ is defined as

$$C_{ij} = \left\langle v(t_i)\, v\left(t_j\right) \right\rangle = \int v(t_i)\, v\left(t_j\right) p(\boldsymbol{v})\, d\boldsymbol{v} \tag{1.11}$$

Since the fluctuations in the signal are Gaussian, the joint probability distribution for the vector $\boldsymbol{v}$ is

$$p(\boldsymbol{v}) = \frac{1}{\sqrt{(2\pi)^N |\mathbf{C}|}} e^{-\frac{1}{2}\boldsymbol{v}^T \mathbf{C}^{-1} \boldsymbol{v}}, \tag{1.12}$$

where $\mathbf{C}^{-1}$ is the inverse of the correlation matrix and $|\mathbf{C}|$ is the determinant of the correlation matrix. With the definition of the probability distribution for the vector $\boldsymbol{v}$ (Eq. 1.12), we can define the differential entropy for the multivariate stochastic continuous variable $\boldsymbol{v}$ following Eq. 1.2

$$H(\boldsymbol{v}) = \int d\boldsymbol{v} p(\boldsymbol{v}) \ln p(\boldsymbol{v}) \quad [\,\text{nats}\,]. \tag{1.13}$$

Note that, for later convenience, we have changed the base of the logarithm, and as a result, the differential entropy has units [nats]. Further, we observe that the differential entropy $H(\boldsymbol{v})$ is not necessarily larger than zero. Since the joint probability distribution $p(\boldsymbol{v})$ for all the elements $v(t_i)$ has Gaussian statistics, we can actually obtain an analytical expression for the entropy $H(\boldsymbol{v})$. Since the correlation matrix $\mathbf{C}$ is a real, symmetric matrix, we write

$$\mathbf{C}^{-1} = \mathbf{Q}\mathbf{D}\mathbf{Q}^T, \text{ with } \mathbf{Q}^{-1} = \mathbf{Q}^T \text{ and } \mathbf{Q}^{-1}\mathbf{Q} = \mathbf{I}, \tag{1.14}$$

where $\mathbf{I}$ is the identity matrix. The matrix $\mathbf{Q}$ has the eigenvectors of $\mathbf{C}^{-1}$ as column vectors.

The matrix $\mathbf{D}$ is a diagonal matrix with the eigenvalues of $\mathbf{C}$ along the diagonal. The matrix $\mathbf{Q}$ also transforms the vector components v_i into new variables u_i:$\Rightarrow \boldsymbol{u} = Q^T \boldsymbol{v}$. The differential entropy of the distribution $p(\boldsymbol{v})$ is

$$H(\boldsymbol{v}) = -\int \frac{1}{\sqrt{(2\pi)^N |\mathbf{C}|}} e^{-\frac{1}{2}\boldsymbol{u}^T \mathbf{D}^{-1} \boldsymbol{u}} \ln\left(\frac{1}{\sqrt{(2\pi)^N |\mathbf{C}|}} e^{-\frac{1}{2}\boldsymbol{u}^T \mathbf{D}^{-1} \boldsymbol{u}}\right) d\boldsymbol{u} \tag{1.15}$$

$$= \frac{1}{\sqrt{(2\pi)^N |\mathbf{C}|}} \int e^{-\frac{1}{2}\boldsymbol{u}^T \mathbf{D}^{-1} \boldsymbol{u}} \left(\ln\left(\sqrt{(2\pi)^N |\mathbf{C}|}\right) + \frac{1}{2}\boldsymbol{u}^T \mathbf{D}^{-1} \boldsymbol{u}\right) d\boldsymbol{u} \tag{1.16}$$

Since D is diagonal, we simplify Eq. 1.16 by substituting

$$\int e^{\frac{1}{2}\boldsymbol{u}^T \mathbf{D} \boldsymbol{u}} d\boldsymbol{u} = \prod_{i=1}^{N} \int du_i e^{-\frac{1}{2}\lambda_i u_i^2} = \sqrt{(2\pi)^N \prod_{i=1}^{N} \frac{1}{\lambda_i}} = \sqrt{\frac{(2\pi)^N}{|\mathbf{D}|}} \tag{1.17}$$

where λ_i is the i$^{\text{th}}$ eigenvalue of $\mathbf{D}$. Since $|\mathbf{D}| = \left|\mathbf{C}^{-1}\right| = 1/|\mathbf{C}|$, we obtain

$$H(\boldsymbol{v}) = \ln\left(\sqrt{(2\pi e)^N |\mathbf{C}|}\right) = \ln\left(\sqrt[N]{2\pi e} \prod_{i=1}^{N} \sqrt{\lambda_i}\right) \tag{1.18}$$

Eq. 1.18 gives the differential entropy for a multivariate, Gaussian distribution. The derivation of Eq. 1.18 can easily be extended to the joint probability distribution of two multivariate, Gaussian processes $\boldsymbol{v}$ and $\boldsymbol{w}$. Introducing the new vector $\boldsymbol{z} = (\boldsymbol{v}, \boldsymbol{w})^T$ with correlation matrix $\mathbf{Z}$

$$\mathbf{Z} = \begin{pmatrix} \mathbf{C}_{vv} & \mathbf{C}_{wv} \\ \mathbf{C}_{vw} & \mathbf{C}_{ww} \end{pmatrix}, \tag{1.19}$$

an equivalent derivation holds. As a result, since we have analytical expressions for the differential entropies $H(\boldsymbol{v}), H(\boldsymbol{w})$ and $H(\boldsymbol{v}, \boldsymbol{w})$, following Eq. 1.6 we write the mutual information between the multivariate, Gaussian distributions $\boldsymbol{v}$ and $\boldsymbol{w}$ as [67]

$$I_t(\boldsymbol{v}; \boldsymbol{w}) = H(\boldsymbol{v}) + H(\boldsymbol{w}) - H(\boldsymbol{v}, \boldsymbol{w}) = \frac{1}{2} \ln\left(\frac{|\mathbf{C_{vv}}| \, |\mathbf{C_{ww}}|}{|\mathbf{Z}|}\right). \tag{1.20}$$

As a simple example let us focus on the situation where N, the number of time points, is 1. Then the vectors $\boldsymbol{v}$ and $\boldsymbol{w}$ reduce to $v(t_0), w(t_0)$ and the correlation matrices reduce to the instantaneous (co)variances, $\sigma_{ij} = \langle i(t_0) \, j(t_0) \rangle$, with $i = v, w$ and $j = v, w$. With

these following Eq. 1.20 we obtain the *instantaneous* mutual information

$$I_t(\boldsymbol{v};\boldsymbol{w}) = \frac{1}{2}\ln\left(\frac{\sigma_{vv}^2\sigma_{ww}^2}{\sigma_{vv}^2\sigma_{ww}^2-\sigma_{vw}^4}\right) = -\frac{1}{2}\ln\left(1-\frac{\sigma_{vw}^4}{\sigma_{vv}^2\sigma_{ww}^2}\right) \tag{1.21}$$

$$= \frac{1}{2}\ln\left(1+\frac{\sigma_{vw}^4}{\sigma_{vv}^2\sigma_{ww}^2-\sigma_{vw}^4}\right) \tag{1.22}$$

Comparing Eq. 1.21 with Eq. 1.10 we observe that [67]

$$\text{signal-to-noise} = \frac{\sigma_{vw}^4}{\sigma_{vv}^2\sigma_{ww}^2-\sigma_{vw}^4} \equiv \frac{g^2}{N}\sigma_{vv}^2 \tag{1.23}$$

where we have defined the instantaneous gain $g^2 \equiv \sigma_{vw}^4/\sigma_{vv}^4$ and noise $N \equiv |\mathbf{Z}|/\sigma_{vv}^2 = \sigma_{ww}^2 - \sigma_{vw}^4/\sigma_{vv}^2 = \sigma_{ww}^2 - g^2\sigma_{vv}^2$. The noise therefore is that part of the variance in w, that is not "explained" by the input variance (the signal) multiplied with the instantaneous gain g^2.

The mutual information rate

We are almost there. In this last section on information theory we derive the mutual information rate $R(\omega)$, which is used extensively in Chapters 2 and 3.

In the previous section we have derived the general expression for the mutual information between multivariate, Gaussian distributions (Eq. 1.18) for time-discretized signals. However, biochemical signals are not discretized in time, but are continuous. Here we make the connection to continuous signals from the time-discretized signals. To simplify the results we assume stationarity, meaning that $\langle V(t)\rangle = \langle V\rangle$ and $\langle v(t)v(t')\rangle = \langle v(\tau)\,v(0)\rangle$, where $\tau = t - t'$ and similar expressions for W.

We sample the traces $v(t)$ and $w(t)$, with length T, at N points with spacing Δ and, again, assume that the multivariate joint distribution $p(\boldsymbol{v},\boldsymbol{w})$ is Gaussian. We write the general eigenvalue equation for the correlation matrix (Eq. 1.11) [61]

$$\sum_k C_{vv}(t_j,t_k)\,e^{-i\omega_0 nk\Delta} = \lambda_n e^{-i\omega_0 nj\Delta}, \tag{1.24}$$

with $\omega_0 = 2\pi/T$. Since we have assumed stationarity the elements of C_{vv} depend only on the absolute time difference $|j-k|$, and therefore the correlation matrix has a Toeplitz structure. In the limit $T\to\infty, N\to\infty$ this means that [5, 66]

$$\lim_{N,T\to\infty}|\mathbf{C}_{vv}| = \frac{1}{4\pi T}\int_{\omega\text{max}}^{\omega\text{max}}\ln S_{vv}(\omega)\,d\omega, \tag{1.25}$$

where $S_{vv}(\omega)$ is the power spectrum of the signal $v(t)$ and $\omega_{\text{max}} = \pi/\Delta$ is the angular Nyquist frequency. With Eq. 1.25 we define the entropy rate h_G as

$$h_G(V) = \lim_{N,T\to\infty}\frac{H_G}{N\Delta} \tag{1.26}$$

Note that in the limit $\Delta \to 0$, the entropy rate h_G is ill-defined. To obtain the mutual information rate, we require the joint entropy rate $h_G(V, W)$ (Eq. 1.6). Although (in general) the correlation matrix $\mathbf{Z}$ is not Toeplitz, it is possible to obtain the mutual information rate in terms of the (cross-) power spectra [61, 65, 67]

$$\lim_{N\to\infty} \frac{I(\boldsymbol{v};\boldsymbol{w})}{N\Delta} = R(\boldsymbol{v};\boldsymbol{w}) = -\frac{1}{4\pi}\int_{\omega\text{max}}^{\omega\text{max}} d\omega \ln\left[1 - \frac{|S_{vw}(\omega)|^2}{S_{vv}(\omega)\, S_{ww}(\omega)}\right] \tag{1.27}$$

$$\equiv \frac{1}{4\pi}\int_{\omega\text{max}}^{\omega\text{max}} d\omega \ln\left[1 + \frac{g^2(\omega)}{N(\omega)} S_{vv}(\omega)\right] \qquad \left[\text{nats s}^{-1}\right]. \tag{1.28}$$

The mutual information rate is the sum over independent Gaussian channels, where for each frequency the (co)variance is given by the (cross-)power spectrum. For independent Gaussian channels, the system should be linear, such that each input at a specific frequency only leads to an output at that same frequency. Since not every network is linear, a linearization procedure is required, and for this we use the Linear-Noise Approximation [68] (see the section: *Linear-Noise Approximation*). The mutual information rate is defined as a bit-rate, like the bit-rate of traditional modems. In Eq. 1.28 a gain and noise are defined again, but here these are frequency dependent.

$$g^2(\omega) \equiv \frac{\left|S^2_{vw}(\omega)\right|^2}{S^2_{vv}(\omega)} \tag{1.29}$$

$$N(\omega) \equiv S_{ww}(\omega) - g^2(\omega)\, S_{vv} \tag{1.30}$$

We note the similarity in structure between Eq. 1.21 and Eq. 1.27, showing that (the) mutual information (rate) in general is proportional to the signal-to-noise and/or gain-to-noise ratio. The important difference between these two ratios is discussed in Chapters 2 and 3.

An example

Let's assume that we have a set of signals, which have a Gaussian distribution such that $p(S) = \mathcal{N}\left(\mu_S, \sigma^2_{SS}\right)$. Each independent signal leads to an output distribution of responses, and the conditional distribution is Gaussian $p(X|S) = \mathcal{N}\left(\mu_{X|S}, \sigma^2_{X|S}\right)$. The mutual information can now be calculated through Eq. 1.5, after specifying the (conditional) mean and variance. For reasons that will become clear shortly, we assume the following forms

$$\begin{aligned} \mu_S = \langle S\rangle = \frac{k}{\lambda},\ \sigma^2_{SS} = \frac{k}{\lambda}, \\ \mu_{X|S} = \langle X|S\rangle = \frac{\rho}{\mu\langle S\rangle},\ \sigma^2_{X|S} = \frac{\rho}{\mu\langle S\rangle} \end{aligned} \tag{1.31}$$

Note here that although the distributions are Gaussian, the channel noise $\sigma^2_{X|S}$ depends on the signal $\mathcal{S}$. This is thus not an additive Gaussian channel, since the noise depends on the signal, which can be observed in Fig. 1.3a,b from the widening of the distribution

$p\left(S/\langle S\rangle, X/\langle X\rangle\right)$ for increasing S. In Fig. 1.4a we show the mutual information as function of μ for constant $\langle S\rangle = 100$, with $\rho = \lambda = 1$. For increasing μ the mutual information $I(S;X)$ decreases. The origin of this decrease is the decrease of the value of the conditional response $\langle X|S\rangle$, leading to an increase of the coefficient of variation (CV) $\sigma^2_{X|S}/\langle X|S\rangle^2$. This is shown in Fig. 1.3a,b, where for two values of μ (resp, $\mu = 0.1, 1$), the probability distribution $p(S/\langle S\rangle, X/\langle X\rangle)$ (gray contour), the dose-response relation (dashed light gray), and most importantly the conditional distribution $p(X/\langle X|S\rangle\,|S = \langle S\rangle)$ (red, logscale) are shown. For larger μ, the conditional distribution is much wider, reflecting the larger CV and therefore the smaller mutual information. We repeat that in Chapters 7 and 8 the mutual information for time-independent input signals is used.

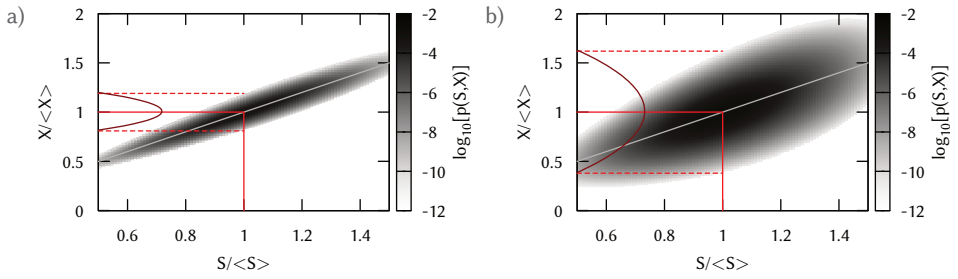


Figure 1.3: The contourplot shows the joint probability distribution p(S/⟨S⟩,X/⟨X⟩) for two different values of μ. Both the conditional mean ⟨X|S⟩ and variance $\sigma^2_{X|S}$ scale with μ^{-1}. The coefficient of variation (CV) $\sigma^2_{X|S}/\langle X|S\rangle^2 \propto \mu$ therefore increases with μ. Indeed, the conditional distribution p(X/⟨X⟩|S/⟨S⟩=1) widens for an increase in μ (red line, logscale). Since the CV increases, the mutual information I(S;X) decreases. **a)** μ=0.1 s^{-1} **b)** μ=1 s^{-1}. Parameters: k=100 s^{-1}, λ=ρ= 1 s^{-1}.

Having observed what the mutual information is for constant signals, we switch to the mutual information for time-varying signals. We take the, arguably, most simple signaling cascade, consisting of a signal S and a response X, which is governed by the following biochemical reactions

$$\emptyset \xrightarrow{k} \mathsf{S}, \quad \mathsf{S} \xrightarrow{\lambda} \emptyset, \quad \mathsf{S} \xrightarrow{\rho} \mathsf{S} + \mathsf{X}, \quad \mathsf{X} \xrightarrow{\mu} \emptyset. \tag{1.32}$$

As input signal we take the variations $s\left(t\right) = S\left(t\right) - \langle S\rangle$ and the output is $x\left(t\right) = X\left(t\right) - \langle X\rangle$. This system is described by the following set of Langevin equations [69]

$$\frac{ds}{dt} = -\mu s\left(t\right) + \eta_s\left(t\right) \tag{1.33}$$

$$\frac{dx}{dt} = \rho s\left(t\right) - \mu x\left(t\right) + \eta_x\left(t\right). \tag{1.34}$$

Here $\eta_x(t)$ denotes the noise due to the production and degradation of X, while $\eta_s(t)$ is not a noise source in the usual sense, but rather defines the ensemble of input signals,

which is assumed to obey Gaussian statistics. They both have the following two properties $\langle \eta_\alpha(t) \rangle = 0$, $\langle \eta_\alpha(t) \eta_\alpha(t') \rangle = A_\alpha \delta(t - t')$, where $\alpha = s, x$ and $A_s = 2\mu\langle S \rangle$, $A_x = 2\lambda\langle X \rangle$ [70] (see the section: *Linear-Noise Approximation*). For this simple system we calculate the instantaneous mutual information (Eq. 1.21) and the mutual information rate (Eq. 1.27).

The steady-state responses are (compare with Eq. 1.31)

$$\langle S \rangle = \frac{k}{\lambda}, \quad \langle X \rangle = \frac{\rho}{\mu} \langle S \rangle = \frac{\rho k}{\lambda \mu}, \tag{1.35}$$

and the steady-state (co)variances are

$$\frac{\sigma_{ss}^2}{\langle S \rangle} = 1, \quad \frac{\sigma_{xx}^2}{\langle X \rangle} = 1 + \frac{\rho}{\mu + \lambda}, \quad \sigma_{sx}^2 = \frac{\rho \langle S \rangle}{\mu + \lambda}, \tag{1.36}$$

The instantaneous mutual information is (Eq. 1.21)

$$I_t(s; x) = \frac{1}{2} \ln \left[1 + \frac{\sigma_{sx}^4}{\sigma_{ss}^2 \sigma_{xx}^2 - \sigma_{sx}^4} \right] \tag{1.37}$$

$$= \frac{1}{2} \ln \left[\frac{(\lambda + \mu)(\lambda + \mu + \rho)}{(\lambda + \mu)^2 + \lambda\rho} \right]. \tag{1.38}$$

The instantaneous mutual information gives the information between X at time t and S at time t. For the system as described in Eq. 1.32 the response time of X is set by μ. We first discuss some interesting limits. In the limit $\lambda \to \infty$ $I_t(s, x) = 0$, since for $\lambda \to \infty$, every variation in s will decay, before a change in x can be established. Indeed, the number of X molecules produced in the lifetime of a single S molecule is ρ/λ, and in the limit $\lambda \to \infty$, $\rho/\lambda \to 0$. In the opposite limit, $\lambda \to 0$, the variations in s decay very slowly, and the instantaneous mutual information saturates ($I_t \propto (\mu + \rho)/\mu$), where the saturation value depends on the relative production and degradation rates of X. In the limit $\rho \to \infty$, the instantaneous mutual information reaches a saturation value ($I_t \propto (\mu + \lambda)/\lambda$). In this limit, the production of X molecules increases and therefore the influence of fluctuations in the copy number of X decreases. Further, in this limit, we observe that an increase in μ, reflecting faster tracking of the signal, increases the saturation value. The instantaneous mutual information is independent of k, the production rate of the signal (as long as $k > 0$). A change in k changes the mean level of $\langle S \rangle$, but the timescale for the decay of variations in s only depends on λ, since the capacity of x to reliably track the variations in s depends only on the dynamics. As a function of μ, the instantaneous mutual information has a maximum, which is obtained for $\mu_{\max} = \sqrt{\lambda(\lambda + \rho)}$,

$$I_{t,\max}(s; x) = \frac{1}{2} \ln \left[\frac{\lambda + \sqrt{\lambda(\lambda + \rho)}}{2\lambda} \right]. \tag{1.39}$$

For $\mu \gg \lambda$, the intrinsic fluctuations in X are much more rapid than variations in X due to variations in S. As a result, the instantaneous statistics of X are those of a Poisson birth-

death process with constant input S and conditional variance $\sigma^2_{x|s} \approx \langle X \rangle \propto \mu^{-1}$. This increase in the noise is not compensated by a similar increase in the instantaneous gain $\sigma^4_{xs}/\sigma^4_{ss} \propto \mu^{-2}$. As a result the instantaneous mutual information scales with μ^{-1}. In the opposite limit, $\mu \ll \lambda$, the response X is much slower than the variations in S, and the response integrates over the fluctuations in S. The actual value of $x(t)$ therefore does not reflect the instantaneous value of $s(t)$, but the time-average of $s(t)$ over a time-window $t_{\text{int}} = \mu^{-1}$; clearly, if $\mu \to 0$ all the variations in s are integrated and x will not respond to changes in s. In Fig. 1.4b the instantaneous mutual information is shown as function of μ for three different values of ρ, where $\lambda = 1\,\text{s}^{-1}$. First, the non-monotonic form with the maximum at μ_{max} is observed. Next, if $\rho \gg 1$, the mutual information is large, since this implies a larger amplification of x following changes in the signal s.

We note here that the instantaneous mutual information and the mutual information for constant signals are not equivalent. The instantaneous mutual information is influenced by temporal correlations of the signal and the response, while the mutual information for constant signals is not. In the limit $\lambda \to 0$ and for equal conditional variances, the instantaneous mutual information approaches the constant mutual information (Fig. 1.4a, gray dashed line).

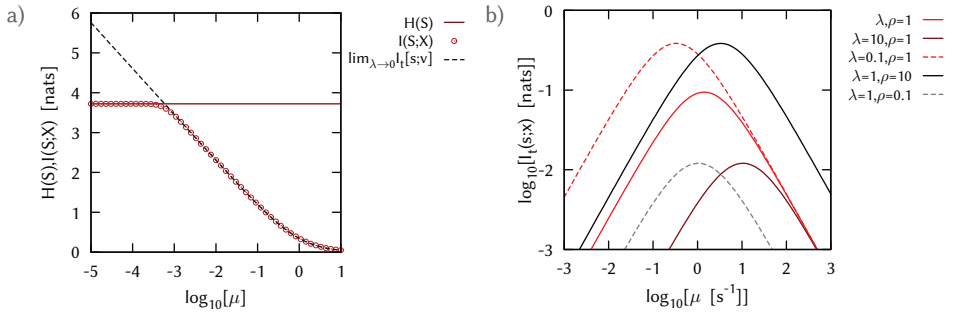


Figure 1.4: a) The difference between I(S;X) (red solid) and H(S) (red dashed), respectively the mutual information and the entropy of the input signalS, for a discrete joint Gaussian distribution p(S,X). I(S;X) decreases with increasing μ, since CV $\sigma^2_{X|S}/\langle X|S\rangle^2 \propto \mu$. The maximal value of the mutual information $\max$ I(S;X)=H(S) is obtained in the limit $\mu \to 0$. The mutual information is equal to the instantaneous mutual information in the limit $\lambda \to 0$ (gray line). For small μ the instantaneous mutual information is larger, since the instantaneous mutual information is obtained assuming continuous signals and therefore the conditional differential entropy H(s|x) can become negative, reflecting infinite precision on the value of s given knowledge of x. **b)** The instantaneous mutual information I_t(s;x) increases for $\rho \gg \lambda$ since this implies that the amplification of x(t) by the variations in s(t) is larger. In the limit $\lambda \to 0$ (red dashed, black solid) the input signal effectively is constant and $\sigma^2_{SX} = \langle X \rangle$, $I_t(s;x) = 1/2 \ln(1+\rho/\mu) = 1/2 \ln(\sigma^2_{XX}/\langle X \rangle)$. In the limit $\lambda \to \infty$ the variations in s decay before these are transmitted to x leading to a small instantaneous mutual information (gray dashed, dark red solid). For large μ , the instantaneous mutual information scales as $\rho\lambda/\mu$. λ, ρ in s^{-1}.

Finally, we look for this simple system at the mutual information rate $R(\boldsymbol{s}, \boldsymbol{x})$ (Eq. 1.27),

the mutual information between complete trajectories $s(t), x(t)$. The mutual information rate R is obtained through the calculation of the (cross-)power spectra, which are

$$S_{SS} = \frac{2k}{\omega^2 + \lambda^2}, \tag{1.40}$$

$$S_{XX} = \frac{\frac{2\rho k}{\lambda}\left(\omega^2 + \lambda\left(\lambda + \rho\right)\right)}{\left(\omega^2 + \lambda^2\right)\left(\omega^2 + \mu^2\right)}, \tag{1.41}$$

$$S_{SX} = \frac{2\rho k}{\left(\omega^2 + \lambda^2\right)\left(i\omega + \mu\right)}, \tag{1.42}$$

and we obtain for $R(\boldsymbol{s}; \boldsymbol{x})$

$$\lim_{T\to\infty} R\left(\boldsymbol{s}; \boldsymbol{x}\right) = \frac{1}{4\pi}\int_{-\infty}^{\infty} d\omega \ln\left(\frac{\omega^2 + \lambda^2}{\omega^2 + \lambda^2 + \rho\lambda}\right) = \frac{\lambda}{2}\left(\sqrt{1 + \frac{\rho}{\lambda}} - 1\right). \tag{1.43}$$

Most striking about this result is that the mutual information rate does not depend on μ, the degradation rate for the fluctuations in X. However reconsidering this observation, this follows from the fact that the degradation events in X do not carry additional information on the actual state of S, given that the production events of X are known [67], since the degradation only depends on the current state of X. We take a closer look at the frequency-dependent gain and noise for this simple system,

$$g^2\left(\omega\right) = \frac{\left|S_{SX}\left(\omega\right)\right|^2}{S_{SS}^2\left(\omega\right)} = \frac{\rho^2}{\omega^2 + \mu^2}, \tag{1.44}$$

$$N\left(\omega\right) = S_{XX}\left(\omega\right) - \frac{\left|S_{SX}\left(\omega\right)\right|^2}{S_{SS}^2\left(\omega\right)} S_{SS}\left(\omega\right) = \frac{\frac{2\rho k}{\lambda}}{\omega^2 + \mu^2}, \tag{1.45}$$

which lead to a constant gain-to-noise ratio

$$\frac{g^2\left(\omega\right)}{N\left(\omega\right)} = \frac{\rho\lambda}{2k} = \frac{\rho}{\langle S\rangle}. \tag{1.46}$$

The gain-to-noise ratio therefore does not depend on the dynamics of the signal $s(t)$, but only on the dynamics of the transmission of the signal to X. This is a property which we will explore in much more detail in Chapters 2 and 3.

Lastly, we look at the frequency-dependent gain from another perspective. Assume we have a signal $s(t)$ which is transferred through a black box, of which we have no knowledge. We do however assume that the black box has a noise-additive effect. Our main question is, whether we can infer from $x(t)$ the original signal $s(t)$. That is, after measuring the response $x(t)$, we estimate the signal $s_{\text{est}}(t)$. Following [48], we use a linear filter $h(t)$ to estimate $s_{\text{est}}(t)$ from $x(t)$. In the frequency domain a linear filter operation is given by

$$S_{\text{est}}\left(\omega\right) = H\left(\omega\right) X\left(\omega\right) \tag{1.47}$$

Minimizing the least square error between the estimated signal S_{est} and the true signal S (min $\left[\left\langle (S_{\text{est}}(\omega) - S(\omega))^2 \right\rangle\right]$) leads to the following form for the filter [48]

$$H(\omega) = \frac{S_{SX}(\omega)}{S_{XX}(\omega)}. \tag{1.48}$$

The noise is defined as the difference between the true signal and the estimated signal $N_e(\omega) = S(\omega) - S_{\text{est}}(\omega)$ and with the noise the signal-to-noise ratio is defined as

$$\frac{S_{S_{\text{est}}S_{\text{est}}}(\omega)}{N(\omega)} = \frac{\left\langle H(\omega) X(\omega) H^{\dagger}(\omega) X^{\dagger}(\omega) \right\rangle}{\left\langle N_e(\omega) N_e^{\dagger}(\omega) \right\rangle} \tag{1.49}$$

$$= \frac{H^2(\omega) S_{XX}(\omega)}{S_{SS}(\omega) + H^2(\omega) S_{XX}(\omega) - \left\langle H^{\dagger}(\omega) X^{\dagger}(\omega) S(\omega) \right\rangle - \left\langle H(\omega) X(\omega) S^{\dagger}(\omega) \right\rangle} \tag{1.50}$$

$$= \frac{|S_{XS}(\omega)|^2}{S_{XX}(\omega) S_{SS}(\omega) - |S_{XS}(\omega)|^2}, \tag{1.51}$$

where † denotes the complex conjugate. The signal-to-ratio in Eq. 1.49, with the gain and noise as defined in respectively Eq. 1.44, Eq. 1.45, is equal to $g^2(\omega)/N(\omega) S_{SS}(\omega)$. The signal-to-noise ratio is thus equivalent to the best estimate of the signal given the output at a specific frequency, or the reciprocal of the uncertainty in $S(\omega)$ given a particular $X(\omega)$.

1.3 Linear-Noise Approximation

As discussed in the previous sections, all biochemical processes are stochastic. A rigorous way to describe stochastic processes is by formulating a master equation for the underlying process [68], which describes the time-evolution of the full probability distribution of a process. Assume we have a simple birth-death process for a protein X, which is described by the following two processes

$$X \overset{k_f}{\rightarrow} X + 1, \; X \overset{k_b X}{\rightarrow} X - 1 \tag{1.52}$$

where X is the actual copy number of X. The master equation that describes the probability to observe X X proteins is

$$\frac{dp(X)}{dt} = -\left(k_f + k_b X\right) p(X) + k_b (X+1) p(X+1) + k_f p(X-1). \tag{1.53}$$

The first term describes the two ways to leave the state X, either by the production or degradation of X, while the other two terms describe the two options to enter the state X. The master equation is a complete description to the underlying stochastic problem and therefore analytic solutions captures all the statistics of the process (see e.g. [34]). However,

analytical solutions are more often than not very difficult to obtain, mostly due to the fact that the processes are non-linear. Even the most simple enzymatic reaction

$$\mathsf{S} + \mathsf{E} \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} \mathsf{ES}, \tag{1.54}$$

$$\mathsf{ES} \overset{k_2}{\rightarrow} \mathsf{E} + \mathsf{P}, \tag{1.55}$$

where S is the substrate molecule, E is the enzyme and P is the product, is described by a non-linear master equation. It is well-known that many biological processes have even stronger non-linear dependencies, like the dependence of the flagellar motor in *E. coli* on switching proteins [71] or the dependence of the expression level of the $\mathrm{P_{GAL1\text{-}D12}}$ promoter on its repressor TetR [72]. While non-linear processes show very rich and fascinating dynamics [73], they are difficult to solve analytically. In general many people resort to Stochastic Simulations Algorithms (SSA) as has been pioneered for chemical kinetics by Gillespie [74, 75]. A different approach is to look for approximations of the master equation that lead to analytically tractable solutions

One of these solutions is the Linear-Noise Approximation (LNA), which we now discuss in more detail. The LNA provides us with a simple form for analytical approximations of the noise in any non-linear stochastic process. Here we will follow the derivation as introduced by Gillespie [69, 76], and directly apply it to an example.

We again start from the biochemical equations, as introduced in Eq. 1.32. We define the vector $\boldsymbol{Y}(t) = (S(t), X(t))$, where $S(t), X(t)$ are the respective copy number of species S, X at time t. Next we define the propensity functions a_j, where $a_j(\boldsymbol{y})\,dt$ is the probability that in time $[t, t+dt]$, given $\boldsymbol{Y}(t) = \boldsymbol{y}$ a reaction j of occurs, such that we have for our example

$$a_0 = k, \quad a_1 = \lambda S(t), \quad a_2 = \rho S(t), \quad a_3 = \mu X(t) \tag{1.56}$$

Next to the propensity function we define the state-change stoichiometry vector $\boldsymbol{\nu}_j$, with ν_{ij} is the change in Y_i given that reaction j occurs, such that we have $\boldsymbol{v}_0 = (1,0)^T$, $\boldsymbol{v}_1 = (-1,0)^T$, $\boldsymbol{v}_2 = (0,1)^T$ and $\boldsymbol{v}_3 = (0,-1)^T$. With the propensity function and stoichiometry vector the chemical master equation can be formulated,

$$Y_i(t+\tau) = y_{t,i} + \sum_{j=1}^{M} K_j(\boldsymbol{Y}(t), \tau)\,\nu_{ji}, \tag{1.57}$$

where $K_j(\boldsymbol{Y}(t), \tau)$ is a random variable which reflects the number of events N in each reaction channel j in a time τ and $y_{t,i}$ is the value of Y_i at time t. Note that Eq. 1.57 is exact, since $\boldsymbol{Y}(t)$ is time-dependent. Now we make two independent assumptions to obtain a simpler expression for the random variables $K_j(\boldsymbol{Y}(t), \tau)$. Suppose that we can define a time $\tau > 0$, which

1. is such that the propensity functions $a_j(\boldsymbol{Y}(t))$ within the time window $[t; t+\tau]$

are approximately constant, meaning that the propensity functions can be written as $a_j(\boldsymbol{y}_t)$ where $\boldsymbol{y}_t$ is the system state at time t, the start of the time window, and

2. is so long that a large number of events $Q \gg 1$ per reaction channel j occur in the time interval $[t, t+\tau]$

We note that these two assumptions contradict each other, since the first assumption requires a small τ, while the second assumption requires a large τ. However, we argue, following [69], that for systems with large copy numbers a time τ can be constructed that suffices both assumptions. Following the first assumption, the random variable $K_j(\boldsymbol{y}_t, \tau)$ follow a Poisson distribution since all reactions are independent, $K_j(\boldsymbol{y}_t, \tau) \propto \mathcal{P}(a_j(\boldsymbol{y}_t)\tau)$, where $a_j(\boldsymbol{y}_t)\tau$ is the average number of events in reaction channel j. The second assumption allows us to write the Poissonian random variables as normal random variables, due to the law of large numbers. The mean and variance of these normal variables are equal, since they originate from Poisson random variables. These two assumptions allow us to write

$$Y_i\left(t+\tau\right) = y_{t,i} + \sum_{j=1}^{M} \nu_{ji} \mathcal{N}_j\left(a_j\left(\boldsymbol{y}_t\right)\tau, a_j\left(\boldsymbol{y}_t\right)\tau\right) \tag{1.58}$$

$$= y_{t,i} + \sum_{j=1}^{M} \nu_{ji} a_j\left(\boldsymbol{y}_t\right)\tau + \sum_{j=1}^{M} \nu_{ji} \sqrt{a_j\left(\boldsymbol{y}_t\right)\tau}\mathcal{N}_j\left(0,1\right), \tag{1.59}$$

since $\mathcal{N}_j(\mu, \sigma^2) = \mu + \sigma\mathcal{N}_j\left(0,1\right)$. In a final transition we rewrite $\tau \to dt$, and introduce the noise source $\Gamma_j\left(t\right)$, where $\langle\Gamma_j\left(t\right)\rangle = 0$ and $\langle\Gamma_j\left(t\right)\Gamma_{j'}\left(t'\right)\rangle = \delta(t-t')(j-j')$, to obtain the Chemical Langevin Equation (CLE)

$$\frac{dY_i\left(t\right)}{dt} = \sum_{j=1}^{M} \nu_{ji} a_j\left(\boldsymbol{Y}\left(\boldsymbol{t}\right)\right) + \sum_{j=1}^{M} \nu_{ji} \sqrt{a_j\left(\boldsymbol{Y}\left(\boldsymbol{t}\right)\right)}\Gamma_j\left(t\right). \tag{1.60}$$

This leads for our example to

$$\frac{dS\left(t\right)}{dt} = k - \lambda S\left(t\right) + \sqrt{k}\Gamma_0\left(t\right) - \sqrt{\lambda S\left(t\right)}\Gamma_1\left(t\right), \tag{1.61}$$

$$\frac{dX\left(t\right)}{dt} = \rho S\left(t\right) - \mu X\left(t\right) + \sqrt{\rho S\left(t\right)}\Gamma_2\left(t\right) - \sqrt{\mu X\left(t\right)}\Gamma_3\left(t\right). \tag{1.62}$$

The CLE provides us with a clear description of the time-dependent noise strength (term(s) involving Γ_j), which not necessarily is linear. The CLE (Eq. 1.60) implies a corresponding

(forward) Fokker-Planck equation for the multivariate probability distribution $p(S, X)$ [69]

$$\begin{aligned}\frac{\partial p\left(\boldsymbol{y}, t|\boldsymbol{y}_0, t_0\right)}{\partial t} = &-\sum_{i=1}^{N}\frac{\partial}{\partial y_i}\left[\left(\sum_{j=1}^{M}\nu_{ji}a_j\left(\boldsymbol{y}\right)\right)p\left(\boldsymbol{y}, t|\boldsymbol{y}_0, t_0\right)\right]\\ &+\frac{1}{2}\sum_{i=1}^{N}\frac{\partial^2}{\partial y_i^2}\left[\left(\sum_{j=1}^{M}\nu_{ji}^2a_j\left(\boldsymbol{y}\right)\right)p\left(\boldsymbol{y}, t|\boldsymbol{y}_0, t_0\right)\right]\\ &+\sum_{i,i'=1,i<i'}^{N}\frac{\partial^2}{\partial y_i y_{i'}}\left[\left(\sum_{j=1}^{M}\nu_{ji}\nu_{ji'}a_j\left(\boldsymbol{y}\right)\right)p\left(\boldsymbol{y}, t|\boldsymbol{y}_0, t_0\right)\right], \quad (1.63)\end{aligned}$$

To continue, we make, following [70, 77], two additional assumptions. The first we describe here, while the second is given below Eq. 1.64. The first assumption is that the ensemble average is independent of time: $\langle S(t)\rangle = \langle S\rangle$. We derive the time-dependence of the stationary (co)variances [70, 78, 79],

$$\frac{d\mathbf{C}}{dt} = \mathbf{AC} + \mathbf{C}^T\mathbf{A}^T + \mathbf{B}, \quad (1.64)$$

where $\mathbf{C}$ is the covariance matrix with entries $C_{ij} = \sigma_{ij}^2$, $\mathbf{A}$ is the Jacobian matrix for the dynamics of the average process, with entries $A_{ij} = \partial/\partial y_i[(\Sigma_j\nu_{ji}a_j(\langle\boldsymbol{y}\rangle))]$, and $\mathbf{B}$ a matrix which describes the diffusion (or random) terms, that depend on the stochastic events in the process, with entries $B_{ik} = \Sigma_j\nu_{ji}\nu_{jk}a_j(\langle\boldsymbol{y}\rangle)$. Importantly, in the derivation of both $\mathbf{A}$ and $\mathbf{B}$ we assume that at steady-state fluctuations around the steady-state can be ignored, which implies

$$\langle SX\rangle = \langle S\rangle\langle X\rangle. \quad (1.65)$$

We stress however that for the calculation of the (co)variances, Eq. 1.65 is not used. This assumption is in general not valid for non-linear systems and application of the LNA therefore should be done with care. However, it is well-known that the Linear-Noise Approximation is accurate as long as the copy numbers $> \mathcal{O}(10)$ ([49] and Fig. 1.5). Since both $\mathbf{A}$ and $\mathbf{B}$ now directly follow from the description of the process, Eq. 1.64 directly allows for the calculation of the variance of every component around steady-state. For our example (Eq. 1.32) we have

$$\mathrm{A}=\begin{pmatrix}-\lambda & 0\\ \rho & -\mu\end{pmatrix}, \quad \mathrm{B}=\begin{pmatrix}2\lambda\langle S\rangle & 0\\ 0 & 2\mu\langle X\rangle\end{pmatrix}. \quad (1.66)$$

This leads to the (co)variances as given in Eq. 1.36 which is as expected, since our example system is linear. However, for non-linear systems, the LNA is not necessarily correct, as is shown in Fig. 1.5.

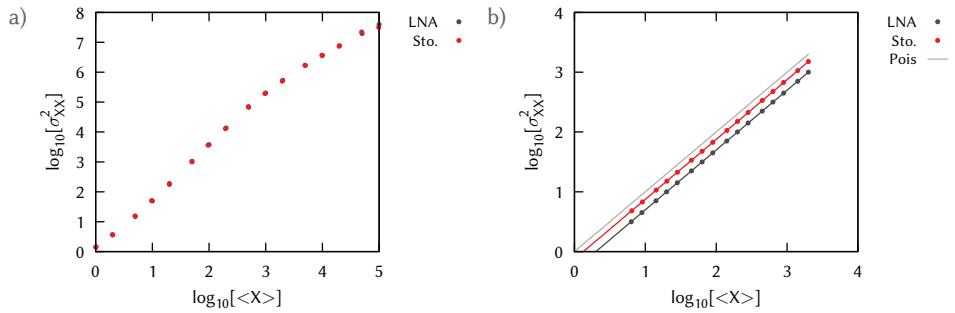


Figure 1.5: The Linear-Noise Approximation (LNA) (black circles) provides accurate predictions for the variance compared with full stochastic simulations (red circles). **a)** In a system, equal to Eq. 1.32, but with non-linear (cooperative) production of X $S^2 \xrightarrow{\rho} S^2+X\emptyset$, where S^2 is a dimer of two S molecules. The LNA (black circles) captures the noise levels in X very good at all copy numbers compared to the full stochastic system (red circles). Parameters: k=10 s^{-1}, λ=1 s^{-1}, ρ=20 s^{-1}, μ is varied to change the average response. **b)** In a birth-death system with non-linear (cooperative) degradation $\emptyset \xrightarrow{k} X$, $X+X \xrightarrow{\mu} \emptyset$, the LNA (black circles) slightly underestimates the noise levels in X compared to the full stochastic system (red circles). Lines show analytical solutions (notes by Andrew Mugler). The use of the LNA should therefore be done with care. Parameters: k=400 s^{-1}, μ is varied to change the average response.

1.4 Numerical optimization

In this last in-depth section I describe two methods of computational optimization, which share many features, but are also somewhat different. Both methods are computational tools and are intensively used in these thesis, which merit a larger background description. But first a small general description as to why we use numerical optimization techniques. In many of the problems that are studied in this thesis the capacity to perform a specific function in a small minimal systems is assessed, like acting as a logic gate (Chapter 5) or multiplexing biochemical signals (Chapters 7 and 8). Although the studied systems are small, they comprise already a large parameter space (more than 10 parameters), which makes it difficult to, from scratch, analytically derive results. However, numerical results can be obtained by performing computational optimization techniques, where numerically the full parameter-space is explored and solutions to the problem at hand are returned. Ideally, one would like to perform brute-force calculations, scanning over all parameter space, but this, given the size of the parameter-space, is in generally unfeasible. Therefore, more elegant algorithms have been developed.

In general, in a multi-dimensional parameter space, the performance function (e.g. the capacity to multiplex) is a highly non-trivial function of all the parameters, showing multiple maxima (peaks) and minima (valleys). In both algorithms a system is a specific point in this parameter space, and each system is characterized by its performance function. The main problem of the optimization algorithms used in this thesis, simulated annealing and numerical evolution, is that they not necessarily obtain the global minimum. Indeed, both algorithms can optimize the system for a local minimum, but not necessarily the global

minimum. Luckily, in many studies in this thesis, I am not necessarily interested in the global optimum, but in a sufficient optimum. In other words, for me it is satisfactory if a specific system can perform a specific function, but not if it is the best system for the specific function.

1.4.1 Simulated Annealing

The first algorithm is simulated annealing (SA). In this method, each point p of the parameter space is analogous to a state of some physical system, and the function $E(p)$ (the performance function) to be minimized is analogous to the internal energy of the system in that state. The goal is to bring the system, from an random initial state in parameter space, to the state with the minimum possible energy (or lowest performance function). At each step, the SA algorithm compares its current state p with a neighboring point p', by comparing the function $E(p)$ and $E(p')$. Based upon the comparison of these two states, two options are possible:

1. If $E(p') < E(p)$ the new state p' is accepted as a more optimal state
2. If $E(p') > E(p)$ the new state p' is accepted following a probabilistic measure, where the likelihood to accept is depends on $\exp(-\beta(E(p') - E(p)))$, where β is an inverse "temperature".

The idea is that, due to the probabilistic acceptance of less optimal states, the simulated annealing algorithm is able to overcome local minima and find the global minimum. The inverse temperature β, which is increased in time, allows for variation of the acceptance of large moves in the uphill direction, thereby stimulating or preventing the algorithm to leave the (local) minimum. The SA algorithm is used in Chapter 7.

1.4.2 Numerical Evolution

The second algorithm is Numerical Evolution. This algorithm in a sense follows the laws of natural evolution for optimization. Instead of the SA algorithm, which modeled a single system, this algorithm models the evolution of a population. The algorithm is based upon ideas of Wright and Fisher [80, 81]. Evolution occurs in discrete, synchronous steps, where the population size remains constant. At each step, each member of the population produces offspring in proportion to its fitness. Then, mutations occur, and the mutated offspring yield the population for the next step.

In general, we have a "population" of S initial systems s (parameter points in the parameter space). Each point has fitness $E(s)$, and the total fitness for the population is $F = \Sigma_s E(s)$. At each step, S new systems ("offspring") are drawn from the distribution $p_s = E(s)/F$, which weights each system by its fitness. Each new system is then "mutated" by changing a single parameter p in the parameter space, at which point selection starts again. Since initially the parameter space is randomly sampled and the assignment of offspring is probabilistic, one expects that the NE algorithm is less prone to converge into a local minimum; however, this is not guaranteed. More background information on this algorithm can be found in [82]. The NE algorithm is used in Chapters 5 and 8.

1.5 Scope of this thesis

The work in this thesis focuses on signal transmission in biochemical networks, with a strong emphasis on dynamical signals. In Chapters 2 and 3 we study the influence of network topology on the information transmission capacity of a specific network. Many recurring motifs are observed experimentally in signaling networks. We study the effect of each motif on the transmission of time-varying signals, the gain, noise and the information rate, which is proportional to the ratio of the signal-to-noise, all as a function of the frequency of the incoming signal. In Chapter 2 we focus on networks with feedback loops and autoregulation, while in Chapter 3 we study networks with a feed-forward architecture. These chapters give insight in how each motif transmits information as a function of frequency. In Chapter 4 we switch gears. Instead of studying the information transmission, we study the precision by which a cell can estimate the concentration by monitoring the occupancy state of a receptor. Indeed, this is an interesting problem; what is the fundamental lower limit on the precision of measuring chemical concentrations by time-averaging? The answer to this question is provided in a seminal paper written by Howard Berg and Edward Purcell [83], but since then however disputed in other publications. In this chapter, together with our co-workers from Japan, we reconsider this fundamental limit using a new analytical approach as well as computer simulations. In Chapter 5 we focus on actual information processing by single molecules. In recent years it has been shown that logic operations can be performed by single molecules [84]. We wonder whether a receptor, by variation of kinetic parameters, can perform all possible logic operations? While this question is interesting on evolutionary timescales, we then focus down onto signaling timescales. Is it possible for cells, by recombination of receptor monomers into dimers, to access all different logic operations. In Chapter 6 we return to one of the original observations, namely that signals are time-varying. This by itself is remarkable in the context of variability reduction, since intuitively one would expect that oscillating signals tend to increase the variability in a response, compared to constant signals. In Chapter 6 we study this question in more detail for a simple model of gene-regulation, pinning down the question whether indeed oscillating signals always increase the variability in a response as we naively expect.

The last two chapters are reserved for the topic reflected in the title of this thesis: Multiplexing Biochemical Signals. It is commonly observed that within cells different signals share a common signaling pathway. In other words, independent signals are transmitted through a common pathway, but ultimately they lead to their own unique response. This raises an interesting question: *can cells simultaneously transmit signals through a common signaling network and yet, respond to each signal uniquely?* We can think of this hypothesis with respect to the telephone system. In this telephone system multiple calls are transmitted through the same wire, but every independent caller-receiver combination can have a lively conversation without interference from any of the other calls in the network. In Chapter 7 we study the idea of multiplexing biochemical signals for signals that are constant in time. We discuss an encoding and decoding strategy that a cell could employ that allows for multiplexing, with absolute fidelity in the signal transmission for each signal. In Chapter 8 we extend the ideas of Chapter 7 by looking not only at signals that are constant in time, but also by studying signals that vary in time. Again we focus on possible encoding

and decoding strategies that allow multiplexing of biochemical signals. We study a possible multiplexing network in more detail, especially looking at the influence of interference from the different signals onto each other, and the influence of noise on the information transmission.

Chapter 2

Effect of feedback on the fidelity of information transmission of time-varying signals

Living cells are continually exposed to environmental signals that vary in time. These signals are detected and processed by biochemical networks, which are often highly stochastic. To understand how cells cope with a fluctuating environment, we therefore have to understand how reliably biochemical networks can transmit time-varying signals. To this end, we must understand both the noise characteristics and the amplification properties of networks. In this chapter, we use information theory to study how reliably signaling cascades employing autoregulation and feedback can transmit time-varying signals. We calculate the frequency-dependence of the gain-to-noise ratio, which reflects how reliably a network transmits signals at different frequencies. We find that the gain-to-noise ratio may differ qualitatively from the power spectrum of the output, showing that the latter does not directly reflect signaling performance. Moreover, we find that auto-activation and auto-repression increase and decrease the gain-to-noise ratio for all of frequencies, respectively. Positive feedback specifically enhances information transmission at low frequencies, while negative feedback increases signal fidelity at high frequencies. Our analysis not only elucidates the role of autoregulation and feedback in naturally-occurring biological networks, but also reveals design principles that can be used for the reliable transmission of time-varying signals in synthetic gene circuits.

Based on manuscript W.H. de Ronde, F. Tostevin and P.R. ten Wolde (**2010**) Effect of feedback on the fidelity of information transmission for time-varying signals. *Physical Review E* **82**. doi:10.1103/PhysRevE.82.031914

2.1 Introduction

Living cells constantly have to respond and adapt to a changing environment. In some cases, such as in response to a changing sugar concentration [85], a cell may wish to integrate out rapid variations and only respond to slow variations of the environmental signal, while in other cases, such as osmo adaptation [86] or bacterial chemotaxis [87], the cell needs to do the opposite — respond to rapid but not slow variations (adaptation). Indeed, to understand how cells cope with a fluctuating environment, we have to understand how cells transduce time-varying signals. Cells detect, process, and transduce signals via biochemical networks, which are the information processing devices of life. However, experiments in recent years have demonstrated that biochemical networks are often highly stochastic [25, 88]. This raises the question how reliably biochemical networks can transmit time-varying signals in the presence of noise.

Interestingly, biochemical networks exploit commonly recurring architectures [10, 89], such as autoregulation, cascades, and feedback, to process signals. These network motifs often implement signal amplification in order to raise the level of the input signal relative to the noise. Amplification can be characterized by the *gain*, the fold-change in the signal amplitude. However, it is important to recognize that such amplification can not only increase the levels of the desired signal, but can also amplify the noise itself. Therefore, to understand the possibilities and limitations of different network motifs for enhancing the fidelity of signal transduction, we need to understand how both the signal and the noise are propagated through these motifs. Specifically, information theory indicates that the reliability of signal transmission is determined by the ratio of the gain of the network to the total noise in the output signal — the gain-to-noise ratio. Moreover, to assess how reliably signals of different temporal characteristics are transduced, we have to understand the frequency dependence of the gain and the noise. Importantly, we expect that different network architectures will affect the frequency-dependence of the gain and the noise differently, which means that we have to study both these quantities. In this chapter, we study the frequency-dependence of the gain-to-noise ratio for simple cascades, and for cascades employing autoregulation and feedback. This allows us to elucidate how autoregulation and feedback can shape the frequency range over which signals can be transduced reliably.

Information theory provides a formalism for quantifying the reliability of information transmission in the presence of noise [45]. A natural measure for the fidelity of signal transmission from an input signal S, with copy number S to an output signal X (the network response), with copy number X, is the mutual information between S and X, which is defined as

$$I(S,X) = H(S) - H(S|X) = \\ -\int dS\, p(S) \log\left[p(S)\right] - \left(-\int dX\, p(X) \int dS\, p(S|X) \log\left[p(S|X)\right]\right). \quad (2.1)$$

Here, $p(S)$ and $p(X)$ are the probability distributions of possible input and output signals respectively, and $p(S|X)$ is the conditional probability of S once X is specified. The mutual information quantifies the reduction in entropy of (or uncertainty about) the signal

after one obtains knowledge of the network response, averaged over all possible responses. In other words, $I(S,X)$ is how much we learn (on average) about S by measuring X. For a deterministic system, every S leads to a unique X (we assume no degeneracy). Measuring X thus precisely specifies S, such that the uncertainty in S after a measurement of X is $H(S|X) = 0$ and $I(S,X) = H(S)$. However, in the presence of noise in the network each input S will lead to a distribution of possible outputs X. As a result, an observed X can correspond to multiple S values and $I(S,X) \leq H(S)$. For completely uncorrelated S and X, $I(S,X) = 0$. By construction, the mutual information is symmetric, such that $I(S,X) = I(X,S)$.

Recently, the mutual information has been used to study the reliability of information transmission in biochemical networks [49, 50, 51, 52]. However, these studies considered only the steady-state response of a network to a distribution of *constant* input signals, which do not change on the timescale of the network response. Yet, in many biological systems, it cannot be assumed that the input signal is constant on the timescale of the network response.

Indeed, in many systems the message is encoded in the *temporal dynamics* of the input signal. A well-known example is bacterial chemotaxis, where the concentration of the intracellular messenger protein depends not on the steady-state ligand concentration, but rather on the change of this concentration in the recent past [90] — the response of the network thus depends on the history of the input signal. Moreover, the extracellular signal may be encoded in the temporal dynamics of the intracellular signal transduction pathway. An interesting example is provided by the rat PC-12 system: while stimulation with a neuronal growth factor gives rise to a sustained response of the Raf-Mek-Erk pathway, stimulation by an epidermal growth factor gives rise to a transient response of this pathway [91]. In all these cases, the message is encoded not in the concentration of some chemical species at a given moment in time, but rather in its concentration as a function of time. This means that to understand how reliably the network can transmit information, we need to know how accurately an input signal as a function of time — the input *trajectory* $s(t)$ — can be mapped onto an output trajectory $x(t)$. We thus need to understand the mutual information between the two trajectories, $I[s(t), x(t)]$.

The ability of a biochemical network to transduce a time-varying input signal depends on the correlation time of the input signal and the architecture and response dynamics of the network. An instructive example is provided by the chemotaxis network of the bacterium *E. coli*. This network employs integral negative feedback [92], as a result of which the intracellular messenger protein can adapt to a constant extracellular ligand concentration. This means that the signaling network cannot respond to changes in ligand concentration that occur on time scales longer than the adaptation time. At the other end of the frequency spectrum, changes in the messenger protein that occur on time scales shorter than the motor switching time will be integrated out; indeed, the network cannot respond reliably to rapidly varying input signals [65]. The architecture and the response dynamics of the processing network thus determines the frequency range over which signals can be transduced reliably.

Recently, we have applied information theory to biochemical networks and studied the mutual information between in- and output trajectories, $I[s(t), x(t)]$ [65]. Here, we apply

this framework to study the propagation of time-varying signals through a number of network motifs—cascades, autoregulation, and feedback. It is known that for *constant* signals (or, to be more precise, signals that do not vary on the time scale of the network response time), the mutual information decreases as a function of cascade length [50]. The same also holds true for time-varying signals. Indeed, the data-processing inequality states that in a cascade with n nodes, the information about the input encoded in the signal at node $i+1$ cannot be greater than the information at node i. Once lost, information about the input cannot be recovered later in the cascade. Simply increasing the length of a signaling cascade therefore can never increase the transmitted information. Conversely, maximizing the total transmitted information cannot be the driving force behind the evolution of such cascades.

Cascades, however, often employ autoregulation and feedback, which can be used to shape the response of the network to signals of different frequencies. Importantly, autoregulation and feedback affect not only the frequency-dependent gain, which describes how strongly an input signal at a particular frequency is amplified in the absence of any biochemical noise, but also the frequency-dependence of the noise. While the frequency-dependence of the gain [30, 93, 94] and the noise [95] have been studied separately, the frequency-dependence of their ratio, the gain-to-noise ratio, has not. However, it is the gain-to-noise ratio which determines how reliably an input signal at a particular frequency can be transmitted [65]. In fact, as we will show, autoregulation and feedback affect the frequency-dependence of the gain and the noise differently, which means that it is essential to study these quantities together.

In this chapter, we study the frequency-dependent gain-to-noise ratio using a Gaussian model. In the next section, we describe this model, and how we can use it to compute the frequency-dependent gain-to-noise ratio and the information transmission rate, which is given by the integral of this ratio over all frequencies [65]. In the section: *Results* we discuss the frequency-dependent gain-to-noise ratio of simple cascades, and cascades employing feedback and autoregulation. Our results highlight the idea that the output power spectrum is not a direct measure for the information content of the output signal—the output power spectrum can differ *qualitatively* from the spectrum of the gain-to-noise ratio. We also show (Fig. 2.1a) that positive regulation tends to increase the gain-to-noise ratio, while negative regulation tends to decrease it. Moreover, we show that the frequency spectra of motifs with negative feedback can exhibit windows in which the gain-to-noise ratio is increased; these motifs can thus act as band-pass filters for information transmission. Finally, we discuss some of the implications of our findings and the limitations of our analysis.

2.2 Methods

We consider information transmission through a biochemical network from an input signal $s\left(t\right)$ to an output signal $x\left(t\right)$. The dynamics of the network can be described mathematically by a set of coupled Langevin equations [69] for the signal, response and an arbitrary number of intermediate components v_i, in vector form $\boldsymbol{v}$. In using the Langevin representation we assume that the copy number of each component is large such that the discrete number of

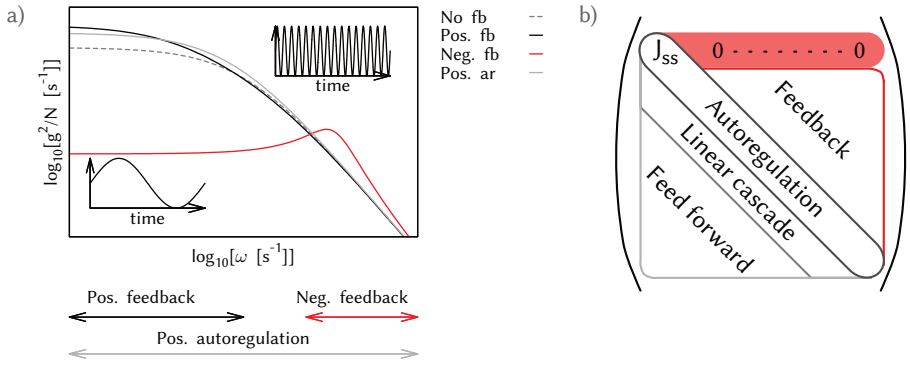


Figure 2.1: a) A schematic drawing of the main conclusions of this chapter. The frequency of the variations of the input is shown on the x-axis. For three different motifs the gain-to-noise ratio is shown. The arrows indicate the specific frequency regime for which each motif performs better with respect to a simple cascade (fb is feedback, ar is autoregulation). **b)** The Jacobian matrix. The entries of the Jacobian matrix encode the structure of the reaction network.

molecules can be approximated by a continuous concentration, and we obtain

$$\frac{ds}{dt} = f_s^+(s) + f_s^-(s) + \Gamma(t), \tag{2.2}$$

$$\frac{d\boldsymbol{v}}{dt} = \boldsymbol{f}_v^+(s,\boldsymbol{v},x) - \boldsymbol{f}_v^-(s,\boldsymbol{v},x) + \boldsymbol{\eta}_v(t), \tag{2.3}$$

$$\frac{dx}{dt} = f_x^+(s,\boldsymbol{v},x) - f_x^-(s,\boldsymbol{v},x) + \eta_x(t). \tag{2.4}$$

Here, f_i^+ and f_i^- contain all the reactions involving the production and degradation of component i, respectively. f_s^+ and f_s^- only depend on s, so that we restrict our analysis to networks that do not feed back onto s itself. In these cases, the gain-to-noise ratio is independent of the input signal [96], as discussed in more detail below Eq. 2.11. $\Gamma(t)$ is a stochastic driving process that serves to define the ensemble of possible input signals. The various noise sources η_i are taken to be independent and Gaussian-distributed [31, 77, 96], such that $\langle \eta_i(t)\,\eta_j(t')\rangle = \langle|\eta_i|^2\rangle\delta_{ij}\delta(t-t')$. Here, we note that the assumption of independent noise sources is only made to simplify the analysis. (Anti)-correlations between noise sources can affect noise propagation [96], and can be included by a straightforward extension of the present discussion. Furthermore, we assume that $\langle|\eta_i|^2\rangle = \langle f_i^+\rangle + \langle f_i^-\rangle = 2\langle f_i^+\rangle$ and $\langle|\Gamma|^2\rangle = 2\langle f_s^+\rangle$ [97], the sum of the production and degradation terms.

We introduce the vector $\boldsymbol{y} = (s;\boldsymbol{v};x)$ and $\boldsymbol{\eta} = (\Gamma, \boldsymbol{\eta}_v, \eta_x)$ and assume the network has a steady state $\langle\boldsymbol{y}\rangle$. Linearizing around this steady state we obtain

$$\frac{d\tilde{\boldsymbol{y}}}{dt} = \mathbf{J}|_{\boldsymbol{y}=\langle\boldsymbol{y}\rangle}\tilde{\boldsymbol{y}} + \boldsymbol{\eta}. \tag{2.5}$$

Here $\tilde{y}_i = y_i - \langle y_i\rangle$ is the deviation of the concentration of component i from its steady-state

value, $\langle y_i \rangle$, and $\mathbf{J}$ is the Jacobian evaluated at the steady state[1]. J_{ij} describes the response of the component i to small changes in component j, while keeping all other components at their steady-state levels. The diagonal element $J_{ii} = -\tau_i^{-1}$ is the relaxation time or dissipative time scale of component i; it describes the time scale on which component i relaxes back to its steady-state value after a perturbation. After linearization, the architecture of the network is encoded in the structure of the Jacobian matrix (see Fig. 2.1b): the diagonal terms correspond to autoregulation, the lower triangular part to downstream (feedforward) regulation and the upper triangular part to upstream (feedback) regulation. Since we restrict ourselves to systems without feedback from the network to the signal itself, we require that all elements on the first row of $\mathbf{J}$ are zero but for J_{ss}.

We take as our input signal the variations $\tilde{s}$. A linear system does not change the frequency of the transmitted signal, but only the amplitude and the phase. Since Eq. 2.5 is linear in $\tilde{\boldsymbol{y}}$, we can calculate exactly the power spectra of the network components [97],

$$\mathbf{P} = [i\omega\mathbf{I} - \mathbf{J}]^{-1}\,\mathbf{Q}\left[-i\omega\mathbf{I} - \mathbf{J}^T\right]^{-1}, \tag{2.6}$$

where $P_{ij}(\omega) = \langle \tilde{Y}_i(\omega)\tilde{Y}_j(-\omega)\rangle$ is the (cross-)power spectrum of $\tilde{y}_i$ and $\tilde{y}_j$, $\tilde{Y}_i(\omega)$ is the Fourier transform of $\tilde{y}_i(t)$, $\mathbf{I}$ is the identity matrix, and $\mathbf{Q}$ is the noise matrix with entries $Q_{ij} = \langle \eta_i(\omega)\eta_j(-\omega)\rangle = \langle |\eta_i|^2\rangle\delta_{ij}$. The power spectrum is a commonly used tool to study time-varying signals, and describes how the total power of a signal is distributed over different frequencies. Power at low frequencies is related to slow variations of the signal, while power at high frequencies corresponds to rapid fluctuations. The integral of the power spectrum over all frequencies equals the total variance of the signal.

The information transmission rate for time-varying signals is [48, 61]

$$\lim_{T\to\infty} \frac{I\left[s\left(t\right), x\left(t\right)\right]}{T} = R\left[s\left(t\right), x\left(t\right)\right] = -\frac{1}{2\pi}\int_0^\infty d\omega \ln\left[1 - \Phi_{sx}(\omega)\right], \tag{2.7}$$

where T is the length of the trajectory and $\Phi_{sx}(\omega)$ is the coherence function, defined as

$$\Phi_{sx}(\omega) = \frac{|P_{sx}(\omega)|^2}{P_{ss}(\omega)P_{xx}(\omega)}. \tag{2.8}$$

$\Phi_{sx}(\omega)$ is a measure of the average correlation between the in- and output signals in the frequency domain. For completely independent in- and output signals, $\Phi_{sx}(\omega) = 0$, while for a noiseless system $\Phi_{sx}(\omega) = 1$.

The power spectrum of the output signal, $P_{xx}(\omega)$, can be decomposed as

$$P_{xx}\left(\omega\right) \equiv \Sigma\left(\omega\right) + N\left(\omega\right) \equiv g^2\left(\omega\right) P_{ss}\left(\omega\right) + N\left(\omega\right). \tag{2.9}$$

Here, $\Sigma(\omega) \equiv g^2\left(\omega\right) P_{ss}(\omega)$ is the transmitted signal, $g^2(\omega) \equiv |P_{sx}(\omega)|^2 / P_{ss}^2(\omega)$ is the frequency-dependent gain, $P_{ss}(\omega)$ is the power spectrum of the input signal and $N(\omega)$ is

[1] We note that with the chosen expression for the size of the random events $\langle \eta_i\eta_j\rangle$, the linearized Langevin equations lead to the Fluctuation-Dissipation theorem or Linear Noise Approximation [68, 70]

the frequency-dependent noise. With these definitions, the coherence function, Eq. 2.8, can be recast as

$$\Phi_{sx}(\omega) = \frac{\Sigma(\omega)}{N(\omega) + \Sigma(\omega)}, \tag{2.10}$$

and the mutual information rate can be rewritten as [65]

$$R\left[s\left(t\right), x\left(t\right)\right] = \frac{1}{2\pi}\int_0^\infty d\omega \ln\left[1 + \frac{g^2(\omega)}{N(\omega)} P_{ss}(\omega)\right]. \tag{2.11}$$

We see that the information transmission rate depends on the power spectrum of the input signal, $P_{ss}(\omega)$, and on the gain-to-noise ratio $g^2(\omega)/N(\omega)$.

As discussed in Ref. [96], in a biological system the reaction that detects the input signal can, depending on the nature of the detection reaction, introduce significant correlations between the variations in the input signal and the intrinsic noise of the reactions that constitute the processing network. These correlations are a consequence of the molecular character of the components and are thus unique to biochemical networks. If the detection reaction does not introduce correlations, then Eq. 2.9 is the spectral-addition rule [96]. The noise $N(\omega)$ is then the intrinsic noise of the processing network and also $g^2(\omega)$ only depends on properties of the processing network. On the other hand, if the detection reaction does introduce correlations, then the output power spectrum $P_{xx}(\omega)$ can be written in the form of Eq. 2.9, but then $N(\omega)$ and $g^2(\omega)$ depend not only on characteristics of the processing network, but also on the statistics of the input signal; conversely, the variations of the input will also be affected by the noise in the processing network [96]. In what follows below, we assume for simplicity that the spectral-addition rule holds, which means that the gain, noise and gain-to-noise ratio are independent of the input signal, and that the input does not need to be specified.

Applying the linearization procedure, as outlined above, may, in general, qualitatively change the dynamics of the network being considered. However, previous studies [32, 49] have shown that the Linear Noise Approximation provides an accurate description of many systems if the average copy numbers are of order 10 molecules or more. For the networks considered in this chapter we also compared the power spectra calculated in the linear approximation with the results of stochastic simulations performed with Gillespie's algorithm [74], and again found good agreement when protein copy numbers are large (see App. 2.A.1). We therefore expect that the linear analysis presented in this chapter provides an accurate description of the signaling characteristics of these networks.

2.3 Results

First we study a simple cascade, where "simple" means that we consider a cascade where each component only regulates the activity of the next component in the cascade; a "simple" cascade is thus a cascade without autoregulation, feedback or feedforward interaction. We analyze this network in detail such that it can serve as an instructive example of the method described above. In addition, we will highlight general features of the results which recur

in more complex networks. We then discuss network motifs including autoregulation and negative feedback loops, which are commonly observed in biochemical networks.

To understand the effects of autoregulation and feedback we will compare information transmission in these motifs to a corresponding simple cascade with the same number of components but without the additional regulation. In order to perform such a comparison of different motifs on an equal footing we constrain the average production rate of every component such that these are equal in the networks under comparison. We argue that from a biological perspective the rate of protein production is a more significant constraint on network design than average protein copy number, since the latter only depends on the *ratio* of the synthesis and degradation rate, while it is the absolute synthesis and degradation rate that determines the cost of having a protein at a particular copy number. This constraint also enforces that the noise strength at each level of the cascade $\langle|\eta_i|^2\rangle = 2\langle f_i^+\rangle$ is the same in the motifs being compared. When comparing two systems with many parameters, equalizing production rates is not a sufficient constraint to uniquely specify all parameter values. To reduce this potential parameter space we will (unless otherwise stated) hold constant as many of the network parameters as possible. For brevity we will only discuss networks in which all regulation occurs via the production reactions, with linear degradation of each component. However, our results are qualitatively unchanged if we instead consider regulation via protein degradation.

We characterize information transmission through these motifs in terms of the gain, noise and gain-to-noise ratio. Since we assume that the spectral-addition rule holds [96], these quantities are intrinsic, signal-independent properties of the network. We also wish to highlight differences between the information transmission characteristics of the network, as determined by the gain-to-noise ratio, and the output power spectrum $P_{xx}(\omega)$, since this is commonly discussed in studies of signal transmission. Since $P_{xx}(\omega)$ depends not only on the processing network but also on the input signal (see Eq. 2.9), we must therefore specify $P_{ss}(\omega)$; for this purpose we assume, for convenience, that the input signal $s(t)$ is generated via a Poisson birth-death process as in Eq. A2.9 (the section: *The simple cascade*).

2.3.1 The simple cascade

Initially we study a simple cascade with a single intermediate component. Extension of the cascade with more intermediate components is straightforward. The appropriate reaction scheme is

$$\frac{ds}{dt} = k_s - \mu_s s + \Gamma(t), \tag{2.12}$$

$$\frac{dv}{dt} = k_v s - \mu_v v + \eta_v(t), \tag{2.13}$$

$$\frac{dx}{dt} = k_x v - \mu_x x + \eta_x(t). \tag{2.14}$$

We reiterate that we assume that there are no cross correlations in the noise; $\langle\eta_\alpha(t)\eta_\beta(t')\rangle = \delta_{\alpha\beta}\delta(t-t')$ and $\langle\Gamma(t)\,\eta_\alpha(t')\rangle = 0$. This means that the reactions are of the type $s \to s+v$ and $v \to v+x$, and not $s \to v$ and $v \to x$, respectively; put differently, the firing of a reac-

tion does not consume a molecule of the reactant, and hence does not affect the fluctuations of the up-stream component [96]. In the section: *Discussion*, we will briefly address some of the limitations of this assumption.

Fourier transformation gives

$$\tilde{X}(\omega) = \overbrace{\frac{k_x k_v \tilde{S}}{(i\omega+\mu_x)(i\omega+\mu_v)}}^{\text{signal}} + \overbrace{\frac{k_x \eta_v(\omega)}{(i\omega+\mu_x)(i\omega+\mu_v)} + \frac{\eta_x(\omega)}{i\omega+\mu_x}}^{\text{noise}}. \tag{2.15}$$

As indicated, we can identify the components of the output which are due to the input $\tilde{S}$ ("signal") and components which are due to intrinsic noise in the network. We obtain for the power spectrum of x,

$$P_{xx}(\omega) = \left\langle \tilde{X}\tilde{X}^* \right\rangle = \overbrace{\frac{k_x^2}{(\omega^2+\mu_x^2)}\frac{k_v^2}{(\omega^2+\mu_v^2)}}^{g^2(\omega)} \overbrace{\frac{2k_s}{(\omega^2+\mu_s^2)}}^{P_{ss}(\omega)} + \overbrace{\underbrace{\frac{k_x^2}{(\omega^2+\mu_x^2)}\frac{2k_v\langle s\rangle}{(\omega^2+\mu_v^2)}}_{N_{v\to x}(\omega)} + \underbrace{\frac{2k_x\langle v\rangle}{(\omega^2+\mu_x^2)}}_{N_x(\omega)}}^{N(\omega)} \tag{2.16}$$

Figure 2.2a shows the output power spectrum of this network $P_{xx}(\omega)$ (dark red solid), as well as its decomposition into the noise $N(\omega)$ (dark red dashed) and the transmitted signal $\Sigma(\omega) = g^2(\omega)P_{ss}(\omega)$ (black dashed) (see also Eq. 2.9). Simple cascades are characterized by a number of "knee" frequencies (vertical dashed), corresponding to the characteristic relaxation rates of the different components of the network (in this case μ_s, μ_v and μ_x). These knee frequencies are the inverse of the response times of the components — e.g. $\mu_v = \tau_v^{-1}$.

In order for the processing network to track variations in the input s on a time scale ω^{-1}, the network should be able to respond on this time scale. If any component of the processing network has a longer response time, this variation in s will be filtered. This filtering can be observed in the transmitted signal $\Sigma(\omega)$, where at frequencies above the first knee frequency, $\Sigma(\omega)$ scales with ω^{-2} and for every consecutive knee frequency, $\Sigma(\omega)$ decays with an additional factor ω^{-2} (Fig. 2.2a). In effect each level of the cascade acts as a low-pass filter, because the incoming signal is averaged over the protein response time. Mathematically, the transmitted signal $\Sigma(\omega)$ can be factored into the input signal $P_{ss}(\omega)$ (black solid), and the total gain $g^2(\omega)$ (Fig. 2.2b, gray solid), which is independent of the input signal (because we assume that the network does not feed back onto s). Moreover, the total gain of the network is the product of the gain of each cascade step: $g^2(\omega) = g^2_{s\to v}(\omega)g^2_{v\to x}(\omega)$; decaying as ω^{-4} for $\omega \gg \mu_v, \mu_x$ (Fig. 2.2b). Consequently, the transmitted signal $\Sigma(\omega)$ decays as ω^{-6} for $\omega \gg \mu_s, \mu_v, \mu_x$.

Since we assume that there are no cross-correlations between the different noise terms, the total noise $N(\omega)$ (dark red dashed in Fig. 2.2a,b) is given by the noise-addition rule [70,

96], which means that $N(\omega)$ is simply given by the sum of two independent contributions, $N_{v\to x}(\omega)$ (Fig. 2.2b, red solid) and $N_x(\omega)$ (red dashed) (see Eq. 2.16). Here, $N_x(\omega)$ is the noise in the concentration of x that arises from the intrinsic stochasticity in the production and decay events of x; $N_x(\omega)$ would be the total variance in the concentration of x if v, the input for x, would not vary over time. However, the upstream component v does vary in time, not only because it is driven by variations in the input s, but also because it fluctuates spontaneously due to the noise in its synthesis and decay events. This noise is propagated to x. Its contribution to the total noise power of x is $N_{v\to x}(\omega)$, which is given by the noise in v, $N_v(\omega)$, multiplied by how much this noise is amplified at the level of x, given by $g^2_{v\to x}(\omega)$: $N_{v\to x}(\omega) = g^2_{v\to x}(\omega)N_v(\omega)$, where $g^2_{v\to x} = k_x^2/(\omega^2+\mu_x^2)$. The "extrinsic" contribution to the noise in x, $N_{v\to x}$, decays as ω^{-4} since the noise in v, decaying as ω^{-2}, is filtered by the finite lifetime of the protein x. The "intrinsic" contribution, $N_x(\omega)$, decays as ω^{-2}, meaning that for $\omega \gg \mu_v, \mu_x$, $N(\omega) \approx N_x(\omega)$. Hence, while the transmitted signal $\Sigma(\omega)$ decays as ω^{-6} for $\omega \gg \mu_s, \mu_v, \mu_x$, the noise $N(\omega)$ decays as ω^{-2} (Fig. 2.2b, dark red dashed). As a result, for frequencies $\omega \gg \mu_s, \mu_v, \mu_x$, the transmitted signal $\Sigma(\omega)$ is completely obscured by the noise and the output $P_{xx}(\omega)$ is simply given by the noise $N(\omega)$ (Fig. 2.2a). Finally, the gain-to-noise ratio (Fig. 2.2b, black solid) is

$$\frac{g^2(\omega)}{N(\omega)} = \frac{k_v k_x \mu_v}{2\langle s\rangle\left(\omega^2 + \mu_v^2 + \mu_v k_x\right)}. \tag{2.17}$$

This expression shows that the simple cascade effectively acts as a low-pass filter for information, meaning that it cannot reliably respond to signals that vary (much) faster than a characteristic cut-off frequency $\omega_c^2 = \mu_v\left(\mu_v + k_x\right)$. We note that the gain-to-noise ratio is independent of μ_x, since both the gain and the noise have the same functional dependence on μ_x. This is a general feature of the biochemical networks we will study: degradation of the output species occurs independently of the upstream components, and therefore provides no additional information about the input [65].

2.3.2 Autoregulation

In this section we consider direct feedback of a component onto its own production, as indicated in Fig. 2.3a and Fig. 2.4a. Autoregulation is one of the most common forms of regulation in signaling networks. It is well known that negative autoregulation speeds up the response time of components, which can also change the response time of the complete signaling cascade [33]. Positive autoregulation slows down the response time and can lead to bistability [33, 98].

Autoregulation alters only the diagonal entries of the Jacobian matrix (Fig. 2.1b). This means that the characteristic timescale for dissipation of small fluctuations — the response time — changes, which is as expected. For the steady state of the system to be stable we require that the diagonal of the Jacobian has only negative terms. Thus autoregulation cannot qualitatively change the form of the output power spectrum $P_{xx}(\omega)$. In fact, once linearized, the dynamics of a network with autoregulation is equivalent to that of a simple cascade with a different degradation rate. In terms of information transmission, however, this is not always true, as we shall see below.

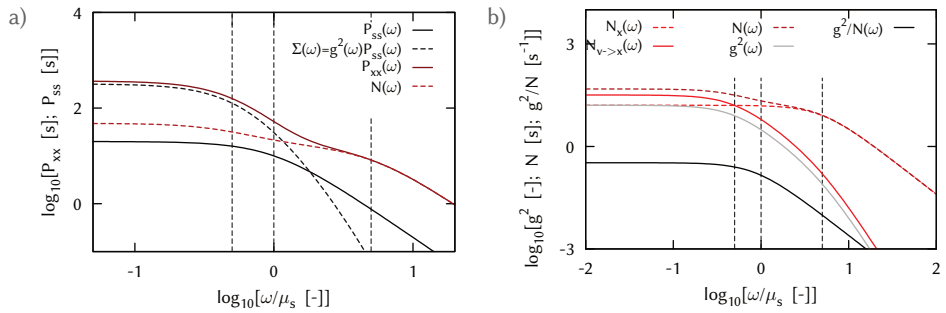


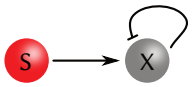
Figure 2.2: Typical power spectra for a linear cascade. **a)** The power spectra of x and s, $P_{xx}(\omega)$ and $P_{ss}(\omega)$, together with the signal $\Sigma(\omega)$ and noise $N(\omega)$ components of the output, for the two-step cascade shown in Eq. 2.12. **b)** The frequency-dependent gain $g^2(\omega)$, noise $N(\omega)$ and gain-to-noise ratio (g^2/N). Red lines indicate the two noise contributions, $N_{v\to x}(\omega)$ (solid) and $N_x(\omega)$ (dashed). Parameters: k_s=10, k_v=10, k_x=1, μ_v=0.5 and μ_x=5. Vertical lines indicate the degradation rates of the three components.

Autoregulation at the response X does not affect information transmission

We first consider autoregulation by the network output X on its own production, as depicted in Fig. 2.3a. For this motif the relaxation time of X is given by $\tau_x = -J_{xx}^{-1} = [\mu_x - \langle \partial/\partial x f(x)s \rangle]^{-1}$, where $f(x)$ describes the effect of the feedback of x onto its own production (see Eq. 2.18 in Fig. 2.3b) For negative regulation $|J_{xx}| > \mu_x$, while for positive regulation $|J_{xx}| < \mu_x$. Negative (positive) regulation therefore reduces (increases) the response time of X to changes in s, compared to the equivalent simple cascade network for which $f(x)$ =constant. In the output power spectrum $P_{xx}(\omega)$ this change in timescale appears as a shift in the knee frequency corresponding to τ_x^{-1}. A corresponding change can also be seen in both the gain and noise (see Eq. 2.20 in Fig. 2.3c).

However, despite these changes in the response time, we find that the gain-to-noise ratio for an autoregulatory network (Eq. 2.20c in Fig. 2.3c) is identical to the gain-to-noise ratio for a simple (two-component) cascade. The effect of changing J_{xx} on the noise and gain is identical (Eq. 2.20a,b in Fig. 2.3b) and therefore cancels in the gain-to-noise ratio (as we also saw previously for the effect of μ_x in the simple cascade, Eq. 2.17). The autoregulation by X of its own production alters the timing of production events. However, our constraint of equal average production means that the mean rate of this process in the two cascades is the same. Moreover, in the linearized regime the production of x is an identical Poissonian process in both simple and autoregulated cascades. Hence, to the extent that the system can be linearized, autoregulation at the output of a network does not affect information transmission. It is conceivable that non-linear effects cause autoregulation of the output component to affect information transmission, but a comparison of our analytical results discussed here with results of Gillespie simulations of the full system, suggest that the linearization approximation is surprisingly accurate (see App. 2.A.1).

a)



b)

$$\frac{dx}{dt} = f(x)\, s - \mu_x x + \eta_x, \tag{2.18}$$

$$f(x) = \frac{\nu\beta}{K + x} \begin{cases} \beta = K, & \text{neg. regulation} \\ \beta = x, & \text{pos. regulation} \end{cases},$$

$$J_{xx} = -\mu_x + \langle s \rangle \left[\frac{\partial f(x)}{\partial x} \right]_{\text{s.s.}}. \tag{2.19}$$

c)

$$g^2 = \frac{J_{xs}^2}{\omega^2 + J_{xx}^2}, \tag{2.20a}$$

$$N = \frac{\left\langle |\eta_x|^2 \right\rangle}{\omega^2 + J_{xx}^2}, \tag{2.20b}$$

$$\frac{g^2}{N} = \frac{J_{xs}^2}{\left\langle |\eta_x|^2 \right\rangle}. \tag{2.20c}$$

Figure 2.3: Autoregulation of the output component. **a)** Schematic representation of the negative autoregulation motif, where s is the input signal and x the output signal, which negatively regulates its own production. **b)** The Langevin equations of the network. **c)** The characteristic equations for the gain, noise and gain-to-noise ratio (see also the section: *Autoregulation*).

Positive autoregulation within the cascade increases the gain-to-noise ratio

In a cascade with autoregulation by an intermediate component V the story is different (Fig. 2.4a and Eq. 2.21 in Fig. 2.4b). First, we reiterate that since we compare the simple cascade and the cascade with autoregulation on the basis of equal average production and degradation rates, the noise strengths $\langle \eta_x^2 \rangle$ and $\langle \eta_v^2 \rangle$ are the same for both cascades. However, as noted above the effective relaxation timescale of component v, $\tau_v = -J_{vv}^{-1}$ (Eq. 2.5), decreases with negative autoregulation and increases with positive autoregulation. This again leads to a reduction (increase) in both the gain (Fig. 2.4d, top left) and the noise (Fig. 2.4d, top right) of the network for negative (positive) autoregulation, as has been reported previously [34, 95]. However, unlike the case of autoregulation of the output X, the gain-to-noise ratio (Fig. 2.4d, bottom left) can change as a result.

Negative autoregulation (Fig. 2.4d, red) leads to a decrease in the response time compared to a simple cascade (gray), corresponding to an increase in $|J_{vv}|$. This leads to a decrease in the gain of the autoregulated component $g_{s\to v}^2(\omega) = J_{vs}^2/(\omega^2 + J_{vv}^2)$ at frequencies $\omega < |J_{vv}|$. Negative autoregulation therefore tends to suppress slowly varying signals relative to the simple cascade. Noise which is introduced upstream of or at the autoregulated component is filtered by the feedback-modified gain in exactly the same way as the signal, whereas noise introduced downstream of v is unaffected. Hence negative autoregulation reduces both the total gain of the network, which is the product of the individual reaction gains $g^2(\omega) = g_{s\to v}^2(\omega) g_{v\to x}^2(\omega)$, and the noise transmitted from v to x, $N_{v\to x}(\omega) = g_{v\to x}^2 N_v(\omega)$, relative to the simple cascade. However, noise in the production and degradation of x is unchanged relative to the simple cascade. Since the total noise (Eq. 2.23b in Fig. 2.4c) is the sum of independent noise contributions, $N(\omega) = N_x(\omega) + N_{v\to x}(\omega)$, the total noise decreases by a smaller factor than the gain, and the gain-to-noise ratio decreases compared with the simple cascade.

Conversely, positive autoregulation (Fig. 2.4d, black) increases the relaxation time of V, which increases $g_{s\to v}^2(\omega)$ at frequencies $\omega < |J_{vv}|$. We can therefore see that positive au-

toregulation amplifies slowly-varying signals. This leads to an increase in the network gain and the noise that is propagated from v to x. However, since the noise that is introduced at x is unchanged, positive autoregulation at V increases the gain-to-noise ratio compared to the simple cascade. Fig. 2.4d shows the comparison between a simple cascade and cascades with positive and negative autoregulation. Hornung and Barkai previously studied transmission of a *constant* signal with additive noise through a deterministic (noiseless) network [99], and found that positive autoregulation can increase the signal-to-noise ratio. Our results for time-varying signals with intrinsic network noise parallel their results.

Given a network with autoregulation, our constraint of equal production of each network component does not define a unique "equivalent" simple cascade. That is, different parameter combinations can be chosen for a simple cascade which satisfy the production constraint. The results in the preceding discussion correspond to one such parameter choice. Specifically, we choose the production rate of V in the simple cascade (Eq. 2.12) to be $k_v = \langle f(v) \rangle$, while taking the same value for μ_v in both networks. A consequence of this choice is that the relaxation time τ_v changes between the two cascades, as discussed above. One can equally well construct a simple cascade for which the diagonal entries of the Jacobian, $J_{\alpha\alpha}$, are equal to those of the autoregulated cascade, so as to hold constant the relaxation time of each component between the two cascades. This is achieved by setting the spontaneous degradation rate for V in Eq. 2.12 to be $\mu_v^{\text{new}} = \mu_v - \langle \partial / \partial v f(v) s \rangle$. By choosing this new rate, the average protein number $\langle v \rangle$ changes in the simple cascade, and as a result also the average production of x. To restore equal production of x we thus also require a rescaling of the kinetic production rate $k_x^{\text{new}} = k_x \mu_v^{\text{new}} / \mu_v$ in the simple cascade (Eq. 2.12). Thus, in this comparison, the diagonal entries of the Jacobian matrices of the autoregulated and simple cascade are the same, while the off-diagonal entry $J_{xv} = k_x$ differs between the two.

Compared to a cascade with positive autoregulation, this new kinetic production rate in the simple cascade is smaller ($k_x^{\text{new}} < k_x$). The reduction in J_{xv} leads to a uniform decrease in $g^2_{v \to x}(\omega)$ at all frequencies. As described above, this affects the signal and also the propagated noise $N_{v \to x}(\omega)$ equally, but not the intrinsic noise at x, $N_x(\omega)$. Thus, compared to a cascade with positive autoregulation, the gain-to-noise ratio is reduced at all frequencies in the simple cascade, as can be seen in Fig. 2.5 (red dashed). Interestingly, the decrease in the gain-to-noise ratio is most pronounced at high frequencies. This is because the propagated noise $N_{v \to x}(\omega)$ only has a significant contribution at frequencies $\omega < \mu_v^{\text{new}}$; at higher frequencies the total noise is dominated by $N_x(\omega)$, as discussed in the section: *The simple cascade*. Thus at these higher frequencies, the gain is reduced relative to the positively-autoregulated cascade, but the noise is not, and so the change in the gain-to-noise is largest. For networks with negative autoregulation, the converse applies: the gain-to-noise ratio is higher in the simple cascade at all frequencies, but by a larger factor for $\omega < \mu_v^{\text{new}}$. Hence, the effect of positive or negative autoregulation is qualitatively the same in both parameterizations.

More generally, even if we relax the production constraint on each component, and instead require only the total production in the two cascades to be the same (i.e., $\langle f_v^+ \rangle + \langle f_x^+ \rangle = \text{constant}$), we see the similar qualitative behavior for the gain-to-noise ratio (see

Eqs. A2.21-A2.23). Positively-autoregulated cascades have a larger gain-to-noise ratio than a simple cascade of the same length, while for a cascade with negative autoregulation the gain-to-noise ratio is smaller. For longer cascades drawing such general conclusions is more difficult. However, if the majority of parameters are kept the same between the simple and autoregulated cascades, as in the cases discussed in detail above, then we again find that positive autoregulation increases and negative autoregulation decreases the gain-to-noise ratio. Furthermore, given a specific simple cascade one can always add positive autoregulation to the network in such a way as to achieve a larger gain-to-noise ratio while maintaining the same total production cost.

We have here considered only autoregulation via the production of the intermediate v. However, for autoregulation via the degradation of V we observe similar results for the gain-to-noise ratio: if V suppresses its own degradation, the decrease in the effective turnover rate leads to a reduction of the noise strength $N_{v\to x}(\omega)$, increasing the gain-to-noise ratio; when V enhances its own degradation rate the transmitted noise is increased, reducing signaling fidelity.

2.3.3 Feedback

Feedback, both positive and negative, corresponds to the upper-triangular part in the Jacobian of the linearized system (see Fig. 2.1b). It is known that negative feedback allows for adaptation as, for example, in the *E. coli* chemotaxis pathway [87, 90, 100]. Feedback can also shift noise to higher frequencies [95]. We will again consider separately the two cases of feedback by the output X onto an upstream component and feedback by an intermediate component onto a component higher up the cascade.

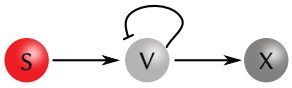
Feedback from X does not affect information transmission

For negative feedback from X to V (Fig. 2.6a and Eq. 2.27 in Fig. 2.6b), the power spectrum of the response $P_{xx}(\omega)$ (Fig. 2.6d, dark red solid) can have a resonance peak while none is present in the input signal (black solid). Surprisingly, this peak does not correspond to an increase in information transmission capabilities at the peak frequency (ω_{peak}), since no peak is present in the gain-to-noise ratio (Fig. 2.6d, red solid). For positive feedback, no peak is present in either $P_{xx}(\omega)$ or the gain-to-noise ratio.

For a system with negative feedback from X to V the gain and noise both show a peak, but these can occur at different frequencies. We consider first the frequency dependence of the gain. At low frequencies the negative feedback leads to destructive interference at v between the input signal $\tilde{S}(\omega)$ and the signal that is fed back, $\tilde{X}(\omega)$. On the other hand, at high frequencies these two signals are exactly out of phase, and hence the interference becomes constructive (since the feedback combines the two signals negatively). However, at frequencies $\omega \gg \mu_v, \mu_x$ the *amplitude* of the fed-back signal decreases, due to averaging over the lifetimes of V and X; hence, even though the two signals interfere constructively, the significance of this interference decreases. Together, these three effects lead to a maximum in the gain. This maximum occurs at

$$\omega_{\text{res}}^2 = -\frac{1}{2}\left[\mu_x^2 + \mu_v^2 + 2J_{vx}J_{xv}\right], \tag{2.25}$$

a)



b)

$$\frac{dv}{dt} = f(v)s - \mu_v v + \eta_v, \tag{2.21a}$$

$$\frac{dx}{dt} = \beta v - \mu_x x + \eta_x, \tag{2.21b}$$

$$f(v) = \frac{\nu\beta}{K+v}\begin{cases}\beta = K, & \text{neg. ar.}\\ \beta = v, & \text{pos. ar.,}\end{cases}$$

$$J_{vv} = -\mu_v + \langle s\rangle\left[\frac{\partial f(v)}{\partial v}\right]_{\text{s.s.}}. \tag{2.21c}$$

c)

$$g^2 = \frac{(J_{xv}J_{vs})^2}{(\omega^2+\mu_x^2)(\omega^2+J_{vv}^2)}, \tag{2.22}$$

$$N = \frac{J_{xv}^2\left\langle|\eta_v|^2\right\rangle + (\omega^2+J_{vv}^2)\left\langle|\eta_x|^2\right\rangle}{(\omega^2+\mu_x^2)(\omega^2+J_{vv}^2)}, \tag{2.23}$$

$$\frac{g^2}{N} = \frac{J_{xv}J_{vs}\mu_v}{2\langle s\rangle\left[\omega^2+J_{vv}^2+J_{xv}\mu_v\right]}. \tag{2.24}$$

d)

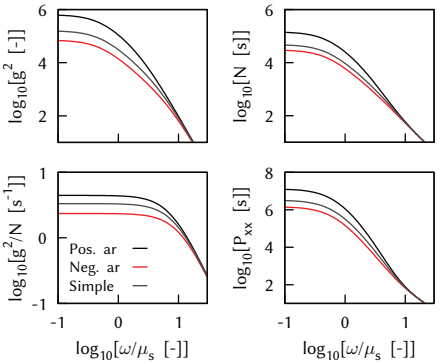


Figure 2.4: A two-step cascade with autoregulation (ar) of the intermediate component. **a)** Cartoon of the negative autoregulation motif, where the intermediate component V negatively regulates its own production. **b)** The Langevin equations describing the network. **c)** The characteristic equations for the gain $g^2(\omega)$, noise $N(\omega)$ and gain-to-noise ratio. **d)** The gain, noise, gain-to-noise ratio (g^2/N) and output power spectrum $P_{xx}(\omega)$ plotted as a function of frequency for three different cascades: simple (gray), positive autoregulation (black) and negative autoregulation (red). Negative autoregulation reduces the gain, noise and gain-to-noise ratio. For positive autoregulation the opposite holds. Positive autoregulation has a smaller knee frequency in the gain-to-noise ratio than negative autoregulation (see also the section: *Autoregulation*). Parameters: k_s=10, k_v=100, k_x=10, μ_v=5, μ_x=0.5, K=$\langle v\rangle$ and ν^a=200, ν^r=200.

which depends on the relaxation rates μ_x, μ_v and the coupling (feedback) loop between v and x, $J_{vx}J_{xv}$. This timescale corresponds to the imaginary part of the eigenvalues of the Jacobian.

The frequency of the peak in the noise depends on the relative strengths of the two noise sources, η_v and η_x. The two noise terms are propagated differently through the network, because η_x originates at the regulator of the feedback loop, while η_v originates at the regulated component. We consider two limiting cases. If the total noise $N(\omega)$ (Eq. 2.28b in Fig. 2.6c) is dominated by the transmitted noise, $N_{v\to x}(\omega)$, both the signal $\Sigma(\omega)$ and the dominant source of noise originate upstream of the feedback loop. Effectively, therefore, the feedback affects both the gain and noise of the network similarly. As a result the peak frequencies of both the noise and the gain are the same. On the other hand, when the total noise is dominated by $N_x(\omega)$, which is located downstream of the regulated component V,

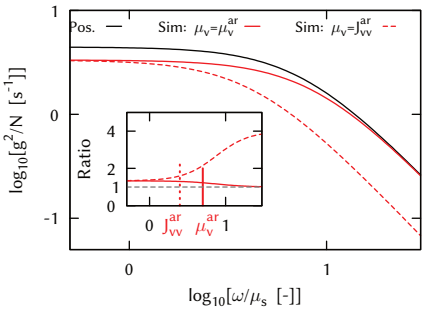


Figure 2.5: The gain-to-noise ratio for a cascade with positive autoregulation (ar, black solid) and two simple cascades (red solid, dashed). Solid: as in **2.4d**, the degradation rate of the simple cascade equals that of the cascade with positive autoregulation, $\mu_V=\mu_V^{ar}$. Hence $|J_{VV}|$ is smaller in the cascade with auto-activation, and the gain-to-noise ratio is larger at low frequencies. Dashed: In the simple cascade we take $\mu_V=J_{VV}^{ar}$, and instead increase the production rate J_{XV}. This decreases the gain-to-noise ratio of the simple cascade with respect to the autoregulated cascade over the full frequency spectrum. Inset: the ratio of the gain-to-noise ratio of the cascade with positive autoregulation to that of the simple cascade; solid: $\mu_V=\mu_V^{ar}$, dashed $\mu_V=J_{VV}^{ar}$. The dotted red vertical line indicates J_{VV}^{ar}, the vertical solid red line μ_V^{ar}, which shows the shift in frequency dependence. Parameters: k_S=10, k_V=100, k_X=10, μ_V=5, μ_X=0.5, K=$\langle V\rangle$ and ν=200.

the feedback loop affects the signal and noise differently. As a result, the noise that is fed back has a different frequency profile than the signal, such that the peaks in the gain and the noise occur at different frequencies (Fig. 2.6d, red dots).

One might therefore expect that when $N_x(\omega) \gg N_{v\to x}(\omega)$ a peak in the gain-to-noise ratio is possible. However, an inspection of the expressions for the gain, Eq. 2.28a, and the noise, Eq. 2.28b (both in Fig. 2.6c), shows that they have the same denominator, such that the gain-to-noise ratio is a monotonically decreasing function of frequency (Eq. 2.28c in Fig. 2.6c). The effect of the negative feedback is canceled. Ultimately, this is due to the fact that the noise in the output x goes back into the feedback loop, such that the peaks in the gain and the noise cannot be controlled separately; in the next section, we show how this can be done. Furthermore, we note that the gain-to-noise ratio is again identical to a simple three-component cascade, as we also saw in the case of autoregulation of X. We conclude that feedback from X onto the cascade also has no effect on information transmission through the network.

This network (Eq. 2.28 in Fig. 2.6b) also highlights the idea that the power spectrum of the output $P_{xx}(\omega)$ may not be indicative of the information that is transmitted at different frequencies. We see in Fig. 2.6d that due to the negative feedback $P_{xx}(\omega)$ can have a peak at non-zero frequencies, even if none is present in the input signal. However, this peak does not correspond to the frequency at which the signal is transmitted most reliably. Instead, we can see that the peak is simply due to resonant amplification of both the signal and the noise at the characteristic frequency of the negative feedback loop.

It has been suggested [30] that a system where a negative feedback loop acts *on* the response component can have a large peak in the gain, such that signals on specific timescales

can be selected for. If we take in Fig. 2.6a not x but v to be the output of the network, we obtain

$$\frac{g^2(\omega)}{N(\omega)} = \frac{J_{vs}^2\left(\omega^2+\mu_x^2\right)}{J_{vx}^2\left\langle|\eta_x|^2\right\rangle + \left(\omega^2+\mu_x^2\right)\left\langle|\eta_v|^2\right\rangle}. \tag{2.26}$$

We observe that the gain-to-noise ratio is a monotonically increasing function of frequency and does not show a peak at any specific frequencies. Furthermore we note that as $\omega \to \infty$ the gain-to-noise ratio becomes equal to the gain-to-noise ratio for the one-step simple cascade ($J_{vs}/2\langle s\rangle$), since for large ω the noise from the downstream component is averaged out. Thus this network motif has a higher gain-to-noise at all frequencies than the cascade with x as the output. However, the information transmitted at low frequencies is less than if X were not present. Following the information processing inequality, the amount of information about s which is encoded in the dynamics of v is always larger than the corresponding information in x. By feeding back x to v we thus do not add more information to the signal, but essentially add an extra source of noise to the pathway from S to V. The strength of this noise is highest at frequencies $\omega < \mu_x$, and hence the effect of the feedback is to obscure the signal at these frequencies. As a result this motif acts as a high-pass filter for information.

Negative feedback within a cascade can lead to a peak in the gain-to-noise ratio

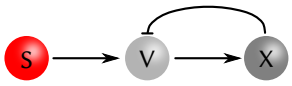
In the section: *Autoregulation* we saw that the gain-to-noise ratio is sensitive to the precise position of autoregulation in a cascade. In this section we therefore study a cascade where the feedback is not from X to V, but between two intermediate components W and V (see Fig. 2.7a and Eq. 2.33 in Fig. 2.7b). This also corresponds to taking the output of the previous feedback cascade (Fig. 2.6a) as the input to another downstream process.

Expressions for the gain, noise and gain-to-noise ratio are given in Fig. 2.7c. For positive feedback the gain, noise and gain-to-noise ratio are once again monotonically decreasing with increasing frequency. However, we find that for a network with strong negative feedback (Hill coefficient $n > 1$, see Eq. A2.35), the gain-to-noise ratio can have a maximum as a function of frequency at

$$\begin{aligned}\omega_{\text{peak}}^2 &= -\frac{1}{2}\left(\overbrace{J_{xw}^2\frac{\left\langle|\eta_w|^2\right\rangle}{\left\langle|\eta_x|^2\right\rangle}}^{\text{noise}} + \overbrace{\mu_v^2+\mu_w^2+2J_{vw}J_{wv}}^{\text{resonance}\left(\omega_{res}^2\right)}\right)\\ &= -\frac{1}{2}\left(J_{xw}^2\frac{\left\langle|\eta_w|^2\right\rangle}{\left\langle|\eta_x|^2\right\rangle} + \mu_v^2+\mu_w^2 - 2\frac{n\mu_w\mu_v\langle w\rangle^n}{K^n+\langle w\rangle^n}\right).\end{aligned} \tag{2.30}$$

This peak frequency depends on the characteristic resonance frequency of the feedback loop, ω_{res}, which is determined by the interactions between v and w: μ_v, μ_w, J_{vw} and

a)



b)

$$\frac{dv}{dt} = f(x)\, s - \mu_v v + \eta_v, \tag{2.27a}$$

$$\frac{dx}{dt} = \beta v - \mu_x x + \eta_x, \tag{2.27b}$$

$$f(x) = \frac{\nu C^n}{K^n + x^n} \begin{cases} C = x, & \text{pos. fb.} \\ C = K, & \text{neg. fb.} \end{cases}.$$

c)

$$g^2 = \frac{(J_{vs} J_{xv})^2}{(\omega^2 + \mu_v^2)(\omega^2 + \mu_x^2) + F(\omega)}, \tag{2.28a}$$

$$N = \frac{\beta^2 \left\langle |\eta_v|^2 \right\rangle + (\omega^2 + \mu_v^2) \left\langle |\eta_x|^2 \right\rangle}{(\omega^2 + \mu_v^2)(\omega^2 + \mu_x^2) + F(\omega)}, \tag{2.28b}$$

$$\frac{g^2}{N} = \frac{(J_{vs} J_{xv})^2}{\beta^2 \left\langle |\eta_v|^2 \right\rangle + (\omega^2 + \mu_v^2) \left\langle |\eta_x|^2 \right\rangle}, \tag{2.28c}$$

$$P_{xx}(\omega) = \frac{J_{vs}^2 J_{xv}^2 P_{ss}(\omega) + J_{xv}^2 \left\langle |\eta_v|^2 \right\rangle + (\omega^2 + \mu_v^2) \left\langle |\eta_x|^2 \right\rangle}{(\omega^2 + \mu_v^2)(\omega^2 + \mu_x^2) + F(\omega)}, \tag{2.28d}$$

with

$$F(\omega) = J_{vx} J_{xv} \left[J_{vx} J_{xv} + 2 \left(\omega^2 - \mu_x \mu_v \right) \right]. \tag{2.29}$$

d)

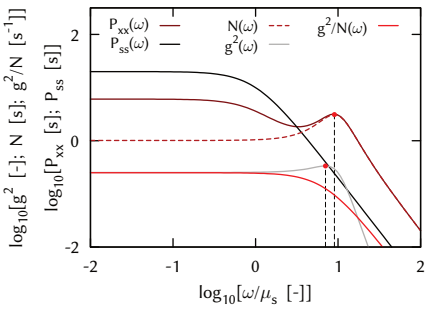


Figure 2.6: Feedback (fb) from the output signal X to an upstream component, discussed in section the section: *Feedback*. **a)** Cartoon of the negative feedback motif, where the output signal X negatively regulates V. **b)** The Langevin equations describing the network. **c)** The characteristic equations: gain $g^2(\omega)$, noise $N(\omega)$ and gain-to-noise ratio $g^2(\omega)/N(\omega)$. **d)** Power spectra of the output, $P_{xx}(\omega)$, input, $P_{ss}(\omega)$, gain $g^2(\omega)$ and noise $N(\omega)$. $P_{xx}(\omega)$, $g^2(\omega)$ and $N(\omega)$ all exhibit a peak due to the negative feedback, while the gain-to-noise ratio is monotonically decreasing. The red dots indicate the peaks of the gain and noise, which occur at different frequencies (see the section: *Feedback*). Parameters: k_s=10, K=0.2⟨x⟩, ν=1260, k_x=5, μ_v=5, μ_x=5 and n=3.

J_{wv}. It is additionally dependent on the relative strengths of the noise introduced into the network at w and at x.

We can understand the appearance of this peak as follows. For a network with negative feedback, $g^2(\omega)$ (Fig. 2.7c, bottom left) has a maximum as a function of frequency at ω_{res}, the characteristic resonance frequency of the feedback loop. Input signals at this frequency are

amplified by the constructive interference between the signal transmitted to v from s and the signal which is fed back from w to v. We note that the resonance frequency has the same form as Eq. 2.25, and depends only on the interactions between V and W. The behavior of the noise power spectrum (Fig. 2.7c, top right) is more complex. We consider two limiting cases in which different noise terms dominate. When the total noise is dominated by noise introduced at v or w, the noise is processed through the feedback loop together with the signal. As discussed in the previous section, $N(\omega)$ therefore shows a peak at a similar frequency to the gain (light gray solid). These two peaks cancel, and hence the gain-to-noise ratio (Fig. 2.7c, top left, light gray solid) is monotonically decreasing with frequency. On the other hand, when the total noise is dominated by $N_x(\omega)$ (top right, black solid) the noise in the network is not affected by the feedback loop. Hence no peak is found in the noise power spectrum. In this limit, the peak in the gain-to-noise ratio corresponds to the peak in the gain at ω_{res} (top left, black solid).

From these arguments we see that the peak in the gain-to-noise ratio becomes more pronounced as the relative contribution of $N_x(\omega)$ to the total noise increases. Additionally, increasing the strength of the negative feedback by reducing K or increasing n leads to a more pronounced peak. However, this increase in the *relative* peak height comes at the expense of a reduction in the value of the gain-to-noise ratio at all frequencies.

How does the gain-to-noise ratio of the network with feedback compare to the corresponding (four-component) simple cascade? We examine the ratio of the gain-to-noise for the network with feedback to the gain-to-noise of the simple cascade,

$$G_{\text{fb}}(\omega) = \left[\frac{g^2(\omega)}{N(\omega)}\right]_{\text{fb}} / \left[\frac{g^2\,(\omega)}{N(\omega)}\right]_{\text{simple}}, \tag{2.31}$$

and find that (Fig. 2.8a,b)

$$G_{\text{pos}}(\omega) > 1 \text{ if } \omega^2 < \mu_v\mu_w\left(1 - \frac{n}{2}\frac{K^n}{K^n + \langle w\rangle^n}\right), \tag{2.32a}$$

$$G_{\text{neg}}(\omega) > 1 \text{ if } \omega^2 > \mu_v\mu_w\left(1 + \frac{n}{2}\frac{\langle w\rangle^n}{K^n + \langle w\rangle^n}\right). \tag{2.32b}$$

Interestingly, for both types of feedback there is a range of frequencies over which the gain-to-noise ratio increases relative to the simple cascade. This contrasts to the results of the section: *Autoregulation*, where we found that autoregulation affected the gain-to-noise ratio in the same way at all frequencies.

This difference can again be understood in terms of the interference of the two signals arriving at V. As described above at low frequencies the signal propagated from s to v and the feedback signal from w to v are in phase, while at high frequencies the two signals are exactly out of phase. Hence for a positive feedback loop (Fig. 2.8a,b; black) the signals combine constructively at low frequencies, increasing the gain, but destructively at high frequencies, decreasing the gain. Recall that, since we are comparing networks with equal production, the noise strengths $\langle|\eta_v|^2\rangle$, $\langle|\eta_w|^2\rangle$ and $\langle|\eta_x|^2\rangle$ are equal in the regulated and simple cascades. In an analogous way to the autoregulation discussed in the section: *Au-*

toregulation, the presence of feedback between W and V affects both the signal and noise introduced upstream of x, but not noise introduced at x. Hence, at low frequencies positive feedback amplifies the signal and the noise introduced at the levels of V and W, but not noise introduced at X. Hence at low frequencies the gain-to-noise ratio increases relative to the simple cascade. At high frequencies, however, positive feedback reduces the gain and the noise upstream of x, but not the intrinsic noise $N_x(\omega)$; consequently, the gain-to-noise ratio is reduced compared to the simple cascade. Conversely, a network with negative feedback (Fig. 2.8a,b; red) reduces the gain at low frequencies, reducing the gain-to-noise ratio. However, at high frequencies, the feedback amplifies the signal but not $N_x(\omega)$, leading to an increase in the gain-to-noise ratio.

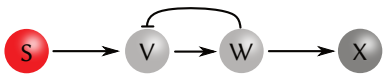
From these results we conclude that if a cell is only concerned with low frequency input signals, it is beneficial in terms of information transmission to add positive feedback within the signaling cascade. If the system wishes to respond specifically to high-frequency signals, negative feedback can be used to increase the fidelity of transmission for these signals. Additionally for a strong negative feedback ($n \gg 1$ or $K \ll \langle w \rangle$, see Eq. A2.43) the gain-to-noise can have a peak in the regime where signaling is more reliable than for a simple cascade, allowing the cell to focus on signals in a particular frequency band. We note that the negative feedback motifs considered here do not lead to perfect adaptation to constant input signals, which is characterized by $g^2(\omega = 0) = 0$ and is necessary for true band-pass behavior. Perfect adaptation requires that the feedback to be implemented via a buffer node or side branch [101]. An example of this network architecture is the *E. coli* chemotaxis pathway [92], for which the gain-to-noise ratio does indeed indicate a band-pass for information [65].

2.4 Discussion

In this chapter we have analyzed information transmission through a number of network motifs, namely cascades, autoregulation and feedback. One of the most important conclusions of our analysis is that to understand how reliably biochemical networks can transmit time-varying signals, we have to study the frequency-dependent gain-to-noise ratio [65]. In particular, the power spectrum of the output signal may not be a good measure for how biochemical networks transduce time-varying input signals. The power spectrum of the output signal depends on the power spectrum of the input signal, the frequency-dependent gain, and the frequency-dependent noise. Only the latter two quantities are intrinsic properties of the network, provided that the network detects the input via biochemical reactions that do not affect the statistics of the input signal [96]. Moreover, we have seen that the power spectrum of the output signal may differ qualitatively from that of the frequency-dependent gain-to-noise ratio. A striking example is provided by the network with negative feedback from the output component, which shows a peak in the output signal (see Fig. 2.6d): while one might be tempted to conclude that input signals at this frequency are transduced most reliably, our analysis shows that this peak in the output spectrum is simply the result of resonant amplification of both the input signal and the noise in the network.

Our analysis leads us to draw the following conclusions on the effect of autoregulation

a)



b)

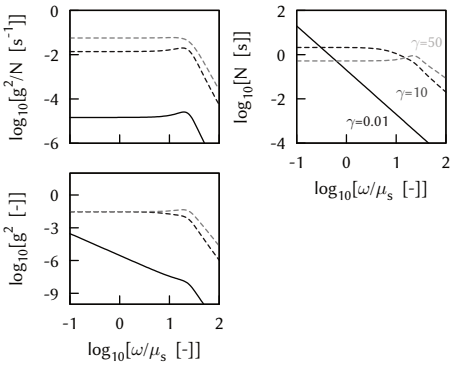
$$\frac{dv}{dt} = f(w)\, s - \mu_v v + \eta_v, \tag{2.33a}$$

$$\frac{dw}{dt} = \beta v - \mu_w w + \eta_w, \tag{2.33b}$$

$$\frac{dx}{dt} = \gamma w - \mu_x x + \eta_x, \tag{2.33c}$$

$$f(w) = \frac{\nu C^n}{K^n + w^n} \begin{cases} C = w, & \text{pos. fb.} \\ C = K, & \text{neg. fb.} \end{cases}$$

c)



d)

$$g^2 = \frac{(J_{vs} J_{wv} J_{xw})^2}{H_x F(\omega)}, \tag{2.34a}$$

$$N = \frac{(J_{xw} J_{wv})^2 \left\langle |\eta_v|^2 \right\rangle + J_{xw}^2 H_v \left\langle |\eta_w|^2 \right\rangle + F(\omega) \left\langle |\eta_x|^2 \right\rangle}{H_x F(\omega)}, \tag{2.34b}$$

$$\frac{g^2}{N} = \frac{(J_{vs} J_{wv} J_{xw})^2}{(J_{xw} J_{wv})^2 \left\langle |\eta_v|^2 \right\rangle + J_{xw}^2 H_v \left\langle |\eta_w|^2 \right\rangle + F(\omega) \left\langle |\eta_x|^2 \right\rangle}, \tag{2.34c}$$

where

$$F(\omega) = \omega^4 + \left(\mu_v^2 + \mu_w^2 + 2 J_{vw} J_{wv}\right) \omega^2 + (J_{vw} J_{wv} - \mu_w \mu_v)^2, \tag{2.35a}$$

with $H_i = \omega^2 + \mu_i^2$.

Figure 2.7: A three-step cascade with feedback (fb) from an intermediate component, discussed in the section: *Feedback*. **a)** Cartoon of a negative feedback motif, where S is the signal and X the response, and W negatively regulates V. **b)** The Langevin equations of this motif. **c)** The effect of changing the strength of the intrinsic noise in x, $N_x(\omega)$, on the spectra of the gain, noise, and gain-to-noise ratio of a cascade with negative feedback. $N_x(\omega)$ is varied by changing $\gamma(=J_{xw})$ and μ_x, in such a way that $\langle x \rangle$ remains constant. Lines show: gray solid, γ=50; gray dashed solid, γ=10; black solid, γ=0.01. Decreasing γ and μ_x in this way increases the relative contribution of $N_x(\omega)$ to the total noise. We see that as γ is reduced the gain and noise decrease at frequencies $\omega < \mu_x$, but the noise increases at lower frequencies. The gain-to-noise ratio decreases at all frequencies. However, the peak in the gain-to-noise ratio becomes more pronounced. Parameters: k_s=10, β=10, ν=330, K=0.5$\langle w \rangle$, n=5, μ_v=10, μ_w=10 (see App. 2.A.6) **d)** The characteristic equations: gain $g^2(\omega)$, noise $N(\omega)$ and gain-to-noise ratio $g^2(\omega)/N(\omega)$.

and feedback on the transmission of time-varying signals: 1) autoregulation of the output component does not affect the gain-to-noise ratio, and hence does not affect information transmission (Fig. 2.3c); 2) positive autoregulation of an intermediate component increases

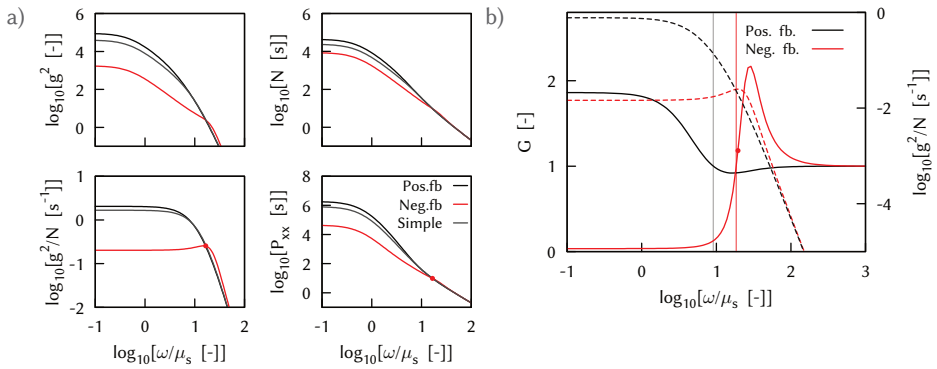


Figure 2.8: a) The spectra of the gain, noise, gain-to-noise ratio, and the output power, $P_{xx}(\omega)$ for a cascade with feedback (fb). For small ω, positive feedback (black) enhances the gain, noise, and gain-to-noise ratio, while negative feedback (red) decreases these. For higher frequencies, negative feedback increases the gain, enhancing the gain-to-noise ratio. With negative feedback a peak in the gain-to-noise ratio is present (denoted by the red dot), while none is present in the output power spectrum $P_{xx}(\omega)$. Parameters: k_s=10, μ_w=10, μ_v=10, μ_x=0.5, β=10, γ=10. For positive feedback: K=0.5$\langle w \rangle$, n=1 and ν=150. For negative feedback: K=0.5$\langle w \rangle$, n=4 and ν=1700. **b)** Solid lines show $G_{fb}(\omega)$ (left axis), the gain-to-noise ratio for networks with positive (black) or negative (red) feedback divided by that of the corresponding simple cascade. Relative to the simple cascade, positive feedback increases the gain-to-noise ratio at low frequencies, while negative feedback increases the gain-to-noise ratio at high frequencies. Vertical lines indicate the frequencies at which $G_{fb}(\omega)$=1 (Eq. 2.32). Dashed lines show the gain-to-noise ratios for the positive (black) and negative (red) feedback motifs (right axis). Parameters: k_s=10, μ_w=10, μ_v=10, μ_x=1, β=10, γ=1 and K=0.5$\langle w \rangle$. For positive feedback: n=1 and ν=150. For negative feedback n=5 and ν=3300.

the gain-to-noise ratio over all frequencies, while negative autoregulation tends to decrease it over all frequencies (Fig. 2.4d); 3) negative feedback from the output component onto an upstream component may lead to a peak in the power spectrum of the output, and those of the gain and the noise; yet, even though the peaks of gain and the noise can be at different frequencies, negative feedback from the output component onto an upstream component can *not* lead to a peak in the spectrum of the gain-to-noise ratio (Fig. 2.6d); 4) positive feedback between upstream components enhances the gain-to-noise ratio at low frequencies, while negative feedback increases the gain-to-noise ratio at high frequencies (Fig. 2.8b). Further, we note that it is possible to achieve a peak in the gain-to-noise ratio via negative feedback between components that are upstream of the output component (Fig. 2.7c); however, this comes at the expense of a reduction in the gain-to-noise ratio for all frequencies. We also note here that stronger band-pass filtering of information can be obtained with networks employing integral feedback in a side branch [65], as found in the networks of osmo adaptation [86] or bacterial chemotaxis [92]. Alternatively, band-pass filters for information transmission can be obtained via feed*forward* loops, which we will discuss in a forthcoming publication.

Taken together these results reveal the following design principles for the use of feedback

and autoregulation in signal transduction cascades (see the schematic drawing Fig. 2.1a). Firstly, feedback and autoregulation can improve information transmission, but only if they occur upstream of the dominant source of noise in the cascade. Feedback or autoregulation downstream of the dominant noise source affects the gain and the noise similarly. Secondly, if signals over the full frequency range have to be transmitted reliably, positive autoregulation is advantageous, while if the cell is concerned only with low- or high-frequency signals, then positive or negative feedback can be employed.

The approach employed here has a number of limitations. Firstly, we have used the linear-noise approximation, and the power spectra calculated using this approximation may deviate from those of the full non-linear system. We argue that this effect does not significantly affect our results, since we find excellent agreement between the power spectra calculated analytically using the linear-noise approximation and those obtained from stochastic simulations of the full system (see App. 2.A.1). Secondly, we stress that the expressions for the information transmission rate, Eq. 2.7 and Eq. 2.11, are exact only for linear Gaussian systems; yet, the information rate calculated in this approximation provides a lower bound on the information transmission rate of the full system [64]. In [65] , we showed how the information transmission rate R can depend on the variance of the input signal. Here, we do not provide such an analysis, because R indeed depends on the statistics of the input signal, while we focus here on the processing network, which is characterized by the gain-to-noise ratio.

Another limitation of our analysis is that to reduce the complexity of the problem, we have assumed that the networks obey the spectral-addition rule [96], meaning that reactants are not consumed during reaction events. However, irreversible modifications of a substrate molecule are common in biochemical networks, and reactions of this type can significantly change the correlations between different network components. For instance, in a cascade of the type $X_0 \rightarrow X_1 \rightarrow \ldots X_{n-1} \rightarrow X_n$, where in each reaction step the reactant is consumed, correlations of the form $\langle \eta_i \eta_{i+1} \rangle = -k \langle X_i \rangle$ appear between different noise terms. As a result, for this cascade the covariance between different components $\langle x_i x_{j \neq i} \rangle = 0$ [96, 102], and hence the mutual information between *instantaneous* levels of components X_i and $X_{j \neq i}$ is zero [65]. This may suggest that these cascades cannot effectively transmit information. Yet, the analysis of [65] indicates that this motif can, in fact, reliably transmit time-varying signals. It would therefore be of interest to study the effect of cross-correlations in the noise on the information transmission in the motifs studied here. We leave this for future work.

Lastly, how could our predictions be tested experimentally? It is increasingly being recognized that stimulating biochemical networks with time-varying signals provides a wealth of information on the dynamics of these networks [35, 86, 87, 103, 104, 105]. These experiments can also be used to study the reliability by which biochemical networks can transmit time-varying signals. By measuring not only the power spectra of the in- and output signals, $P_{ss}(\omega)$ and $P_{xx}(\omega)$, but also their cross-power spectrum $P_{sx}(\omega)$, one can obtain the frequency-dependent gain $g^2(\omega) \equiv |P_{sx}(\omega)|^2 / P_{ss}(\omega)^2$ and the frequency-dependent noise $N(\omega)$ (see Eq. 2.9), and hence the gain-to-noise ratio. Stimulating synthetic gene circuits or existing signal transduction pathways and gene regulation networks of known

architecture with time-varying signals, for example using microfluidic devices, would make it possible to test our predictions on the effect of feedback and autoregulation on information transmission.

2.5 Acknowledgements

I thank Martin Frimmer and Nils Becker for a careful reading of this chapter.

2.A Supplementary Information

All cascades have the following linear birth-death process for the signal

$$\frac{ds}{dt} = k_s - \mu_s s + \Gamma(t). \tag{A2.1}$$

2.A.1 Gillespie Simulations

The linearization used in the derivation can change the characteristics of the frequency response. A linearized system does not change the frequency of the transmitted signal. However, this may not be the case for a non-linear system. To study this, we performed Gillespie simulations of the full system. The positive and negative regulation in our networks arises from Hill-like interactions between components. In the Gillespie simulation we calculated the propensities for every reaction with identical expressions. For example, in the network with negative feedback from W to V, we model reactions like Eq. 2.33a in Fig. 2.7b as

$$\mathsf{S} \xrightarrow{r} \mathsf{S} + \mathsf{V}, \tag{A2.2}$$

where r is

$$r = \frac{\nu K^n s}{K^n + w^n}. \tag{A2.3}$$

In these equations the actual copy number w is used, and not $\langle w \rangle$, as in the linearized expressions (Eq. 2.34 in Fig. 2.7c).

The power spectra are calculated using 2^{11} (2048) exponentially distributed frequencies from $\omega = 10^{-3}\ \mathrm{s}^{-1}$ to $\omega = 10^{3}\ \mathrm{s}^{-1}$ and averaged over 2^4 neighboring frequencies to obtain a single data point. In total we have 2^7 datapoints. The length of the simulation is 10^6 seconds, or a maximum of 10^9 events. For every run 50 blocks are averaged.

The positive feedback loops considered here display bistability. For the positive feedback loops a constant low level production is added to drive the system to the stable state with high copy numbers, instead of the stable state where the copy number equals zero. The positively autoregulated component thus is described by

$$\frac{dv}{dt} = -\mu_v v + \eta_v + \begin{cases} \dfrac{\nu v s}{K + v} \text{ if } v \neq 0 \\ \dfrac{1}{1000} \text{ if } v = 0 \end{cases}. \tag{A2.4}$$

Linearizing this we find that the fluctuations follow

$$\frac{d\tilde{v}}{dt} = -\mu_v \tilde{v} + \eta_v - \frac{\nu K \langle s \rangle}{K + \langle v \rangle}\tilde{v} + \frac{\nu K}{K + \langle v \rangle}\tilde{s}, \tag{A2.5}$$

which is equivalent to the linearization of Eq. 2.18 in Fig. 2.4a. The addition of the basal expression therefore drives the system to a specific steady state, but does not change the dynamical behavior around this steady state. For positive feedback within the cascade, the

motif is described by

$$\frac{dw}{dt} = a + k_w v - \mu_w w + \eta_w. \tag{A2.6}$$

Taking different values for $a = 0.1, 1, 10$ does not lead to qualitatively different answers (see Fig. 2.11). Again, the basal production changes the steady state, but not the dynamical behavior around the steady state. All results are shown in Figs. 2.9-2.11.

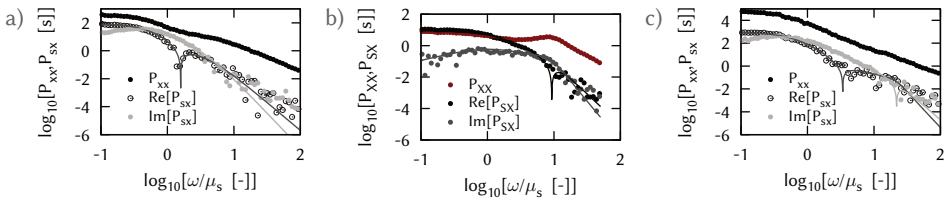


Figure 2.9: The results (symbols) of the Gillespie simulations together with the results of the linear noise approximation (lines) as employed in the main text. **a)** For the linear cascade (Eq. 2.12), Kinetic rates as in figure Fig. 2.2. **b)** For a network with negative feedback from X to V (Eq. 2.27 in Fig. 2.6b), Kinetic rates as in Fig. 2.6d. **c)** For a network with negative feedback from W to V (Eq. 2.33 in Fig. 2.7b), Kinetic rates as in Fig. 2.7d with negative feedback.

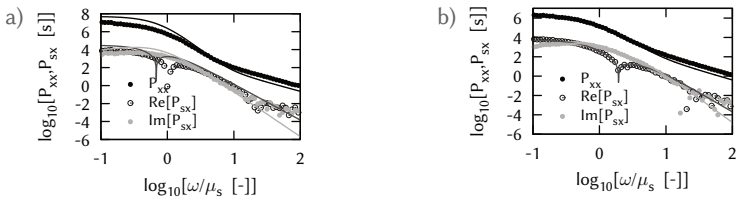


Figure 2.10: The results (symbols) of the Gillespie simulations together with the results of the linear noise approximation (lines) as employed in the main text. **a)** For the network with positive autoregulation of V (Eq. 2.21 in Fig. 2.4b) (kinetic rates as in Fig. 2.4c with positive autoregulation),To drive the system to the non-zero steady state, basal production of V is present (Eq. A2.4). The steady state of the full non-linear Gillespie simulation is slightly different from the steady state derived from the mathematical expressions for s,v and x. This causes the slight difference between the results of the linearization and the simulations. **b)** For a network with negative autoregulation on V (Eq. 2.21 in Fig. 2.4b), kinetic rates as in Fig. 2.4d with negative autoregulation.

2.A.2 Linear cascades

The one step linear cascade is described by

$$\frac{dx}{dt} = k_x s - m_x x + \eta_x\,(t) \tag{A2.7}$$

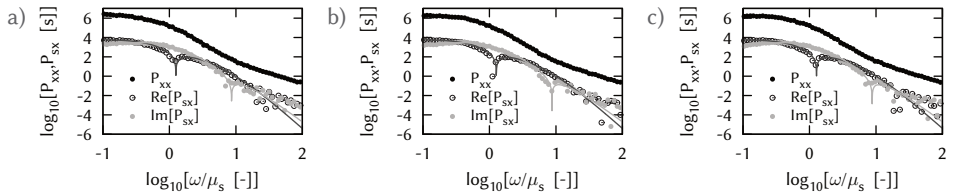


Figure 2.11: The results (symbols) of the Gillespie simulations for a network with positive feedback from W to V (Eq. 2.33 in Fig. 2.7b), together with the results of the linear noise approximation (lines) as employed in the main text. Kinetic rates as in Fig. 2.7d with positive feedback. To drive the system to the non-zero steady state, basal production of w is present (Eq. A2.6), **a)** with a=0.1, **b)** with a=1, **c)** with a=10.

with the following characteristic equations

$$g^2(\omega) = \frac{k_x^2}{\omega^2 + m_x^2}, \tag{A2.8a}$$

$$N(\omega) = \frac{\left\langle |\eta_x|^2 \right\rangle}{\omega^2 + m_x^2}, \tag{A2.8b}$$

$$\frac{g^2}{N} = \frac{k_x^2}{\left\langle |\eta_x|^2 \right\rangle}, \tag{A2.8c}$$

with $\langle |\eta_x|^2 \rangle = k_x \langle s \rangle + m_x \langle x \rangle = 2k_x \langle s \rangle$.

The two-step linear cascade is described by

$$\frac{dv}{dt} = k_v s - m_v v + \eta_v(t), \tag{A2.9a}$$

$$\frac{dx}{dt} = k_x v - m_x x + \eta_x(t), \tag{A2.9b}$$

with the following characteristic equations

$$g^2(\omega) = \frac{k_v^2 k_x^2}{(\omega^2 + m_x^2)(\omega^2 + m_v^2)}, \tag{A2.10a}$$

$$N(\omega) = \frac{k_x^2 \left\langle |\eta_v|^2 \right\rangle + (\omega^2 + m_v^2) \left\langle |\eta_x|^2 \right\rangle}{(\omega^2 + m_x^2)(\omega^2 + m_v^2)}, \tag{A2.10b}$$

$$\frac{g^2}{N} = \frac{k_v^2 k_x^2}{k_x^2 \left\langle |\eta_v|^2 \right\rangle + (\omega^2 + m_v^2) \left\langle |\eta_x|^2 \right\rangle}. \tag{A2.10c}$$

The linear cascades are used as a reference. For the kinetic rates of the linear cascades we use roman symbol (k and m). For the kinetic rates of the cascades with feedback regulation we use Greek symbols.

2.A.3 Autoregulation

Autoregulation by X

The process and characteristic equations (gain, noise, *GNR*) are listed in Fig. 2.3b,c. For equal average production, as the simple two-step cascade, (production rate k_x), we chose

$$\langle f(x)\, s\rangle \equiv f(\langle x\rangle)\,\langle s\rangle = k_x\,\langle s\rangle\,, \tag{A2.11}$$

where the definition expresses the fact that we assume that the average rates can be expressed by the rates at the deterministic steady state, thus ignoring fluctuations. Thus

$$k_x = J_{xs} = \frac{\partial f(\langle s\rangle, \langle x\rangle)}{\partial \langle s\rangle} = \frac{\nu\kappa}{(K+\langle x\rangle)}, \tag{A2.12}$$

and $\langle|\eta_x|^2\rangle = 2k_x\langle s\rangle$. Expressed in terms of the kinetic rates of the one-step linear cascade, the autoregulated cascade has the following form

$$\frac{g^2}{N} = \frac{k_x^2}{2k_x\langle s\rangle} = \frac{k_x}{2\langle s\rangle} \tag{A2.13}$$

which is identical to Eq. A2.8c. The power spectrum of x for the autoregulated cascade is

$$P_{xx}(\omega) = \frac{J_{xs}^2\left\langle|\Gamma|^2\right\rangle + \left(\omega^2+\mu_s^2\right)\left\langle|\eta_x|^2\right\rangle}{\left(\omega^2+\mu_s^2\right)\left(\omega^2+\left(\mu_x - \frac{\partial f(\langle s\rangle,\langle x\rangle)}{\partial\langle x\rangle}\right)^2\right)}. \tag{A2.14}$$

Following a rescaling of the kinetic degradation rate μ_x, such that $\mu_x^{\text{new}} = \mu_x - J_{xx}$, we observe that the power spectrum of the simple cascade and the autoregulated cascade agree. This is because the noise term η_x depends on the mean rate of the production and degradation events. In steady state the average number of production events equals the average number of degradation events. Since by the rescaling the production is not changed, the noise η_x is constant. The change in $\mu_x \to \mu_x^{\text{new}}$ will lead to a new steady state value $\langle x\rangle$, but not to a different number of degradation events.

2.A.4 Autoregulation by V

For autoregulation of one of the intermediate components the network and characteristic equations (gain, noise, *GNR*) are listed in Fig. 2.4b,c. We equalize the production of v between the autoregulated cascade and the two-step simple cascade (with rates k_v and k_x) to obtain

$$k_v = J_{vs} = \frac{\nu\kappa}{(K+\langle v\rangle)}, \tag{A2.15}$$

and the *GNR* for the autoregulated cascade expressed in terms of the kinetic rates of the simple cascade (thus using k_v, k_x and m_v, m_x where applicable) is

$$\begin{aligned}\frac{g^2}{N} &= \frac{\left(J_{vs}J_{xv}\right)^2}{\beta^2\left\langle|\eta_v|^2\right\rangle + (\omega^2 + J_{vv}^2)\left\langle|\eta_x|^2\right\rangle} \\ &= \frac{k_v^2\beta^2}{\beta^2\left\langle|\eta_v|^2\right\rangle + \left(\omega^2 + \left(\mu_v - \frac{\partial f(\langle s\rangle,\langle v\rangle)}{\partial\langle v\rangle}\right)^2\right)\left\langle|\eta_x|^2\right\rangle}.\end{aligned} \tag{A2.16}$$

We keep all kinetic rates equal in the autoregulated and simple cascade that do not influence the constraint condition (Eq. A2.15) (thus $\mu_v = m_v$ and $\beta = k_x$). We obtain

$$\frac{g^2}{N} = \frac{k_v^2 k_x^2}{k_x^2\left\langle|\eta_v|^2\right\rangle + \left(\omega^2 + \left(m_v - \frac{\partial f(\langle s\rangle,\langle v\rangle)}{\partial\langle v\rangle}\right)^2\right)\left\langle|\eta_x|^2\right\rangle}. \tag{A2.17}$$

We note that for positive autoregulation $|J_{vv}| < \mu_v$ while for negative autoregulation $|J_{vv}| > \mu_v$. Thus the *GNR* is larger for the positively autoregulated than the two-step cascade, especially for $\omega < J_{vv}$. For the negatively autoregulated cascade the opposite holds.

The constraint does not lead to a unique relation between an autoregulated and a non-autoregulated cascade. An alternative choice would be a simple two-step cascade for which the degradation rate μ_v is equivalent to the "effective" degradation rate in the autoregulated cascade. Thus $m_v = J_{vv}^{\text{ar}}$. The production of x is then

$$\overbrace{\beta\frac{k_v}{\mu_v}}^{\text{autoregulated}} = \overbrace{k_x\frac{k_v}{m_v}}^{\text{two-step cascade}} \tag{A2.18}$$

Equalizing this leads to

$$\beta\frac{k_v}{\mu_v} = k_x\frac{k_v}{m_v} \Rightarrow k_x = \beta\frac{m_v}{\mu_v}. \tag{A2.19}$$

And we obtain

$$\begin{aligned}\frac{g^2}{N} &= \frac{\left(\beta\frac{\mu_v^{\text{casc}}}{\mu_v}k_v\right)^2}{\left(\beta\frac{m_v}{\mu_v}\right)^2\left\langle|\eta_v|^2\right\rangle + \left(\omega^2 + m_v^2\left\langle|\eta_x|^2\right\rangle\right)} \\ &= \frac{(\beta k_v)^2}{\beta^2\left\langle|\eta_v|^2\right\rangle + \left(\frac{\mu_v}{m_v}\right)^2\left(\omega^2 + m_v^2\left\langle|\eta_x|^2\right\rangle\right)}\end{aligned} \tag{A2.20}$$

for the *GNR* of the two-step simple cascade. Since for positive feedback $m_v < \mu_v$, the *GNR*

of the positively autoregulated cascade is larger than that of the simple cascade, especially if $\omega \gg m_v$ or $\omega^2 \gg \beta^2 \langle |\eta_v|^2 \rangle$.

If we allow for even more differences between the kinetic rates, but require equal production, we obtain the following equations (we still assume the signal to be identical in both cases)

$$\mu_v = C m_v \text{ and } \beta = C k_x, \tag{A2.21}$$

where C is an arbitrary constant. We note that the mean level of v differs between the autoregulated and the simple cascade

$$\langle v \rangle^{\text{ar}} = \frac{1}{C} \langle v \rangle^{\text{simple}}. \tag{A2.22}$$

As a result we derive for the gain-to-noise ratio for the regulated cascade (using Eqs. A2.16, A2.21, and A2.22)

$$\begin{aligned} \frac{g^2}{N} =& \frac{(C k_x k_v)^2}{(C k_x)^2 \left\langle |\eta_v|^2 \right\rangle + \left(\omega^2 + \left(C \mu_v - \frac{\partial f(\langle s \rangle, \langle v \rangle)}{\partial \langle v \rangle} \right)^2 \right) \left\langle |\eta_x|^2 \right\rangle} \\ =& \frac{(k_x k_v)^2}{k_x^2 \left\langle |\eta_v|^2 \right\rangle + \left(\frac{\omega^2}{C^2} + \left(\mu_v - \frac{1}{C} \frac{\partial f(\langle s \rangle, \langle v \rangle)}{\partial \langle v \rangle} \right)^2 \right) \left\langle |\eta_x|^2 \right\rangle} \end{aligned} \tag{A2.23}$$

For small ω the conclusions on positive and negative feedback are still valid, but for $\omega \to \infty$ the ratio of the *GNR* for positive feedback and a two-step cascade is a function of C. Similar arguments can be made about comparing negative and positive feedback for $\omega \to \infty$, where again the ratio of the gain-to-noise ratio's depends on C.

2.A.5 Feedback

Feedback from X to V

For feedback of the response onto the intermediate components the network and characteristic equations (gain, noise, *GNR*) are listed in Fig. 2.6b,c. The power spectrum is

$$P_{xx}(\omega) = \frac{(J_{vs}\beta)^2 P_{ss}(\omega) + \beta^2 \left\langle |\eta_v|^2 \right\rangle + \left(\omega^2 + \mu_v^2 \right) \left\langle |\eta_x|^2 \right\rangle}{\left(\omega^2 + \mu_v^2 \right) \left(\omega^2 + \mu_x^2 \right) + J_{vx}\beta \left[J_{vx}\beta + 2 \left(\omega^2 - \mu_v \mu_x \right) \right]}, \tag{A2.24}$$

where

$$J_{vx} = \frac{\partial \langle f(x) s \rangle}{\partial \langle x \rangle} = \frac{\partial \frac{\nu \langle \kappa \rangle^n \langle s \rangle}{K^n + \langle x \rangle^n}}{\partial \langle x \rangle}. \tag{A2.25}$$

We note that the *GNR* is independent of J_{vx}. The peak in P_{xx} only exist if $J_{vx} < 0$, since the numerator is independent of J_{vx}, while the denominator is monotonic increasing for

$J_{vx} > 0$. Only for negative feedback a peak can exist in the power spectrum, gain and noise (since the same argument applies to gain and noise).

The frequency of the maximum of the gain can easily be obtained, since it coincides with the minimum of the denominator D

$$D = \left(\omega^2 + \mu_v^2\right)\left(\omega^2 + \mu_x^2\right) + J_{vx}\beta\left[J_{vx}\beta + 2\left(\omega^2 - \mu_x\mu_v\right)\right]. \tag{A2.26}$$

This frequency, where the gain has a maximum, is

$$\omega_{\text{res}}^2 = -\frac{1}{2}\left[\mu_v^2 + \mu_x^2 + 2J_{xv}\beta\right], \tag{A2.27}$$

such that we require $\mu_v^2 + \mu_x^2 + 2J_{xv}\beta < 0$. As a check we note that $D > 0$ for ω_{res} so divergence is not possible. The maximum frequency for the noise is not the minimum of D, due to the ω-dependence in the numerator. If $\beta^2\langle|\eta_v|^2\rangle \gg \langle|\eta_x|^2\rangle$, the ω-dependence in the noise is less strong, and the frequency of the peak of the noise shifts to the frequency of the peak in the gain. Although a peak in P_{xx} can be derived analytically ($dP_{xx}/d\omega$ is 4^{th} order in ω^2), it is not insightful. We note that P_{xx} is the sum of the noise ($N(\omega)$) and the signal ($\Sigma(\omega)$), such that if one of these two dominates in P_{xx} the peak is likely to coalesce with the peak of the dominating term. We also note that the signal Σ depends on μ_s, so the peak in P_{xx} is not likely to coincide exactly with the peak in the gain, since the gain is independent of μ_s.
Compared with a two-step cascade (rates k_v, k_x) , requiring equal production, we note that

$$k_v = \frac{\nu\kappa^n}{K^n + \langle x\rangle^n} \tag{A2.28}$$

and the two-step cascade has an identical *GNR* as the cascade with regulation.

Feedback from W to V

For feedback of an intermediate onto another intermediate component the network and characteristic equations (gain, noise, *GNR*) are listed in Fig. 2.7b,d. The *GNR* (Eq. 2.34c) is described by $C/a\,(\omega)$, i.e. a constant divided by a function of ω. For this to have an extremum, the denominator should have an extremum. We differentiate and obtain

$$\omega_{\text{peak}}^2 = -\frac{1}{2}\left(\gamma^2\frac{\left\langle|\eta_w|^2\right\rangle}{\left\langle|\eta_x|^2\right\rangle} + \mu_v^2 + \mu_w^2 + 2J_{vw}\beta\right). \tag{A2.29}$$

Since this expression is negative, to have $\omega^2 > 0$ we require negative feedback. Explicitly writing $J_{wv}J_{vw}$, we have for the requirement that a peak exists

$$2J_{wv}J_{vw} = 2\frac{\nu n \langle w\rangle^n K^n \langle s\rangle}{\langle w\rangle \left(K^n + \langle w\rangle^n\right)^2}\beta > \left(\gamma^2 \frac{\left\langle |\eta_w|^2 \right\rangle}{\left\langle |\eta_x|^2 \right\rangle} + \mu_v^2 + \mu_w^2\right) \tag{A2.30}$$

$$\langle w\rangle = \frac{\beta \langle v\rangle}{\mu_w} = \frac{\beta}{\mu_v\mu_w}\frac{\nu K^n \langle s\rangle}{K^n + \langle w\rangle^n}, \tag{A2.31}$$

which gives n solutions for $\langle w\rangle$ (of which only one is real and positive). If we constrain the production rate of v and w to be constant — and we assume $\langle v\rangle = \left(k_v\langle s\rangle\right)/m_v$ — then we obtain

$$\frac{\nu K^n}{K^n + \langle w\rangle^n} = k_v,$$

and the following expression for Eq. A2.31

$$\langle w\rangle = \frac{\beta \langle v\rangle}{\mu_w} = \frac{k_v\beta \langle s\rangle}{\mu_w\mu_v}. \tag{A2.32}$$

We rewrite the coupling strength

$$J_{vw}^{\text{pos}} = -J_{vw}^{\text{neg}} = \frac{\nu n \langle w\rangle^{n-1} K^n \langle s\rangle}{\left(K^n + \langle w\rangle^n\right)^2}, \tag{A2.33}$$

and substitute $\langle w\rangle$

$$J_{vw}^{\text{pos}} = \frac{n\langle s\rangle k_v}{\langle w\rangle}\frac{K^n}{K^n + \langle w\rangle^n} = \frac{n\langle s\rangle k_v}{\langle w\rangle}\frac{(\mu_w\mu_v K)^n}{(\mu_w\mu_v K)^n + (k_v k_w \langle s\rangle)^n}, \tag{A2.34a}$$

$$J_{vw}^{\text{neg}} = -\frac{n\langle s\rangle k_v}{\langle w\rangle}\frac{\langle w\rangle^n}{K^n + \langle w\rangle^n} = -\frac{n\langle s\rangle k_v}{\langle w\rangle}\frac{(k_v k_w \langle s\rangle)^n}{(\mu_w\mu_v K)^n + (k_v k_w \langle s\rangle)^n}. \tag{A2.34b}$$

For $K \ll \langle w\rangle$ positive regulation is maximized and J_{vw}^{pos} is maximal, while negative regulation is greatly suppressed and $\left|J_{vw}^{\text{neg}}\right|$ is minimal. The limit $n \to \infty$ is more complicated. If $K < \langle w\rangle$, $J_{vw}^{\text{neg}} \to -\infty$, while $J_{vw}^{\text{pos}} \to \infty$ for $\langle w\rangle < K$. In the opposite scenario's the limits tend to zero. This is only valid if while changing n, $\langle w\rangle$ remains constant, which is true due the constraint.

With Eq. A2.34b we can study ω_{peak} in more detail and we obtain

$$\begin{aligned} 2\frac{\nu n \langle w\rangle^n K^n \langle s\rangle}{\langle w\rangle \left(K^n + \langle w\rangle^n\right)^2}\beta &> \left(\gamma\mu_w + \mu_v^2 + \mu_w^2\right) \\ \frac{2k_v\beta n \langle s\rangle}{\langle w\rangle}\frac{\langle w\rangle^n}{K^n + \langle w\rangle^n} &> \left(\gamma\mu_w + \mu_v^2 + \mu_w^2\right) \\ 2\frac{\langle w\rangle^n}{K^n + \langle w\rangle^n}\mu_v\mu_w n &> \left(\gamma\mu_w + \mu_v^2 + \mu_w^2\right), \end{aligned} \tag{A2.35}$$

which, interestingly, only has a solution for $n > 1$.
The power spectrum of x is

$$P_{xx}(\omega) = \frac{J_{vs}^2\beta^2\gamma^2 P_{ss} + \gamma^2\beta^2\left\langle|\eta_v|^2\right\rangle + \gamma^2\left(\omega^2 + \mu_v^2\right)\left\langle|\eta_w|^2\right\rangle + F(\omega)\left\langle|\eta_x|^2\right\rangle}{\left(\omega^2 + \mu_x^2\right)F(\omega)}, \tag{A2.36}$$

which depends on μ_s through P_{ss} and therefor has a peak for a different ω than the *GNR*. The *GNR* for a simple three-step cascade is

$$\frac{g^2(\omega)}{N(\omega)} = \frac{(k_v k_w k_x)^2}{k_x^2\left[k_w^2\left\langle|\eta_v|^2\right\rangle + \left(\omega^2 + m_v^2\right)\left\langle|\eta_w|^2\right\rangle\right] + \left(\omega^2 + m_v^2\right)\left(\omega^2 + m_w^2\right)\left\langle|\eta_x|^2\right\rangle} \tag{A2.37}$$

$$= \frac{(k_v k_w k_x)^2}{D}, \tag{A2.38}$$

where we chose ν such that

$$k_v = \frac{\nu\kappa^n}{K^n + \langle w\rangle^n}, \tag{A2.39}$$

to obtain equal production. We obtain for the ratio of the *GNR* of the feedback cascade and a simple cascade

$$G = \frac{\left[g^2/N\right]_{\text{fb}}}{\left[g^2/N\right]_{\text{simple}}} = \frac{D}{D + J_{vw}\beta\left[J_{vw}\beta + 2\left(\omega^2 - \mu_v\mu_w\right)\right]\left\langle|\eta_x|^2\right\rangle}. \tag{A2.40}$$

So that the feedback is larger if $J_{vw}\beta\left[J_{vw}\beta+2\left(\omega^2-\mu_v\mu_w\right)\right]\left\langle|\eta_x|^2\right\rangle<0$.
The result of this inequality is

$$G_{\text{pos}}(\omega)>1 \text{ if } \omega^2<\mu_v\mu_w\left(1-\frac{n}{2}\frac{K^n}{K^n+\langle w\rangle^n}\right), \tag{A2.41}$$

$$G_{\text{neg}}(\omega)>1 \text{ if } \omega^2>\mu_v\mu_w\left(1+\frac{n}{2}\frac{\langle w\rangle^n}{K^n+\langle w\rangle^n}\right). \tag{A2.42}$$

which is Eq. 2.32 from the article. The peak for the negative feedback occurs at ω_{peak} (Eq. A2.29). The negative feedback cascade is larger than the three-step simple cascade if $\omega>\omega_{\text{switch}}$ (Eq. 2.32). Thus if $\omega_{\text{peak}}>\omega_{\text{switch}}$ the *GNR* for the negative feedback at the peak is larger than the three-step cascade

$$\omega^2_{\text{peak}}>\mu_v\mu_w\left[1+\frac{n}{2}\frac{(k_v\beta\langle s\rangle)^n}{(\mu_v\mu_wK)^n+(k_v\beta\langle s\rangle)^n}\right]$$

$$\left(\gamma^2\frac{\left\langle|\eta_w|^2\right\rangle}{\left\langle|\eta_x|^2\right\rangle}+\mu_v^2+\mu_w^2+2J_{vw}\beta\right)<-\mu_v\mu_w\left[2+n\frac{(k_v\beta\langle s\rangle)^n}{(\mu_v\mu_wK)^n+(k_v\beta\langle s\rangle)^n}\right]$$

$$n\mu_v\mu_wM>(\mu_v+\mu_w)^2+2\gamma^2\frac{\left\langle|\eta_w|^2\right\rangle}{\left\langle|\eta_x|^2\right\rangle}, \tag{A2.43}$$

which is possible for large n and large $M=\langle w\rangle^n/(\langle w\rangle^n+K^n)$, which indicates that $K\ll\langle w\rangle$, in both cases representing a strong negative feedback.

2.A.6 Comments on Fig. 2.7c

Here we list some additional explanation on Fig. 2.7c. In this figure, we keep the parameters μ_v, μ_w, ν, $\beta\,(=J_{wv})$ and K constant, since they dictate the feedback cycle (Eq. 2.34 in Fig. 2.7b). We vary J_{xw} and μ_x, so that in this case, not the average production rate of x is constrained, but the average copy number $\langle x\rangle$.

To understand the dependence of the gain, noise and gain-to-noise ratio on $\gamma=J_{xw}$ and μ_x,we note that $g^2\sim\gamma^2g^2_{s\to w}/(\mu_x^2+\omega^2)$ and $N\sim\gamma^2/(\mu_x^2+\omega^2)N_v(\omega)+\gamma^2/(\mu_x^2+\omega^2)N_w(\omega)+N_x(\omega)$, where $N_v(\omega)$ and $N_w(\omega)$ are independent of γ and μ_x and $N_x(\omega)=2\gamma\langle w\rangle/(\mu_x^2+\omega^2)$ (with $\langle w\rangle$ being independent of γ and μ_x).

For $\omega\ll\mu_x$, the contributions of v and w to $N(\omega)$ are proportional to γ^2/μ_x^2, while the contribution of x is given by $N_x(\omega)\propto\gamma/\mu_x^2$. Hence, for $\omega\ll\mu_x$, the contributions of v and w to the noise are constant, while the contribution of x decreases with increasing γ and μ_x, leading to a decrease of $N(\omega)$. Since the gain is constant in this regime, the gain-to-noise ratio increases with increasing γ and μ for $\omega\ll\mu_x$. For $\omega\gg\mu_x$, the gain, and the contributions of v and w to the noise increase with γ^2 while the contribution of x to the noise increases with γ, meaning that also in this regime the gain-to-noise ratio increases with γ and μ_x.

Chapter 3

Information transmission in networks with feed-forward loops or diamond motifs

Using a Gaussian model, we study the transmission of time-varying biochemical signals through feed-forward motifs and diamond motifs. To this end, we compute the frequency dependence of the gain, the noise, as well as their ratio, the gain-to-noise ratio, which measures how reliably a network transmits signals at different frequencies. We find that both coherent and incoherent feed-forward motifs can either act as low-pass or high-pass filters for information: the frequency dependence of the gain-to-noise ratio increases or decreases with increasing frequency, respectively. Our analysis of diamond motifs reveals that cooperative activation of the output component can increase the gain-to-noise ratio. This means that from the perspective of information transmission, it can be beneficial to split the input signal in two and recombine the two propagated signals at the output. Cooperative activation can be implemented via the formation of homo- or heteromultimers that then bind and activate the output component or via the binding of individual molecules of the intermediate species to the output component.

Based on manuscript W.H. de Ronde, F. Tostevin and P.R. ten Wolde. Feed-forward loops and diamond motifs allow for tunable information transmission *Submitted*.

3.1 Introduction

Cells live in a highly dynamic environment. While in some cases cells may wish to ignore rapid fluctuations and only respond to persistent changes, in other cases they may have to do the opposite. For example, in the case of chemotaxis or osmo-adaptation, cells respond to changes in the stimulus but are insensitive to the absolute level of the stimulus. In contrast, in response to a changing sugar concentration, cells respond to the absolute steady-state sugar level, but may wish to integrate out rapid fluctuations of the sugar level. In general, to understand how cells cope with a changing environment, we have to understand how cells transduce time-varying signals. Moreover, given the observation that the biochemical networks which process the signals are stochastic in nature, we have to understand how reliably biochemical networks can process time-varying signals in the presence of noise.

Cells use recurring network motifs to specifically respond to temporal characteristics of the input signal. Negative feedback or incoherent feed-forward loops may be used to only respond to rapid variations and not to slow changes in the environment [101], while coherent feed-forward loops can be used to filter out transient fluctuations in the input and only respond to persistent changes in the environment [106]. To understand how specifically and reliably these motifs can respond to inputs with distinct temporal dynamics, we have to understand how they amplify input signals as a function of their frequency [65], which is characterized by the frequency-dependent gain. Moreover, we have to understand how they propagate biochemical noise as a function of frequency [65]. Indeed, information theory [45] tells us that the fidelity by which a signal of a given frequency is transmitted, is determined by the gain-to-noise ratio at that frequency [65]. We have recently shown for motifs with different types of feedback regulation that different network architectures affect the frequency dependence of the gain and the noise differently [107], which means that both of these quantities have to be studied together in order to understand how reliably a network transmits time-varying signals.

In this chapter, we use a Gaussian model to study the frequency dependence of the gain, noise and gain-to-noise ratio [65, 107] of feed-forward loops and diamond motifs. Both are common motifs in biochemical networks [11, 98]. Feed-forward motifs have been shown to regulate many different cellular processes, and, indeed, they exhibit very rich dynamics. Feed-forward motifs can act as sign-sensitive circuits [17], perform adaptation [101], provide fold-change detection [108] or attenuate extrinsic noise [109].

While the mean-field response [11, 17, 98, 101, 108, 109] and the noise characteristics [37, 41] of feed-forward loops and diamond motifs have been well characterized, how reliably they propagate time-varying signals has not been studied. This is of specific interest since the data-processing inequality dictates that information transmission decreases with the length of the cascade. If information transmission is the only constraint, the shortest cascade is the most reliable solution. However, as discussed above, cells may wish to respond to specific frequencies in the input signal and it may have to do so reliably. This is precisely what feed-forward loops and diamond motifs can achieve, in contrast to simple cascades.

We find that coherent feed-forward motifs and diamond motifs typically act as low pass filters for information: they transmit input signals of low frequency more reliably than input signals of high frequency. In contrast, incoherent feed-forward motifs tend to act as

high-pass filters for information. These results are not surprising: while for the coherent motifs the gain is large at low frequencies, for incoherent motifs the gain is large at high frequencies. We also show that, in contrast to the intuitive expectation, a coherent feed-forward loop can also act as a high-pass filter, while an incoherent feed-forward loop can also act as a low-pass filter for information. Our results also reveal that diamond motifs can have a higher gain-to-noise ratio over all frequencies than simple two-level cascades, when the total cost of making the proteins is the same in the two networks under comparison. This means that from the perspective of information transmission, it is beneficial to split the signal in two and combine the two transmitted signals again at the output. This could be considered as a form of coincidence detection. Interestingly, a diamond motif is not necessary: The same effect can also be achieved via cooperative activation of the output via an intermediate component. For example, the input may stimulate the formation of a homodimer or a homomultimer, which then activates the output; alternatively, the input activates a messenger, for example a transcription factor, which then activates the output, the gene promoter, in a cooperative fashion. Our analysis suggests that the gain-to-noise ratio increases with the level of cooperativity.

3.2 Methods

In this section, we briefly discuss the mathematical background of our analysis. A more in depth analysis is presented in [65, 67, 107]. The biochemical network consists of the components S, $\mathsf{V_i}$ (intermediate(s)) and X. Here S is the input and X is the output. We model the time evolution of these components using a set of coupled Langevin equations [69], which can be non-linear, e.g:

$$\frac{ds}{dt} = f_s(s) - \mu_s s + \Gamma(t), \tag{3.1}$$

$$\frac{dv_i}{dt} = f_{v_i}(s, \boldsymbol{v}) - \mu_{v_i} v_i + \eta_{v_i}(t), \tag{3.2}$$

$$\frac{dx}{dt} = f_x(s, \boldsymbol{v}, x) - \mu_x x + \eta_x(t). \tag{3.3}$$

For simplicity, we assume linear degradation of each component. The various noise sources (η_i, η_x) in Eqs. 3.2,3.3 are assumed to be independent and Gaussian distributed [31, 77, 96]. We take the noise strength $\langle|\eta_{v_i}|^2\rangle, \langle|\eta_x|^2\rangle$ as the sum of the average number of production and degradation events per unit amount of time for component v_i, x [68, 70, 97]. We assume the noise source $\Gamma(t)$ to be a Gaussian white noise. It generates an ensemble of input trajectories $s(t)$ with Gaussian statistics. The "forces" $f_{s,v_i,x}(s, \boldsymbol{v}, x)$ model all the reactions involving the production events of s, v_i and x.

We assume that the network has a steady state and linearize about this steady state, so that we get a dynamical equation for the fluctuations of each component, $\tilde{v}_i(t) = v_i(t) - \langle v_i\rangle$, and similarly for the input s and output x. In the linearized form, the relation between the components i, j is described by the coupling coefficient J, which are the Jacobian

elements, e.g.

$$J_{v_i v_i} = \frac{\partial \frac{dv_i}{dt}}{\partial v_i} = -\mu_{v_i}, \;\; J_{x v_i} = \frac{\partial \frac{dx}{dt}}{\partial v_i}. \tag{3.4}$$

We take as the input signal the variations of $s\,(t)$ around its mean $\langle s \rangle$, $\tilde{s}\,(t)$, and as the output the variations $\tilde{x}\,(t)$ of $x\,(t)$ around its mean $\langle x \rangle$. The mutual information rate between the in- and output trajectories, $s(t)$ and $x(t)$ respectively, is defined as [65]

$$R\left[\tilde{s}\,(t)\,, \tilde{x}\,(t)\right] = \frac{1}{2\pi} \int_0^\infty d\omega \ln \left[1 + \frac{g^2\,(\omega)}{N\,(\omega)} P_{ss}\,(\omega)\right], \tag{3.5}$$

where P_{ss} is the power spectrum of the signal,

$$P_{ss}\,(\omega) = \langle \tilde{s}\,(\omega)\, \tilde{s}\,(-\omega) \rangle\,. \tag{3.6}$$

The gain $g^2\,(\omega)$ and noise $N\,(\omega)$ are defined through the power spectra

$$g^2\,(\omega) \equiv \frac{|P_{sx}\,(\omega)|^2}{P_{ss}^2\,(\omega)}, \tag{3.7}$$

$$N\,(\omega) \equiv P_{xx}\,(\omega) - g^2\,(\omega)\, P_{ss}\,(\omega)\,. \tag{3.8}$$

These definitions are prescribed by using Eq. 3.5. A large gain-to-noise ratio $g^2\,(\omega)\,/N\,(\omega)$ (*GNR*) leads to a high mutual information rate (Eq. 3.5) and this implies reliable information transmission.

We have made a number of assumptions to obtain Eq. 3.5. First, we assume that the linearized system is an accurate representation of the non-linear system. Second, we assume that the variations $\tilde{s}\,(t)$ and $\tilde{x}\,(t)$ can be described by a Gaussian joint-probability distribution. Third, and last, we assume the signal $\tilde{s}\,(t)$ to be modular from the underlying network. Modularity of the signal with respect to the network indicates that no correlations exists between the variations in the input signal and the intrinsic noise of the reactions that constitute the processing network; it also implies that there is no feedback from the network onto the input signal. If signal modularity holds, then Eq. 3.8 is equal to the spectral addition rule [96]. In this case, the gain-to-noise ratio $g^2\,(\omega)\,/N\,(\omega)$, does not depend on the input signal, but only on the information transmission characteristics of the processing network; it describes the ability of the network to reliably propagate input signals as a function of their frequency. As an additional simplification we assume that no (anti)-correlations between the different noise sources are present [107]. While these may quantitatively change the results presented below, they do not qualitatively change them. In the next section we will describe the effect of the feed-forward motif on the information transmission through a biochemical network. We will characterize the gain, noise and gain-to-noise ratio, since these are intrinsic, signal-independent, properties of the network, when the spectral-addition rule holds [96]. We compare different motifs with simple linear one-step ($\mathsf{S} \rightarrow \mathsf{X}$) and linear two-step cascades ($\mathsf{S} \rightarrow \mathsf{V} \rightarrow \mathsf{X}$). We will compare the different

networks, unless specified otherwise, under the constraint that the average production rates of the respective components are the same in the networks under comparison.

Lastly, we will comment on the the differences between the gain-to-noise ratio , which describes information transmission and $P_{xx}(\omega)$, which describes power transmission.

3.3 Results

3.3.1 Simple Cascades

Simple cascades form the building blocks of the feed-forward motif. The feed-forward motif consists of two cascades, one in which the input S directly regulates the output X via a one-step cascade and one where S indirectly regulates the output X via a two-step cascade with an intermediate component V. A diamond motif consists of two two-step cascades which start and end at the same component, S and X, respectively. Since we will compare the behavior of these networks with simple cascades consisting of one or two steps, it will be useful to briefly recall their main transmission characteristics. A more detailed discussion can be found in [107] and Chapter 2.

For a one-step cascade the gain is given by $k_x^2/(\omega^2+\mu_x^2)$ where $k_x = J_{xs}$ is the coupling between s and x and μ_x is the lifetime of X , while the noise is given by $\langle|\eta_x|^2\rangle/(\omega^2+\mu_x^2)$. Consequently, the *GNR* of a one-step cascade is constant for all frequencies (Eq. A3.5). For the two-step cascade the gain is $(k_v k_x)^2/(\omega^2+m_x^2)(\omega^2+m_v^2)$, with two corner frequencies, μ_v, μ_x, the noise is $k_x^2\langle|\eta_v|^2\rangle/[(\omega^2+m_v^2)(\omega^2+m_x^2)]+\langle|\eta_x|^2\rangle/(\omega^2+m_x^2)$ and as a result the *GNR* decays with ω^{-2} for $\omega > (\mu_v + k_x\langle|\eta_v|^2\rangle/\langle|\eta_x|^2\rangle)^{1/2}$. For more details see App. 3.A.1.

3.3.2 The feed-forward motif

Two different types of feed-forward (ff) motifs exist. If the total regulatory effect of S on X and S via V on X are of the same nature, both either active or inactive, the motif is referred to as *coherent* feed-forward (cff). If the regulatory effects are opposing, the motif is referred to as *incoherent* feed-forward (iff). For the coherent motif we can further differentiate with respect to the integration strategy at the reporter X. If both S **and** V are required to produce X, the node X acts as an AND-gate and we refer to the motif as AND coherent feed-forward (acff), while if either S **or** V is sufficient to produce X, we refer to the motif as OR coherent feed-forward (ocff). The AND type is observed in the *ara* system of *E. coli* [17], while the OR type is present in the biosynthesis of the flagellar motor [110]. The motifs are shown in Figs. 3.1a (ocff), 3.2a (acff) and 3.3a (iff). In this chapter we study an iff motif for which the negative regulation is always at the response X. As a result, a distinction between AND and OR regulation in the iff is not made. We will discuss this assumption in more detail in the section discussing the iff. We now start by studying some general characteristics of the feed-forward motif.

General characteristics

We first study the gain. The gain for the feed-forward is

$$g^2(\omega) = \overbrace{\frac{J_{xs}^2}{\omega^2 + J_{xx}^2}}^{g^2_{s\to x}} + \overbrace{\frac{(J_{vs}J_{xv})^2}{(\omega^2 + J_{xx}^2)(\omega^2 + J_{vv}^2)}}^{g^2_{s\to v\to x}} - \overbrace{\frac{2J_{xs}J_{vs}J_{xv}J_{vv}}{(\omega^2 + J_{xx}^2)(\omega^2 + J_{vv}^2)}}^{\text{coherence}}, \tag{3.9}$$

The first term is the gain due to the direct regulation of X by S ($g^2_{s\to x}$), the second term is due to the pathway S $\to$ V $\to$ X ($g^2_{s\to v\to x}$) and the third term is a term due to the interaction between the two pathways.

The first term $g^2_{s\to x}$ is the gain of a one-step cascade in which the input s regulates the output x. It depends on the coupling constant J_{xs} and the lifetime of the protein X, $\mu_x^{-1} = -J_{xx}^{-1}$. If the lifetime μ_x^{-1} of X is longer than the timescale ω^{-1} on which the input signal varies, $\omega \gg \mu_x$, then variations in the input $s(t)$ are filtered out by the slow response of X.

The second term $g^2_{s\to v\to x}$ is the gain of a two-step cascade in which the input s regulates the output x via an intermediate v. It is seen that the gain of the two-step cascade is the product of the gain in each cascade step. This gives rise to two corner frequencies in $g^2_{s\to v\to x}$, one at $J_{vv} = -\mu_v$ and another at $J_{xx} = -\mu_x$, where μ_v and μ_x are the degradation rates of proteins V and X, respectively. The gain is large for $\omega \ll \mu_v, \mu_x$, since in this frequency regime both V and X can respond rapidly on the time scale of the signal variations, ω^{-1}, while for frequencies $\omega \gg \mu_v, \mu_x$, the gain decreases strongly, scaling as ω^{-4}, because in this regime the input variations are filtered by the finite lifetime of both V and X. Note that a lower degradation rate of the proteins increases the gain at low frequencies, but also reduces the corner frequencies beyond which the gain rapidly drops. This is a generic trade-off between the bandwidth of information transmission (the frequency range, bounded by the corner frequency, with large *GNR*) and the magnitude of the gain in the band.

The third term describes the coherence of the interaction between the signal transmitted via the direct pathway S $\to$ X and the signal transmitted via the indirect pathway S $\to$ V $\to$ X. Both signals originate from the source signal $s(t)$, which means that their variations are correlated. If $\mathrm{sgn}(J_{xs}) = \mathrm{sgn}(J_{vs}J_{xv})$ the third term is positive (since by construction $J_{vv} = -\mu_v$) and we have a coherent interaction. Such a coherent interaction is present in coherent feed-forward networks and leads to an increase in the gain. If $\mathrm{sgn}(J_{xs}) \neq \mathrm{sgn}(J_{vs}J_{xv})$, as in the incoherent feed-forward motifs, the coherence term is negative and thus the gain $g^2(\omega)$ is decreased.

It is instructive to compare the phase of the direct pathway $\phi_{s\to x}$ and the indirect pathway ($\phi_{s\to v\to x}$) at x

$$\phi_{s\to x} = \phi_s, \tag{3.10}$$

$$\phi_{s\to v\to x}(\omega) = \phi_s - \arctan\left[\frac{\omega}{\mu_v}\right]. \tag{3.11}$$

For $\omega = 0$, both signals are in phase. As the frequency increases, the phase of the signal that is transmitted via the indirect pathway decreases with respect to that which is transmitted via the direct pathway. At the corner frequency $-\mu_v = J_{vv}$ the phase difference is $\phi_{s\to v\to x}(J_{vv}) - \phi_s = -\pi/4$ and for $\omega \to \infty$, the signals are even more out of phase, $\phi_{s\to v\to x} - \phi_s = -\pi/2$. Combining both the phase and amplitude information, we see that the coherence between the two signals decreases with increasing ω for two related reasons. First, there is a decrease in the coherence term for large ω due to the time-averaging over the fast signal fluctuations resulting from the limited response time of X and V. Second, the coherence decreases because the phase difference increases (see App. 3.A.8).

Before we consider the noise, it is useful to briefly consider the three terms of the gain together. While the gain of the direct pathway scales for $\omega \gg \mu_x$ as ω^{-2}, the gain of the indirect pathway and the coherence term scale for $\omega \gg \mu_v, \mu_x$ as ω^{-4}. This means that for frequencies $\omega \gg \mu_v$ the gain of the direct pathway dominates.

The noise in the linearized feed-forward motif is

$$N(\omega) = \overbrace{\frac{J_{xv}^2 \left\langle |\eta_v|^2 \right\rangle}{(\omega^2 + J_{vv}^2)(\omega^2 + J_{xx}^2)}}^{N_{v\to x}(\omega)} + \overbrace{\frac{\left\langle |\eta_x|^2 \right\rangle}{(\omega^2 + J_{xx}^2)}}^{N_x(\omega)} \tag{3.12}$$

$$= g_{v\to x}^2(\omega) N_v(\omega) + \frac{\left\langle |\eta_x|^2 \right\rangle}{(\omega^2 + J_{xx}^2)} \tag{3.13}$$

It is seen that the total noise in x is independent of the regulatory effect of either pathway, since all terms are positive. The expression also reveals that the noise is the sum of two noise sources. One is the intrinsic noise arising from the stochastic production and decay of X, given by $N_x(\omega) = \langle|\eta_x|^2\rangle/(\omega^2 + J_{xx}^2)$. The other is the extrinsic noise coming from the stochastic production and decay of V, which is given by the intrinsic noise of v, $N_v(\omega) = \langle|\eta_v|^2\rangle/(\omega^2 + J_{vv}^2)$, multiplied by a frequency-dependent gain $g_{v\to x}^2 = J_{xv}^2/(\omega^2 + J_{xx}^2)$ which reflects how the noise from v is amplified by J_{xv} and integrated by x as a result of its finite lifetime. While the intrinsic noise in x, $N_x(\omega)$, scales as ω^{-2} for $\omega \gg \mu_x = -J_{xx}$, the extrinsic noise $g_{v\to x}^2(\omega)N_v(\omega)$ scales as ω^{-4} for $\omega \gg \mu_v, \mu_x$. Indeed, for $\omega \gg \mu_v$, the noise that originates from v in the indirect pathway becomes negligible.

Finally, we obtain the *GNR* for this three component motif

$$\frac{g^2(\omega)}{N(\omega)} = \frac{\left(\omega^2 + J_{vv}^2\right) J_{xs}^2 + \left(J_{vs}J_{xv}\right)^2 - 2J_{xs}J_{vs}J_{xv}J_{vv}}{J_{xv}^2 \left\langle |\eta_v|^2 \right\rangle + \left(\omega^2 + J_{vv}^2\right) \left\langle |\eta_x|^2 \right\rangle}. \tag{3.14}$$

While the noise is independent of the regulatory effect in the pathways, the gain, and hence the gain-to-noise ratio, depends on the total regulatory effect of each pathway (either S $\to$ X or S $\to$ V $\to$ X). This indicates that it is not important which of the reactions in a specific pathway acts negatively, only the overall effect of the pathway is important for information transmission.

We are now in a position to study in more detail the feed-forward motif. As discussed

above, for $\omega \gg \mu_v = -J_{vv}$, the gain of the direct pathway dominates the total gain, because the finite lifetime of V averages out the variations in the signal that are transmitted via the indirect pathway. Also the noise that originates at v in the indirect pathways becomes negligible in the total noise in x. Therefore, in this frequency regime, the direct pathway is dominant and the *GNR* becomes that of a one-step cascade, which means that it approaches a constant value, independent of frequency. For smaller frequencies, the behavior of the feed-forward motif depends on the relation between J_{xs}, J_{vs}, J_{xv} and $J_{vv} = -\mu_v$, which are the coupling constants of the direct and indirect pathway and the degradation rate of V, respectively (Eq. 3.4). When $|J_{xs}| \ll |J_{vs}J_{xv}/J_{vv}|$, then for $\omega \ll \mu_v$, the signal is transmitted more strongly via the indirect pathway than via the direct pathway, in which case the feed-forward motif resembles a two-step cascade (Figs. 3.1c, 3.2c, 3.3c, black solid line); clearly, in the limit that J_{xs} reaches zero, the feed-forward motif becomes a two-step cascade for all frequencies, in which case the the *GNR* (Eq. A3.10) is constant up to a cut-off frequency $\omega_c^2 = \mu_v(\mu_v + J_{xv})$, falling off as ω^{-2} for frequencies much larger than that [107]. When on the other hand $|J_{xs}| \gg |J_{vs}J_{xv}/J_{vv}|$, then the direct pathway dominates the gain for all frequencies (Figs. 3.1c, 3.2c, 3.3c, red solid line). Indeed, in terms of the gain, the feed-forward network effectively becomes a one-step cascade. However, the noise via the indirect pathway still contributes to the total noise and therefore this pathway effectively acts as a noise source. This has interesting consequences for the gain-to-noise ratio, as we describe in the next paragraph, since this allows any feed-forward motif to function as either a high-pass or a low-pass filter for information,

The gain-to-noise ratio (Eq. 3.14) of a feed-forward motif varies monotonically with frequency, but it can either be a decreasing or increasing function of ω. Indeed, both a coherent and an incoherent feed-forward motif can either act as a high-pass or low-pass filter for information. We can determine whether the *GNR* increases or decreases monotonically with frequency, by comparing the *GNR* at $\omega \to \infty$ to that at $\omega = 0$ (see Eq. 3.14). The GNR is monotonically increasing, meaning that the network acts as a high-pass filter for information, if

$$\left\langle |\eta_v|^2 \right\rangle > \left\langle |\eta_x|^2 \right\rangle \left(\frac{J_{vs}^2}{J_{xs}^2} + \frac{2\mu_v J_{vs}}{J_{xs}J_{xv}} \right). \tag{3.15}$$

The frequency of the inflection point is

$$\omega_{\text{ip}} = \sqrt{\frac{J_{xv}^2 \left\langle |\eta_v|^2 \right\rangle + J_{vv}^2 \left\langle |\eta_x|^2 \right\rangle}{3 \left\langle |\eta_x|^2 \right\rangle}}. \tag{3.16}$$

In the next two sections, we will discuss these conditions in more detail for the coherent and incoherent feed-forward motif, respectively.

The data processing inequality states that information that is lost cannot be recovered. Consequently, increasing the length of a cascade reduces information transmission. For this reason, for equal total production rate of the components within a cascade, the *GNR* of a one-step cascade is always larger than that of a feed-forward motif. The gain by itself

can be larger in a feed-forward motif than in a one-step cascade (Fig. 3.2c, dashed and solid red line). However, the intermediate component V in the indirect pathway introduces an additional noise source, which is not present in the one-step cascade. The coherence term in the gain is not large enough to compensate for this increase in the noise.

We have now specified some general characteristics. In the following we study the coherent and incoherent motif separately.

The coherent feed-forward motif

We will first compare the cff-motif, both the `OR` and `AND` types, with two-step cascades. Next we will compare the two motifs with each other. But first, we will start with a couple of observations which apply both to the acff and the ocff. Unless specified otherwise, we assume equal degradation rates for the respective components in the respective motifs and cascades.

The steady-state gain is determined by the average copy number X ($= \partial\langle x\rangle/\partial\langle s\rangle$). As a result, if the production of x in a cff equals the production of x in a two-step simple cascade, the gain at zero frequency ($\omega = 0$) is equal. Next, in the linear-noise approximation assumed here, the intrinsic noise arising from the production and degradation events of X, $N_x(\omega)$, is equal for the cff and the two-step cascade if the production rates of x are equal (see Eq. 3.12), since we assume throughout that the degradation rates are the same. If the production of v is equal as well the noise $N_v(\omega)$ is equal in the cff and the two-step cascade. However, the transmitted noise $N_{v\to x}(\omega)$, and thus the total noise $N(\omega)$ can be different, as discussed in more detail below. With respect to the *GNR*, we observe that for large ω, the cff always has a larger *GNR* than the two-step cascade, while for small ω, no general results can be presented. Comparing to the one-step cascade, the *GNR* of the cff is always smaller than that of a one-step cascade if the total production in both networks is equal. This is because η_v acts as an additional noise source that corrupts the signal (Figs. 3.1c,3.2c, bottom).

The *GNR* of a coherent feed-forward motif either increases monotonically with frequency or decreases monotonically with frequency, as mentioned in the previous section. If Eq. 3.15 is satisfied, it increases monotonically, and the motif acts as a high-pass filter for information. We can intuitively understand the terms in Eq. 3.15 as follows. A decrease in the ratio J_{vs}/J_{xs} means that the input signal s is relayed more to x directly than to x via v. However, while the direct pathway S $\to$ X contributes to information transmission at all frequencies — its *GNR* is flat — the indirect pathway S $\to$ V $\to$ X only contributes at low frequencies — the *GNR* of a two-step cascade falls of as ω^{-2} for high frequencies. The effect of the indirect pathway, both the gain and the noise, at high frequencies becomes negligible. Indeed, in the limit that J_{vs}/J_{xs} reaches zero, the signal is transmitted completely via the direct pathway only; yet, while the *GNR* of the direct pathway, a one-step cascade, is flat, the *GNR* of the cff increases with frequency, because the indirect pathway still adds noise to the signal, especially in the low frequency regime. The presence of the second term on the right-hand side of Eq. 3.15 can be understood by noting that it arises from the interplay between the noise (Eq. 3.12) and the coherence term in the gain (Eq. 3.9). We can understand the dependence on $\mu_v = -J_{vv}$ by noting that at high frequencies $\omega \gg \mu_v$ the coherence and the noise coming from v hardly contribute to the gain and the total noise $N(\omega)$ respectively, while at $\omega = 0$ the coherence decreases with increasing degradation rate as μ_v^{-1}

(Eq. 3.9) while the noise coming from v decreases as μ_v^{-2} (Eq. 3.12). Similarly, at $\omega = 0$, the coherence increases with J_{xv} while the noise coming from v increases with J_{xv}^2. Thus at low frequencies decreasing the degradation rate μ_v and/or increasing the coupling between V to X, J_{xv}, increases the noise more than it does the gain, thus reducing the *GNR* at low frequencies, while at high frequencies the influence on the *GNR* is negligible.

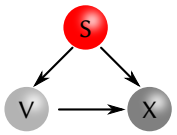
We now discuss the ocff, which combines the two pathways at X according to OR logic. We first compare the ocff with a two-step cascade on the footing of equal production costs for each of the components separately, For $\omega > 0$, the gain of the ocff is always larger than the gain of the two-step cascade (Table. 3.2): for $0 < \omega < \mu_v, \mu_x$ the gain of the cff is boosted by the coherent interaction between the signals propagated via the direct and indirect pathway, while for $\omega > \mu_v$ the signal is attenuated in the two-step cascade by the finite lifetime of V whereas it can still be propagated in the ocff via the direct pathway. The noise $N(\omega)$ in the ocff motif is smaller than that in the two-step cascade for all ω. The intrinsic noise $N_x(\omega)$ is equal in both networks. However, the extrinsic noise coming from v, $g^2_{s\to v}(\omega)N_v(\omega)$ (see Eq. 3.12) is not. Since the intrinsic noise in v, $N_v(\omega)$, is the same in both networks because the production of v is equal, this means that the difference must lie in how this noise is propagated to x, which is given by $g^2_{s\to v} = J^2_{xv}/(\omega^2 + J^2_{xx})$; since J_{xx} is the same, this means that the coupling between V and X in the ocff motif, J^{cff}_{xv}, must be less than that in the two-step cascade, J^{ts}_{xv}. Indeed, in the ocff the production of x depends on both s and v, while in the two-step cascade it only depends on v. Together with the constraint that the production rates of X are equal, this indeed implies that $J^{\text{ocff}}_{xv} < J^{\text{ts}}_{xv}$: $J^{\text{ts}}_{xv}\langle v\rangle = J^{\text{ocff}}_{sv}\langle s\rangle + J^{\text{ocff}}_{xv}\langle v\rangle$, showing that $J^{\text{ocff}}_{xv} < J^{\text{ts}}_{xv}$. In the ocff network, the noise in x is thus smaller because less noise is propagated from v because of the smaller amplification of the transmitted signal between v and x. The higher gain and the lower noise means that the gain-to-noise ratio of the ocff motif is higher than that of a two-step cascade, as shown in Fig. 3.1c. The ocff is thus able to signal more reliably than the corresponding two-step cascade.

It can be shown for the ocff motif that for equal production of the species separately in the network, the plateau value of the *GNR* in the high-frequency regime is higher for the high-pass filter than for the low-pass filter. Therefore, if the goal is to transmit signals reliably at high frequencies, then a suitable parameter set can be chosen that yields a high-pass filter with a high *GNR* at high frequencies, irrespective of the *GNR* behavior at low frequencies (Fig. 3.1d, red dashed line). For different constraints on the production, for example constraining the production of only x, or constraining the total production of all components, the analysis is more difficult. But, as discussed above, for large ω, the ocff always has a larger *GNR* than a two-step cascade. The results are summarized in Tables. 3.1,3.2.

In the next paragraphs we study the acff motif. The acff motif combines the two paths at X according to AND-logic and is described by Eq. 3.20. The linearized system has an identical structure to the ocff, and therefore the results are qualitatively similar to those of the ocff. However, quantitatively the results can be different.

First we compare the acff with the corresponding two-step cascade. For equal production of v and x individually the gain of the acff (Eq. A3.29) is always larger than the gain of the two-step cascade (Table. 3.3). This again is due to the coherent interaction between the

a)

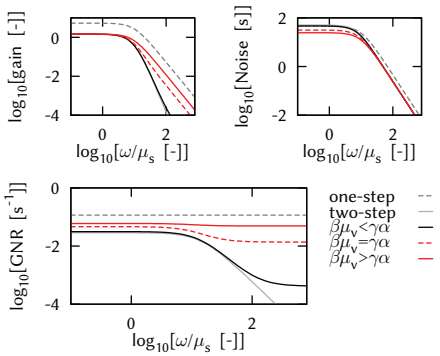


b)

$$\frac{dv}{dt} = \alpha s - \mu_v v + \eta_v(t), \tag{3.17a}$$

$$\frac{dx}{dt} = \beta s + \gamma v - \mu_x x + \eta_x(t). \tag{3.17b}$$

c)



d)

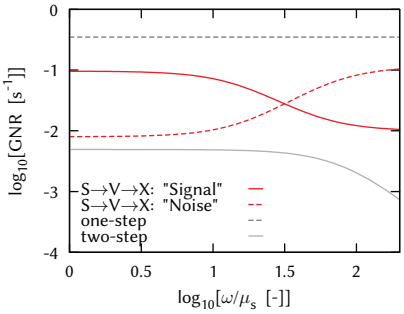


Figure 3.1: The OR coherent feed-forward motif. **a)** In the ocff, either S and V is required to produce X. **b)** The Langevin equations for the ocff. **c)** For equal production in the total cascade the *GNR* is shown for different weighings of the two pathways. If the pathway S→X dominates ($\beta\mu_V > \alpha\gamma$) (solid red), the ocff is similar to a one-step cascade (gray dashed) although with smaller gain. If the pathway S→V→X dominates (black solid) two-step cascade (gray solid) behavior is obtained. But for large ω the ocff is similar to the one-step cascade (scaling as ω^0), because the signal and noise fluctuations through v are averaged out due to the finite response time of V. Parameters: k_S=10, μ_V=10, μ_X=10, k_X^1=23.1, k_V^2=11, k_X^2=11, α=11, γ=10.1,5.8,1.1 and β=1,5.8,10.1 for respectively red solid, red dashed, black solid lines. β sets the production of x in the cascade, respectively μ_S sets the timescale. **d)** The gain-to-noise ratio for a cff-motif can both have high-pass and low-pass characteristics. For high-pass characteristics, the indirect pathway functions as a "noise" source at low frequencies, while at high frequencies the noise is filtered out due to the finite response time of V. Parameters: k_S=100, μ_V=5, μ_X=10, k_X^1=70, k_V^2=1, k_X^2=690, α=20,0.3, γ=20,1000,β=10,40, respectively solid, dashed.

direct and indirect pathway at low frequencies, $\omega \ll \mu_v, \mu_x$, while for high frequencies, $\omega \gg \omega_v, \omega_x$, the gain of a two-step cascade falls of more rapidly with frequency than the gain of the direct pathway. Interestingly, the mathematical dependence of the total noise $N(\omega)$ on the network parameters ($J_{xv}, \mu_v, \ldots$) of the acff is the same as that of the two-step cascade. Since $N_v(\omega)$ and $N_x(\omega)$ are equal in both networks, this implies that the noise propagated from v to x ($N_{v\to x}(\omega)$) is equal:

$$N_{v\to x,\texttt{AND}} = \frac{\beta^2 \langle s \rangle^2 \left\langle |\eta_v|^2 \right\rangle}{(\omega^2 + \mu_x^2)(\omega^2 + \mu_v^2)}, \tag{3.18}$$

$$N_{v\to x,\text{ts}} = \frac{k_x^2 \left\langle |\eta_v|^2 \right\rangle}{(\omega^2 + \mu_x^2)(\omega^2 + \mu_v^2)}. \tag{3.19}$$

The total rate of production of x by v in the two-step cascade is $k_x\langle v\rangle$ (Eq. A3.7), while, in the linear-noise approximation used here, the average total rate of production of x by s and v in the acff motif is $\beta\langle s\rangle\langle v\rangle$ (Eq. A3.28); the latter means that the coupling of x to v is given by $J_{xv} = \beta\langle s\rangle$. For equal rate of production of x in the two-step cascade and the acff — $k_x\langle v\rangle = \beta\langle s\rangle\langle v\rangle$ — the coupling strengths $J_{xv} = k_x = \beta\langle s\rangle$ in the two networks are the same, meaning that the extrinsic noise coming from v, $N_{v\to x}$ is indeed the same. Since the gain is larger, but the noise is the same, the *GNR* of the acff is always larger than that of a two-step cascade.

Again, we can choose different constraints for the production. Let's constrain the total production of both v and x together to be the same in the two-step cascade and the acff, but we do not require equal production of v and x in both networks separately. For large ω the *GNR* of the acff is always larger than that of the two-step cascade. With this constraint this is not necessarily true for small ω. At low frequencies the two-step cascade can have a larger *GNR*. Intuitively, this is possible if one of the two steps in the two-step cascade is fast, such that it effectively reduces to a one-step cascade at small frequencies, or if the indirect pathway acts as a noise source and is strongly coupled to the response X (see Eq. A3.34).

The acff can not transmit information better at small frequencies than a one-step cascade, given equal total production. For equal production of x — thus at the cost of producing more V — the acff can perform better than the one-step cascade (Eq. A3.32). This is in contrast to the ocff, which also under this constraint has a smaller *GNR*. The results are summarized in Tables. 3.3,3.4.

Finally, we compare the two motifs ocff and acff for equal separate production of V and X. The acff has a larger gain, because it has a steeper response function at equal production. Not only the gain, but also the intrinsic noise ($N_{v\to x}$) is larger in the acff. These terms have an opposite effect on the *GNR*. For $\omega \to \infty$ the acff has a larger *GNR*, since the noise fluctuations from v are averaged out and only the gain remains important. For small ω, however, no general conclusion for the *GNR* can be drawn, but a few observations can be made. An increase in the coupling J_{xs} of S $\to$ X (increasing β) leads to a larger *GNR* for the ocff than for the acff. In this case, the direct coupling between S and X dominates in the ocff, whereas this connection is not directly present in the acff due to the AND-logic integration strategy. The ocff 'averages' the effects of the direct and indirect pathway, while for the acff-gate the pathways are more connected (Fig. 3.2d). Indeed, one necessary requirement for the ocff to have a larger *GNR* than the acff at low frequencies is that the direct pathway dominates the output. Another necessary requirement for this is that the *GNR* of the acff has to have high-pass characteristics (Eqs. A3.48,A3.49).

The incoherent feed-forward motif

The last class of feed-forward loops that we will consider is the incoherent feed-forward (iff) motif (Fig. 3.3a). We will discuss only two different types of iff motifs: one in which negative regulation is implemented in the direct pathway and one in which it is present in the indirect pathway; in the latter case, we always take the negative interaction onto X, even though we could also have chosen the alternative in which S acts negatively onto V (and V acts positively on X). It is known that the precise topology can influence the noise behavior [38, 111]; however, the qualitative behavior of the *GNR* of the different possible architectures

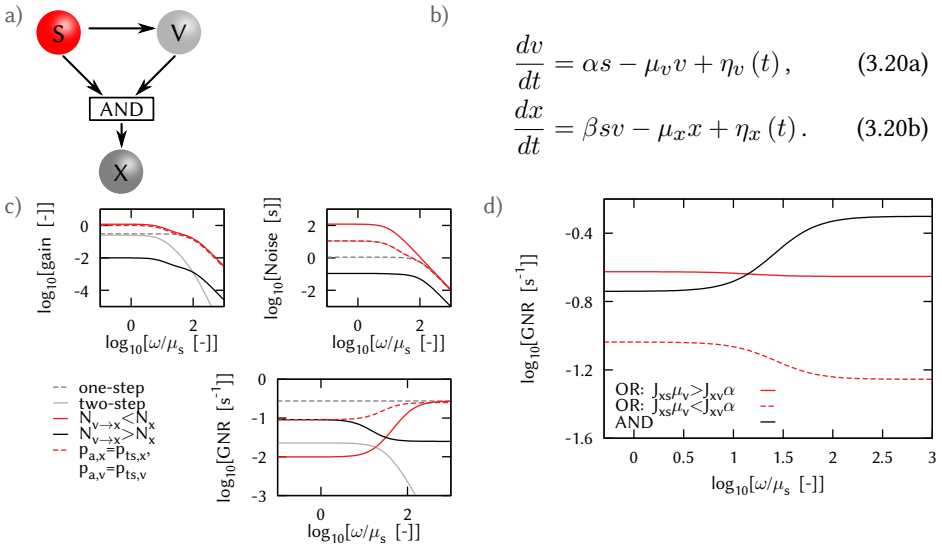


Figure 3.2: The AND coherent feed-forward loop. **a)** In the acff, both S and V are required to produce X. **b)** The Langevin equations for the acff. **c)** The *GNR* of the acff for different weighings (changing α) of the two pathways for equal total production and equal degradation rates. First, for large ω the direct pathway is dominant and both gain and noise (Eqs. A3.29, A3.30) scale with ω^{-2}, leading to a constant *GNR*, which is larger than that of a two-step cascade. If α is small (red solid), the production of v, p_v is small. To compensate for the small production of v, the production of x should be large, leading to a large β. Note that $\alpha\beta\neq1$. The gain $g^2_{v\to x}$ scales with $\beta^2\alpha^2$, while the noise $N_{v\to x}\propto\beta^2\alpha$, and $N_x\propto\beta\alpha$. For small ω, therefore the *GNR* scales as α, while for $\omega\to\infty$, $N_{v\to x}$ is averaged out and *GNR* scales as $\alpha\beta^0$. The dependence of the gain and the noise on μ_v is slightly different. Therefore, a small bandwidth exists for which the gain decreases more slowly than the noise for increasing ω (red solid). For large α the opposite reasoning holds. The red dashed line shows the acff with equal production of v and x individually, as the two-step cascade. Both the gain, noise and *GNR* are larger. It is interesting to note that for small frequencies and equal total production, the acff can have a smaller gain-to-noise ratio than the simple two-step cascade (red solid and gray solid). This is the case if $N_{v\to x}(\omega)$ becomes dominant. Parameters: k_s=100, μ_v=10, μ_x=100, k^1_x=55,k^2_x=100 k^2_v=5, α=0.5,5,50 respectively red solid, red dashed, black solid. β such that the total production is equal, respectively β=10.9,1,0.01. μ_s set the timescale. **d)** The acff (black) for large ω always has a larger *GNR* than the ocff (red solid and dashed). However, for small ω, depending on kinetic rates, the *GNR* of the ocff can be larger than the acff (red solid). For this to occur, in the ocff the direct pathway S→X should dominate the indirect S→V→X pathway. Then, for small ω the ocff acts as an one-step cascade. For the acff this is not the case, because the coupling between s and v is more complex. Parameters:k_s=100, μ_v=10, μ_x=100, $J^{and}_{vs}=J^{or}_{vs}$=10, β=1, J^{or}_{xv}=16,67 and J^{or}_{xs}=83,33 respectively solid and dashed. μ_s sets the timescale.

is similar to that of the networks studied here. For a more detailed discussion, we refer to App. 3.A.5.

An inspection of the gain in Eq. 3.9 reveals that the coherence term is negative in both types of incoherent feed-forward motifs, since either J_{xs} or the product $J_{xv}J_{vs}$ is negative, respectively. This leads to a reduction of the gain on timescales that are smaller than the

response times of V and X ($\omega < \min\{\mu_v, \mu_x\}$). For $\omega < \mu_v$, the two pathways are exactly out of phase, i.e the phase difference is $-\pi$, while as ω increases, the phase difference reduces to $-\pi/2$ (Eq. 3.10).

We model the incoherent regulation using a repressive Hill function (Eq. 3.23b), which is a commonly used coarse-grained description of protein interactions or gene or enzyme regulation. If the direct pathway is repressive, the repression strength depends on the ratio $\langle s\rangle/K$, with K being the value of s at which X is reduced to half its maximal value. In the limit that $K \gg \langle s\rangle$, the repression is very weak, and the influence of the direct pathway is negligible. The iff then effectively reduces to a two-step cascade. In the opposite limit, repression is very strong and the iff becomes adaptive [101]. This means that $\langle x\rangle$ does not depend on $\langle s\rangle$. Indeed,

$$\langle x\rangle = \frac{\nu}{\mu_x}\frac{K\langle v\rangle}{K+\langle s\rangle} = \frac{\nu\alpha}{\mu_x\mu_v}\frac{K\langle s\rangle}{K+\langle s\rangle} \approx \frac{K\nu\alpha}{\mu_x\mu_v}. \tag{3.21}$$

In terms of the frequency response, adaptation to constant signals corresponds to a zero-gain at zero frequency; $g^2(\omega = 0) = 0$

$$\begin{aligned}\lim_{K\ll\langle s\rangle} g^2(\omega=0) &= \frac{(J_{vs}J_{xv} - J_{vv}J_{xs})^2}{\mu_v^2\mu_x^2} \\ &= \frac{K^2}{(K+\langle s\rangle)^2}\frac{(\alpha\nu)^2}{\mu_v^2\mu_x^2}\left[1 - \frac{\langle s\rangle}{K+\langle s\rangle}\right]^2 \approx 0.\end{aligned} \tag{3.22}$$

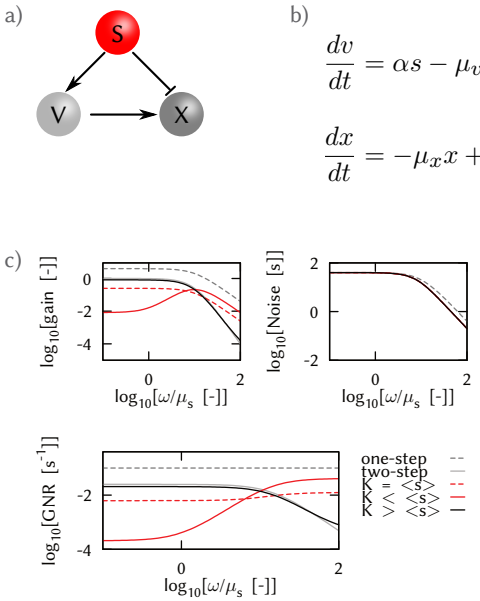
For large frequencies, only the direct pathway, which in this example is repressive, transmits information. For information transmission it is not important whether the pathway acts negatively or positively on X. The variations in s still affect the variations in x, but with an opposing sign. For the motif where the repression occurs in the indirect pathway, similar conclusions hold, but now the motif functions as an one-step cascade in the case that repression is weak.

An incoherent feed-forward motif with a strong negative interaction in one of the two pathways acts as a high-pass filter for information. This is because of the destructive interference of the two pathways at small frequencies, $\omega \ll \mu_v, \mu_x$. For higher frequencies, the gain increases because the phase difference between the two pathways decreases and also because the indirect pathway becomes less important as the finite lifetime of V increasingly averages out the variations in S. The gain therefore has high-pass characteristics. Since the noise (Eq. 3.12) is not affected by the destructive interference and has low-pass characteristics, the *GNR* is high-pass (Fig. 3.3c, red line).

For the incoherent motif, a low-pass *GNR* is observed only if the negative regulation is small (e.g. $K \gg \langle s\rangle$) and the direct pathway is negligible compared to the indirect pathway. Then the motif for small ω resembles a two-step cascade (Fig. 3.3c, black solid line), which indeed exhibits a low-pass *GNR*. For large ω the direct pathway will dominate, which means that in contrast to a two-step cascade, the *GNR* reaches a constant as a function of frequency for $\omega \gg \mu_v, \mu_x$.

Finally, we compare, for completeness, the iff-motif to a two-step cascade. For equal

production of v and x separately, the iff has a lower *GNR* for small ω than the two-step cascade (Table. 3.6). This is because of the destructive interference between the direct and indirect pathway in the iff, which reduces the gain at low frequencies.



$$\frac{dv}{dt} = \alpha s - \mu_v v + \eta_v(t), \tag{3.23a}$$

$$\frac{dx}{dt} = -\mu_x x + \eta_x(t) + \begin{cases} \dfrac{\nu K}{K+s} v, & \text{repressive direct} \\ \dfrac{\nu K}{K+v} s, & \text{repressive indirect} \end{cases}. \tag{3.23b}$$

Figure 3.3: The incoherent feed-forward motif. **a)** The incoherent motif. Left: The pathway S→V→X is positive, while the pathway S→X is negative. **b)** The Langevin description of the network **c)** The coherence terms acts destructively on small timescales, reducing the gain, but not the noise, so that the *GNR* is reduced. The gain-to-noise ratio of the incoherent motif with a repressive direct link S→X has different characteristics, for three different values of K (Eq. 3.23). If K>⟨s⟩ (black solid) the negative regulation by S diminishes and the iff motif resembles a two-step cascade. For large ω, the indirect pathway is averaged out; at these frequencies one-step characteristics are observed. If K<⟨s⟩ (red solid), the iff becomes purely adaptive, leading to g^2=0 for ω=0. Note that the lines in the noise plot are overlapping. gray dashed: one-step cascade, gray solid: two-step cascade. Parameters: k_s=100, α=1, μ_v=1, μ_x=1 and ν=110,20,11 for K/⟨s⟩=0.1,1,10. μ_s set the timescale.

3.3.3 Multimerization

In this section we examine multimerization of intermediate signaling components. In this motif, an intermediate component v is activated by the input signal s, which then cooperatively activates the output component x (Fig. 3.4a). The intermediate components could form a protein complex that then binds and activates the output, but it need not be: the intermediate component could also bind the output component, which could be a gene promoter or an enzyme, separately but cooperatively, thereby activating it. This is a common motif in gene regulation and also enzyme regulation.

This system is described by

$$\frac{dv}{dt} = \alpha_v s\left(t\right) - \mu_v v\left(t\right) + \eta_v\left(t\right), \tag{3.24}$$

$$\frac{dx}{dt} = \gamma_v \frac{v^n\left(t\right)}{K^n + v^n\left(t\right)} - \mu_x x\left(t\right) + \eta_x\left(t\right) \tag{3.25}$$

$$\approx \gamma_n v^n\left(t\right) - \mu_x x\left(t\right) + \eta_x\left(t\right). \tag{3.26}$$

Here, n is a measure for the cooperativity, the number of V molecules that are required to activate X. We assume that the concentration v is very low , $K \gg v$, in which case Eq. 3.25 reduces to Eq. 3.26, with $\gamma_n \simeq \gamma_v/K^n$. For this network, the concentration of X and the coupling between v and x, J_{xv}, depends on the degree of cooperativity n. The concentration X is given by $\langle x\rangle = \gamma_n\langle v^n\rangle/\mu_x$. In the linear-noise approximation this concentration is given by $\langle x\rangle \simeq \gamma_n\langle v\rangle^n/\mu_x$, and the coupling $J_{xv,n} = n\gamma_n\langle v\rangle^{n-1}$, where the subscript n in $J_{xv,n}$ indicates that we consider the coupling between v and x when the degree of cooperativity is n. Since we compare the networks on the footing of equal productions costs and the degradation rates are kept constant, we find that the coupling constant $J_{xv,n}$ for a system in which n V molecules are required to activate X is related to the coupling constant $J_{xv,n=1}$ for a system in which only one V molecule is required to activate X, simply via $J_{xv,n} = nJ_{xv,n=1}$.

For this motif, the gain is given by

$$g^2\left(\omega\right) = \frac{n^2 J_{xv,n=1}^2 J_{vs}^2}{\left(\omega^2+\mu_x^2\right)\left(\omega^2+\mu_v^2\right)}. \tag{3.27}$$

It is seen that the gain increases with the cooperativity n.

The noise for this motif is given by

$$N\left(\omega\right) = \overbrace{\frac{n^2 J_{xv,n=1}^2\left\langle\left|\eta_v\right|^2\right\rangle}{\left(\omega^2+\mu_v^2\right)\left(\omega^2+\mu_x^2\right)}}^{N_{v\to x}(\omega)} + \overbrace{\frac{\left\langle\left|\eta_x\right|^2\right\rangle}{\left(\omega^2+\mu_x^2\right)}}^{N_x(\omega)} \tag{3.28}$$

$$= g_{v\to x}^2\left(\omega\right) N_v(\omega) + \frac{\left\langle\left|\eta_x\right|^2\right\rangle}{\left(\omega^2+\mu_x^2\right)}. \tag{3.29}$$

Clearly, increasing n increases the extrinsic noise in x that originates in v, $N_{v\to x}(\omega) = g_{v\to x}^2 N_v(\omega)$ but not the intrinsic noise in x, $N_x(\omega)$. Increasing the coupling $J_{xv,n} = nJ_{xv,n=1}$ does not change the intrinsic noise, but it does affect how fluctuations in v are amplified at the level of x.

The *GNR* then reads

$$GNR = \frac{J_{vs}^2 J_{xv,n=1}^2}{J_{xv,n=1}^2\left\langle\left|\eta_v\right|^2\right\rangle + \frac{1}{n^2}\left(\omega^2+\mu_v^2\right)\left\langle\left|\eta_x\right|^2\right\rangle}. \tag{3.30}$$

From Eq. 3.30 it is clear that the *GNR* increases with the cooperativity n. This is because while the overall gain $g_{s\to x}^2$ and and the extrinsic noise $N_{v\to x}(\omega)$ both increase with n, the intrinsic noise $N_x(\omega)$ does not. Interestingly, not only the amplitude of the *GNR* increases, but also the knee frequency and thus the bandwidth for reliable information transmission. Indeed, the knee frequency ω_c is set by $\omega_c^2 = \mu_v^2 + n^2 J_{xv,n=1}^2 \langle|\eta_v|^2\rangle / \langle|\eta_x|^2\rangle = \mu_v(\mu_v + n^2\gamma_{n=1})$, showing that it increases with n (Fig. 3.4b).

To summarize, we observe that information transmission can be increased by cooperatively activating the output. This could either be achieved via homomultimerization of the intermediate component, or by separate binding of the intermediate molecules to the output component. We note that an increase in n increases the non-linearity of our system and therefore the approximation might break down. However, for the parameters used here (Fig. 3.4b), numerical simulations of the non-linear system agree very well with the linear theory (see App. 3.A.7).

3.3.4 Diamond motif

The multimerization motif discussed above could be considered to be a special case of a diamond motif, in which the intermediate components are identical. Here, we consider the general scheme in which they are different: the diamond motif [13, 102, 105, 112, 113]. We will compare the *GNR* of this motif to that discussed above. Moreover, we will compare the performance of this motif to that of a two-step cascade; if the *GNR* of the diamond motif is higher than that of a two-step cascade, then this indicates that from the perspective of information transmission it is beneficial to split the signal after the input and recombine the signals downstream, at the output.

We will consider a diamond motif in which the two intermediate components U and W can either form a homodimer U_2 or W_2, respectively, or a heterodimer UW; note that in this network the intermediate components effectively activate the output via `AND` logic. We will compare the performance of this motif to that of a homodimer motif in which there is only one intermediate component V, which forms a homodimer V_2; this corresponds to the scenario discussed before with $n = 2$. The diamond motif is described by

$$\begin{aligned}
\frac{du}{dt} &= \alpha s\,(t) - \mu_u u\,(t) + \eta_u\,(t)\,, \\
\frac{dw}{dt} &= \beta s\,(t) - \mu_w w\,(t) + \eta_w\,(t)\,, \\
\frac{dx}{dt} &= \gamma_u u^2\,(t) + \gamma_w w^2\,(t) + 2\gamma_{uw} u\,(t)\,w\,(t) - \mu_x x\,(t) + \eta_x\,(t)\,.
\end{aligned} \tag{3.31}$$

The factor 2 is introduced so that the diamond motif reduces to the motif with only one intermediate component V when the properties of the components U and W are identical and equal to those of V. If $\gamma_u = \gamma_w = 0$, only heterodimer activation is possible. In general, the values of the γ's might be different depending on the respective binding kinetics of the components U and W.

The gain of the diamond motif of Eq. 3.31 is

$$g^2(\omega) = \overbrace{\frac{(J_{xu}J_{us})^2}{(\omega^2+\mu_x^2)(\omega^2+\mu_u^2)}}^{g^2_{s\to u\to x}} + \overbrace{\frac{(J_{xw}J_{ws})^2}{(\omega^2+\mu_x^2)(\omega^2+\mu_w^2)}}^{g^2_{s\to w\to x}} + \overbrace{\frac{2J_{xu}J_{us}J_{xw}J_{ws}\left(\omega^2+\mu_u\mu_w\right)}{(\omega^2+\mu_x^2)(\omega^2+\mu_u^2)(\omega^2+\mu_w^2)}}^{\text{coherence}}. \tag{3.32}$$

Here, $J_{us}=\alpha$, $J_{ws}=\beta$, $J_{xu}=2\gamma_u\langle u\rangle+2\gamma_{uw}\langle w\rangle$, $J_{xw}=2\gamma_w\langle w\rangle+2\gamma_{uw}\langle u\rangle$. The first two terms in Eq. 3.32 describe the gain due to the transmission of the input signal via the pathways containing U and W, respectively, while the third term describes the coherence of their interaction at the output X. It can be verified that when U and W are identical and equal to V, meaning that $\alpha=\beta$ and $\alpha_v=2\alpha$ and $J_{us}=J_{ws}=J_{vs}/2$ and $\langle v\rangle=\langle w\rangle+\langle u\rangle$, $\gamma_{uw}=\gamma_u=\gamma_w=\gamma_v=\gamma$ and $J_{xu}=J_{xw}=J_{xv,n=2}=2\gamma\langle v\rangle$, and $\mu_u=\mu_w=\mu_v$, the gain of the diamond motif equals that of the homodimer motif.

The noise of the diamond motif of Eq. 3.31 is

$$\begin{aligned} N(\omega) &= \overbrace{\frac{J_{xu}^2\left\langle|\eta_u|^2\right\rangle}{(\omega^2+\mu_u^2)(\omega^2+\mu_x^2)}}^{N_{u\to x}(\omega)} + \overbrace{\frac{J_{xw}^2\left\langle|\eta_w|^2\right\rangle}{(\omega^2+\mu_w^2)(\omega^2+\mu_x^2)}}^{N_{w\to x}(\omega)} + \overbrace{\frac{\left\langle|\eta_x|^2\right\rangle}{(\omega^2+\mu_x^2)}}^{N_x(\omega)} \\ &= g^2_{u\to x}(\omega)\,N_u(\omega) + g^2_{w\to x}(\omega)\,N_w(\omega) + N_x(\omega). \end{aligned} \tag{3.33}$$

where we have exploited that $\langle\eta_u(\omega)\eta_w(-\omega)\rangle=0$ when U and W are different. If the properties of U and W are identical and equal to V, then $J_{xu}=J_{xw}=J_{xv,n=2}$, $\mu_u=\mu_w=\mu_v$, $\langle|\eta_u|^2\rangle=\langle|\eta_w|^2\rangle=\langle|\eta_v|^2\rangle/2$ and the noise of the diamond motif is indeed equal to that of the homodimer motif.

The gain-to-noise ratio for the diamond motif is

$$\frac{g^2}{N} = \frac{(J_{xu}J_{us})^2 H(\omega,\mu_w) + (J_{xw}J_{ws})^2 H(\omega,\mu_u) + 2J_{xu}J_{vs}J_{xw}J_{ws}\left(\omega^2+\mu_u\mu_w\right)}{H(\omega,\mu_w)\,J_{xu}^2\left\langle|\eta_u|^2\right\rangle + H(\omega,\mu_u)\,J_{xw}^2\left\langle|\eta_w|^2\right\rangle + H(\omega,\mu_u)\,H(\omega,\mu_w)\left\langle|\eta_x|^2\right\rangle}, \tag{3.34}$$

where $H(x,y)=x^2+y^2$.

We now compare the *GNR* of the diamond motif to that of the homodimer motif. To compare on equal footing, we will assume in what follows below that the production rate of x is equal in the two motifs – $\gamma_v\langle v\rangle^2 = 2\gamma_{uw}\langle u\rangle\langle w\rangle+\gamma_u\langle u\rangle^2+\gamma_w\langle w\rangle^2$ – and that the production rate of v in the homodimer motif equals the sum of that of u and w in the diamond motif; $\alpha_v\langle s\rangle=\alpha\langle s\rangle+\beta\langle s\rangle$. If the degradation rates of the intermediate components in the two motifs are equal – $\mu_v=\mu_u=\mu_w$ – and the coupling between the intermediate components and the output is the same – $\gamma_{uw}=\gamma_u=\gamma_w$ – then the *GNR*

of the diamond motif is equal to that of the homodimer motif as it should, since there is no distinction between the components. If the components U and W are different, leading, for example, to different couplings ($\gamma_u \neq \gamma_w$, still assuming equal degradation rates), the homodimer motif has a larger *GNR* than the diamond motif at small $\omega < \mu_u, \mu_w, \mu_v$. In this case, the gain is equal for both processes because the concentration of X is taken to be the same in both networks. Hence, the difference in the *GNR* originates from the noise. The noise term $N_X(\omega)$ is equal in both networks, because the production and degradation rates of X are taken to be the same in the two networks. However, the extrinsic noise propagated from the intermediate components is larger in the diamond motif, $N_{u\to x} + N_{w\to x} > N_{v\to x}$. It can be shown that under the constraints that a) the production of v equals the total production of u and w and b) the production of x is the same in both motifs, the extrinsic noise is minimized when the coupling of u and w to x are identical. We stress, however, that the gain-to-noise ratio of this diamond motif is higher than that of a simple two-step cascade, with one intermediate component V that does not activate X in a cooperative manner. Indeed, for a diamond motif the coherent interaction between the two pathways plays a crucial role. While for a motif with non-cooperative activation of the output by the two pathways, this coherence between the two pathways exactly compensates for the decrease of the gain of each independent pathway, for a network with cooperative interaction, the coherence term increases the *GNR* over that of a simple two step cascade. In other words, splitting the input signal into two and then recombining them with `AND` logic at the output does increase the gain-to-noise ratio. Finally, we study a motif in which only the heterodimer UW, and not the homodimers U_2 and W_2, can activate the output X. We thus consider the case that $\gamma_w = \gamma_u = 0$, and consider what happens if the degradation rates of U and W are allowed to be different. With unequal degradation rates, a bandpass filter for information is possible, if the coupling of one pathway to x is stronger than the coupling of the other pathway. We take $\mu_w > \mu_u$, such that the pathway S $\to$ W $\to$ X is capable of transmitting information on faster timescales than the pathway S $\to$ U $\to$ X. If $\beta \gg \alpha$, the input signal is relayed more strongly via the pathway containing W, and the other pathway acts as a noise source. Consequently, for small frequencies, $\omega < \mu_u$, signal transmission will be corrupted by noise originating at u, but for $\mu_u < \omega < \mu_w$ this noise is averaged out. We thus obtain a band-pass filter for information transmission (Fig. 3.4c, red solid). By actively changing the degradation rates μ_u and μ_w the cell can tune the frequency range of the band.

3.4 Discussion

Our analysis reveals that feed-forward motifs and diamond motifs are very rich information processing devices. More specifically, our study shows that both coherent and incoherent feed-forward motifs can either act as low- or high-pass filters for information. This behavior can be understood by noting that while at high frequencies the direct pathway always dominates the output signal, at low frequencies the contribution of each pathway to the output varies between networks, depending on the coupling constants between the components in the network; moreover, at low frequencies, the output strongly depends on the nature of the

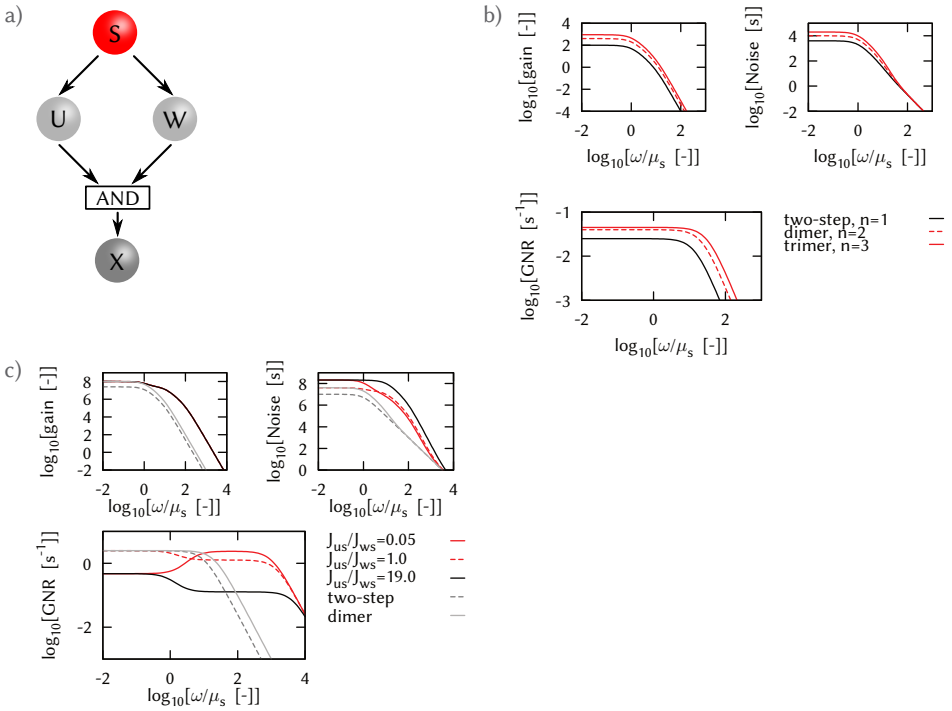


Figure 3.4: The diamond motif. **a)** The diamond motif. The diamond motif combines two pathways (S→V→X and S→W→X), originating from the same source at the response X. With homodimerization U and W correspond to two molecules of the species V. **b)** The difference between the gain, noise and *GNR* for a network with different levels of cooperativity (Eq. 3.24). For increasing cooperativity (n increases), the gain and the noise both increase. The increase in the noise is smaller than in the gain, since N_X is unaffected by the cooperative interactions and therefore the *GNR* increases. Next, also the knee frequency, and thus the bandwidth, increases, since the *GNR* scales with ω^{-2} for $\omega^{-2}>n^2\mu_v$. Parameters: k_s=100,k_v=10,k_x=10,μ_v=100,μ_x=1, μ_s sets the timescale. **c)** The *GNR* for the diamond motif where U and W have different degradation rates $\mu_u<\mu_w$. If $J_{us}>J_{ws}$ the *GNR* is low-pass (black solid), since on long timescales ($t>\mu_u^{-1}$, the signal is transmitted while on shorter timescales most transmission is corrupted by the intrinsic noise. If $J_{us}<J_{ws}$ (red solid line), the slow signal variations transmitted via w are corrupted by noise from u. For faster variations, the noise from u is averaged out and the signal can be transmitted with larger reliability. Note that the gain for all three parameter sets is equal, and the lines thus overlap. The bandpass characteristic for the *GNR* is thus due to the different dependence of the noise on ω. Parameters: k_s=100, μ_x=10, two-step cascade: μ_v=10,k_v^{ts}=500, k_x^{ts}=100, α to equalize production of v, γ_v to equalize production of x, diamond motif: μ_u=1,μ_w=100, γ_u=γ_w=0, γ_{uw} to equalize production of x, β to equalize $\langle v^{ts}\rangle=\langle u^{dm}\rangle+\langle w^{dm}\rangle$ and α=0.05k_v^{ts},0.5k_v^{ts},0.95k_v^{ts}, respectively red solid, red dashed, black solid. μ_s sets the timescale.

interaction between the two pathways. If the interaction is coherent, as in coherent feed-forward motifs, then the frequency dependence of the gain, noise and gain-to-noise ratio tends to have low-pass characteristics. If, however, not only at high frequencies, but also at low frequencies the direct pathway dominates the response, then the gain-to-noise ratio

can have high-pass characteristics; in this case, the signal is predominantly transmitted via the direct pathway, while the indirect pathway acts as a noise source, masking this signal at low frequencies. For an incoherent feed-forward motif, the gain is low at low frequencies because of the destructive interference between the two pathways; consequently, the frequency dependence of the gain-to-noise ratio of incoherent feed-forward motifs tends to have high-pass characteristics. However, an incoherent motif can also act as a low pass filter for information. This scenario arises when the direct pathway acts weakly on the output at low frequencies; then at low frequencies the output is dominated by the the indirect pathway, which exhibits low-pass signal filtering, while at high frequencies it is dominated by the direct pathway. More generally, our analysis demonstrates that by changing the coupling constants between the components the frequency-dependence of the gain, noise and gain-to-noise ratio can be sculpted in almost any desirable manner.

For equal total production cost of all molecules in the network a coherent feed-forward motif has a lower information transmission capacity than a simple one-step motif. However, if we allow for a higher production cost in the feed-forward motif (e.g. we require equal production of x, but allow for the additional production of the intermediate component V), interestingly a coherent feed-forward motif in which the two pathways are combined following `AND` logic can have a larger *GNR* than the one-step cascade. Combining the pathways following `OR` logic has always a smaller *GNR* than the one-step cascade, even for larger total production. This demonstrates for these simple cascades a possible advantage of coincidence detection.

Our results also underscore the important observation that the power spectrum of the output signal is not a good measure for information transmission [107]. The results on the incoherent feed-forward motif provide a concrete illustration of this idea: while the frequency dependence of the gain exhibits band-pass characteristics, the frequency dependence of the gain-to-noise ratio shows high-pass characteristics; indeed, at high frequencies not only the gain, and hence the output, decreases, but also the noise. The coherent feed-forward motif with a high-pass gain-to-noise ratio provides another striking example: while the gain and hence the output decreases with frequency, the gain-to-noise ratio increases; this is because the high gain at low frequencies is masked by the high noise. Our results thus show that in order to draw any conclusion on how reliably a network can transmit time-varying signals, one needs to measure not only the power spectrum of the output $P_{xx}(\omega)$, but also the power spectrum of the input $P_{ss}(\omega)$ and their cross-power spectrum $P_{sx}(\omega)$: from these quantities one can obtain the gain $g^2(\omega) \equiv |P_{sx}(\omega)|^2/P_{ss}^2(\omega)$ and the frequency dependence noise $N(\omega)$, and hence the gain-to-noise ratio (see Eq. 3.7 and Eq. 3.8).

Finally, our analysis of the diamond motifs reveals that cooperatively activating the output can markedly enhance the gain, as well as the gain-to-noise ratio. The latter is due to coincidence detection: While variations in the input signal lead to correlated variations in the intermediate components that tend to boost the output, noise generates uncorrelated fluctuations in the intermediate components, which couple less strongly to the output. We emphasize that this mechanism is very generic. Indeed, cooperative activation of the output can be implemented in many ways: via the formation of homo- or heteromultimers that then bind and activate the output component, or via the individual binding of the inter-

mediate components to the output component. While cooperative activation of the output via one and the same type of intermediate component, as in the case of homodimerization, increases the overall gain-to-noise ratio, cooperative output activation via components that are different, as in the case of heterodimerization of the intermediate components, makes it possible to mold the frequency dependence of the gain-to-noise ratio, even allowing for band-pass filters for information.

3.5 Acknowledgements

I would like to thank Philippe Nghe for a critical reading of this chapter.

3.A Supplementary Information

All cascades have the following simple (linear) birth-death process for the signal

$$\frac{ds}{dt} = k_s - m_s s + \eta_s(t). \tag{A3.1}$$

3.A.1 Simple cascades

Simple cascades, cascades without feed-forward interaction, are described using roman symbols for the kinetic rates, while motifs with feed-forward interactions, are described using Greek symbols. The network for the one-step simple cascade is described by

$$\frac{dx}{dt} = k_x s - m_x x + \eta_x(t). \tag{A3.2}$$

Gain, noise and gain-to-noise are

$$g^2(\omega) = \frac{k_x^2}{\omega^2 + m_x^2}, \tag{A3.3}$$

$$N(\omega) = \frac{\left\langle |\eta_x|^2 \right\rangle}{\omega^2 + m_x^2}, \tag{A3.4}$$

$$\frac{g^2(\omega)}{N(\omega)} = \frac{k_x^2}{\left\langle |\eta_x|^2 \right\rangle} = \frac{k_x}{2\langle s \rangle}. \tag{A3.5}$$

The network for the two-step simple cascade is described by

$$\frac{dv}{dt} = k_v s - m_v v + \eta_v(t), \tag{A3.6}$$

$$\frac{dx}{dt} = k_x v - m_x x + \eta_x(t). \tag{A3.7}$$

Gain, noise and gain-to-noise are

$$g^2(\omega) = \frac{(k_v k_x)^2}{(\omega^2 + m_x^2)(\omega^2 + m_v^2)}, \tag{A3.8}$$

$$N(\omega) = \frac{k_x^2 \left\langle |\eta_v|^2 \right\rangle}{(\omega^2 + m_v^2)(\omega^2 + m_x^2)} + \frac{\left\langle |\eta_x|^2 \right\rangle}{(\omega^2 + m_x^2)}, \tag{A3.9}$$

$$\frac{g^2(\omega)}{N(\omega)} = \frac{(k_v k_x)^2}{k_x^2 \left\langle |\eta_v|^2 \right\rangle + (\omega^2 + m_v^2) \left\langle |\eta_x|^2 \right\rangle}. \tag{A3.10}$$

Production constraints

In the following sections we will compare the one-step (os) and two-step (ts) cascades with different feed-forward motifs. We always assume that degradation rates for proteins are

equal, unless specified otherwise. We use three different constraints, such that the comparison is performed on an equal footing. These constraints are

1. Equal production of x, free production of v.

$$p_X^{\text{ts}} = p_X^{\text{ff}}, p_X^{\text{os}} = p_X^{\text{ff}}. \tag{A3.11}$$

   We note that this also implies $\langle|\eta_x^{\text{ts}}|^2\rangle = \langle|\eta_x^{\text{os}}|^2\rangle = \langle|\eta_x^{\text{ff}}|^2\rangle$

2. Equal total production of x and v,

$$p_X^{\text{ts}} + p_V^{\text{ts}} = p_X^{\text{ff}} + p_V^{\text{ff}}, p_X^{\text{os}} = p_X^{\text{ff}} + p_V^{\text{ff}}. \tag{A3.12}$$

3. Equal production of x and v separately (this constraint has no meaning for the one-step cascade)

$$p_X^{\text{ts}} = p_X^{\text{ff}}, p_V^{\text{ts}} = p_V^{\text{ff}}. \tag{A3.13}$$

   We note that this also implies $\langle|\eta_x^{\text{ts}}|^2\rangle = \langle|\eta_x^{\text{ff}}|^2\rangle$, $\langle|\eta_v^{\text{ts}}|^2\rangle = \langle|\eta_v^{\text{ff}}|^2\rangle$.

Unless specified otherwise the degradation rates are equal for components in either cascade, e.g $m_v = \mu_v$.

3.A.2 OR coherent feed-forward

The network is described by

$$\frac{dv}{dt} = \alpha s - \mu_v v + \eta_v(t), \tag{A3.14}$$

$$\frac{dx}{dt} = \beta s + \gamma v - \mu_x x + \eta_x(t), \tag{A3.15}$$

and the gain, noise and gain-to-noise the gain, noise and gain-to-noise (see Eqs. 3.9,3.12 and 3.14)

$$g^2(\omega) = \frac{(\alpha\gamma + J_{vv}\beta)^2 + J_{xs}^2\omega^2}{(\omega^2+\mu_v^2)(\omega^2+\mu_x^2)}, \tag{A3.16}$$

$$N(\omega) = \frac{\gamma^2\left\langle|\eta_v|^2\right\rangle + (\omega^2+\mu_v^2)\left\langle|\eta_x|^2\right\rangle}{(\omega^2+\mu_v^2)(\omega^2+\mu_x^2)}, \tag{A3.17}$$

$$\frac{g^2(\omega)}{N(\omega)} = \frac{(\alpha\gamma+\mu_v\beta)^2 + \beta^2\omega^2}{\gamma^2\left\langle|\eta_v|^2\right\rangle + (\omega^2+\mu_v^2)\left\langle|\eta_x|^2\right\rangle}. \tag{A3.18}$$

In Fig. 3.10 (left) the dependence of the gain and the phase difference between the direct S $\to$ X and indirect S $\to$ V $\to$ X pathway as a function of frequency is shown. For $\omega < \mu_v$ the two pathways have no phase difference, but for an increase in frequency the phase

Cons.	**one-step**	*GNR* **one-step**
E.p.x	$k_x = \beta + \dfrac{\alpha\gamma}{\mu_v}$	$\dfrac{g^2}{N} = \dfrac{(\mu_v\beta+\alpha\gamma)^2}{\mu_v^2\left\langle\lvert\eta_x\rvert^2\right\rangle} > GNR^{\text{ocff}}$
E.t.p.		Always larger
	two-step	*GNR* **two-step at** $\omega = 0$
E.p.x	$k_x k_v = \beta + \dfrac{\alpha\gamma}{\mu_v}$	$\dfrac{g^2}{N} = \dfrac{(\mu_v\beta+\alpha\gamma)^2}{\mu_v^2\left\langle\lvert\eta_x\rvert^2\right\rangle + \frac{2\langle s\rangle}{k_v}(\mu_v\beta+\alpha\gamma)^2}$
E.t.p.	$(k_x+\mu_v)k_v = (\alpha+\beta)\mu_v + \gamma\alpha$	$\dfrac{g^2}{N} = \dfrac{k_x(\beta\mu_v+\alpha(\gamma+\mu_v))}{(k_x+\mu_v)^2}$
E.s.p.	$k_v = \alpha,$ $k_x = \gamma + \dfrac{\beta\mu_v}{\alpha}$	$g^2 = \dfrac{(\mu_v\beta+\alpha\gamma)^2}{\mu_v^2\mu_x^2}$ $N = \dfrac{\frac{(\mu_v\beta+\alpha\gamma)^2}{\alpha^2}\left\langle\lvert\eta_v\rvert^2\right\rangle + \mu_v^2\left\langle\lvert\eta_x\rvert^2\right\rangle}{\mu_v^2\mu_x^2}$ $\dfrac{g^2}{N} = \dfrac{(\mu_v\beta+\alpha\gamma)^2}{\frac{(\mu_v\beta+\alpha\gamma)^2}{\alpha^2}\left\langle\lvert\eta_v\rvert^2\right\rangle + \mu_v^2\left\langle\lvert\eta_x\rvert^2\right\rangle}$

Table 3.1: The results for the gain, noise and gain-to-noise ratio for the one-step cascade and the two-step cascade in parameters of the ocff, given the constraint conditions. For the two-step only the $\omega\to 0$ limit is given, since for large ω the ocff always has a larger *GNR*. We assume always $\mu_v = m_v$. E.p.x: Equal production of x, e.t.p.:equal total production, e.s.p.:equal separate production.

difference increases and as a result the magnitude of the coherent term in the gain decreases. The magnitude of the coherent term also decreases due to the time-averaging over the finite lifetime of the intermediate component V. Even for a constant phase-difference, at large ω the gain decreases.

In Table. 3.1 and Table. 3.2 we list the expression for the *GNR* for different production constraints. The *GNR* of the one-step cascade given equal production of x is always larger than that of the ocff. Compared to the two-step cascade, for $\omega \to \infty$ the ocff motif always has a larger *GNR*, since the noise source $N_{v\to x}$ is averaged out, and only an effective one-step cascade with constant *GNR* remains. Thus we only compare the ocff with the two-step cascade for $\omega \to 0$.

The various constraints lead to different results. The first scenario, where we independently constrain p_v and p_x and take $\mu_v = m_v$ in both cascades leads to an expression for the *GNR* in the two-step cascade that is smaller than that for the ocff for all ω. The gain in both the ocff and the two-step cascade are precisely equal (not shown), and therefore the larger *GNR* is due to the increase in the noise ($N(\omega)$) in the two-step cascade. A constraint that allows for more freedom in the network is such that only the production of x is constrained,

Cons.	**one-step vs ocff**	**two-step vs ocff**
Equal production of x	$GNR^{\text{os}} > GNR^{\text{ocff}}$	Parameter dependent
Equal total production	$GNR^{\text{os}} > GNR^{\text{ocff}}$	Parameter dependent
Equal separate production		$GNR^{\text{ts}} < GNR^{\text{ocff}}$

Table 3.2: Summary of the results for the ocff.

but there is free production of v. We observe that for $\omega = 0$, the *GNR* for the ocff is larger than the two-step if

$$\gamma^2 \left\langle |\eta_v|^2 \right\rangle^{\text{ocff}} < \frac{2\langle s \rangle}{k_v} (\mu_v \beta + \alpha\gamma)^2 \tag{A3.19}$$

$$2\alpha \langle s \rangle \gamma^2 < \frac{2\langle s \rangle}{k_v} (\mu_v \beta + \alpha\gamma)^2, \tag{A3.20}$$

$$\alpha\gamma^2 < \frac{(\mu_v \beta + \alpha\gamma)^2}{k_v}. \tag{A3.21}$$

This inequality is not satisfied if e.g. $k_v \to \infty$, since then the two-step cascade effectively becomes an one-step cascade and as a result, has a larger *GNR* than the ocff.

For $\omega \to \infty$ the *GNR* of the ocff with high-pass filter characteristics can have a larger plateau value than an ocff with low-pass characteristics, under the constraint that the production of v and x separately are equal. This can be shown by solving for the inequality

$$\lim_{\omega\to\infty} GNR^{\text{hp}} > \lim_{\omega\to\infty} GNR^{\text{lp}} \to \left(\frac{\beta^{\text{hp}}}{\mu_v^{\text{hp}}}\right)^2 > \left(\frac{\beta^{\text{lp}}}{\mu_v^{\text{lp}}}\right)^2, \tag{A3.22}$$

$$\text{Low-pass (lp)} \to \beta^{\text{lp}} < \left(\beta^{\text{lp}} + \frac{\gamma^{\text{lp}}\alpha^{\text{lp}}}{\mu_v^{\text{lp}}}\right) \frac{\alpha^{\text{lp}}}{\beta^{\text{lp}}} + \frac{2\mu_v^{\text{lp}}}{\gamma^{\text{lp}}}, \tag{A3.23}$$

$$\text{High-pass (hp)} \to \beta^{\text{hp}} < \left(\beta^{\text{hp}} + \frac{\gamma^{\text{hp}}\alpha^{\text{hp}}}{\mu_v^{\text{hp}}}\right) \frac{\alpha^{\text{hp}}}{\beta^{\text{hp}}} + \frac{2\mu_v^{\text{hp}}}{\gamma^{\text{hp}}}. \tag{A3.24}$$

The first line describes the inequality. The second line describes the requirement for β^{lp} and the third line the requirement for β^{hp}, which are obtained following substitution of $\langle |\eta_v|^2 \rangle$

and $\langle|\eta_x|^2\rangle$ in Eq. 3.15. From the constraints of equal separate production we obtain

$$\left\langle\left|\eta_v^{\text{lp}}\right|^2\right\rangle = \left\langle\left|\eta_v^{\text{hp}}\right|^2\right\rangle \to \alpha^{\text{lp}} \equiv \alpha^{\text{hp}}, \tag{A3.25}$$

$$\left\langle\left|\eta_x^{\text{lp}}\right|^2\right\rangle = \left\langle\left|\eta_x^{\text{hp}}\right|^2\right\rangle \to \beta^{\text{lp}} + \frac{\gamma^{\text{lp}}\alpha^{\text{lp}}}{\mu_v^{\text{lp}}} \equiv \beta^{\text{hp}} + \frac{\gamma^{\text{hp}}\alpha^{\text{hp}}}{\mu_v^{\text{hp}}}. \tag{A3.26}$$

Solving the system of inequalities Eqs. A3.22-A3.24, with Eqs. A3.25,A3.26 using Mathematica, it can be shown that these can always be fulfilled. However, the full expressions are unwieldy to present here. Even if $\beta^{\text{lp}} \to \infty$, reflecting that the low-pass filter effectively is a one-step cascade, parameters can be found for which the high-pass filter can still have a larger *GNR* for large ω. However, the difference between the low-pass and high-pass filter is negligible.

3.A.3 AND coherent feed-forward

The network is described by

$$\frac{dv}{dt} = \alpha s - \mu_v v + \eta_v(t), \tag{A3.27}$$

$$\frac{dx}{dt} = \beta v s - \mu_x x + \eta_x(t). \tag{A3.28}$$

and the gain, noise and gain-to-noise (Eq. 3.9,3.12 and 3.14)

$$g^2(\omega) = \frac{\beta^2\langle v\rangle^2\left(\omega^2+4\mu_v^2\right)}{\left(\omega^2+\mu_v^2\right)\left(\omega^2+\mu_x^2\right)}, \tag{A3.29}$$

$$N(\omega) = \frac{\beta^2\langle s\rangle^2\left\langle|\eta_v|^2\right\rangle + \left(\omega^2+\mu_v^2\right)\left\langle|\eta_x|^2\right\rangle}{\left(\omega^2+\mu_v^2\right)\left(\omega^2+\mu_x^2\right)}, \tag{A3.30}$$

$$GNR(\omega) = \frac{\beta^2\langle v\rangle^2\left(\omega^2+4\mu_v^2\right)}{\beta^2\langle s\rangle^2\left\langle|\eta_v|^2\right\rangle + \left(\omega^2+\mu_v^2\right)\left\langle|\eta_x|^2\right\rangle}. \tag{A3.31}$$

For equal production of x and for $\omega \to \infty$, the *GNR* of a one-step cascade is equal to the *GNR* of the acff. For $\omega = 0$, we obtain the following relation

$$GNR^{\text{os}}(\omega=0) < GNR^{\text{acff}}(\omega=0) \tag{A3.32}$$

$$\frac{\beta^2\langle v\rangle^2}{\left\langle|\eta_x|^2\right\rangle} < 4\frac{\beta^2\langle v\rangle^2}{\left\langle|\eta_x|^2\right\rangle + \left(\frac{\beta\langle s\rangle}{\mu_v}\right)^2\left\langle|\eta_v|^2\right\rangle}, \tag{A3.33}$$

where the last line holds for $\beta\langle v\rangle/\alpha < 3\langle|\eta_x|^2\rangle/\langle|\eta_v|^2\rangle$. This inequality can be sufficed for example in the limit that $\beta \to 0$. Then the acff has a larger *GNR* than the one-step cascade for $\omega = 0$. This limit corresponds to a situation where the noise $N_{v\to x}$ (Eq. 3.12)

Cons.	**one-step**	***GNR* one-step**
E.p.x	$k_x = \beta\langle v\rangle$	$\frac{g^2}{N} = \frac{\beta^2\langle v\rangle^2}{\left\langle \lvert\eta_x\rvert^2\right\rangle}$
E.t.p.	$k_x = \alpha + \beta\langle v\rangle$	$\frac{g^2}{N} = \frac{(\alpha+\beta\langle v\rangle)^2}{\left\langle \lvert\eta_x^{\text{os}}\rvert^2\right\rangle} > GNR^{\text{acff}}$
	two-step	***GNR* two-step at $\omega = 0$**
E.p.x	$k_x k_v = \beta\langle v^{\text{acff}}\rangle\mu_v$	$\frac{g^2}{N} = \frac{(\langle s\rangle\alpha\beta)^2}{\left\langle \lvert\eta_x\rvert^2\right\rangle \mu_v^2 + \frac{(\langle s\rangle\alpha\beta)^2}{k_v^2}\left\langle \lvert\eta_v\rvert^2\right\rangle}$
E.t.p.	$(k_x+\mu_v)\,k_v = \mu_v\left(\alpha + \beta\langle v^{\text{acff}}\rangle\right)$	$\frac{g^2}{N} = \frac{k_x\alpha\,(\beta\langle s\rangle + \mu_v)}{2\langle s\rangle\,(k_x+\mu_v)^2}$
E.s.p.	$k_v = \alpha,$ $k_x = \frac{\mu_v}{k_v}\beta\langle v\rangle$	$g^2 = \frac{(\langle s\rangle\alpha\beta)^2}{\mu_v^2\mu_x^2}$ $N = \frac{\left\langle \lvert\eta_x\rvert^2\right\rangle \mu_v^2 + (\langle s\rangle\beta)^2\left\langle \lvert\eta_v\rvert^2\right\rangle}{\mu_v^2\mu_x^2}$ $\frac{g^2}{N} = \frac{(\langle s\rangle\alpha\beta)^2}{\left\langle \lvert\eta_x\rvert^2\right\rangle \mu_v^2 + (\langle s\rangle\beta)^2\left\langle \lvert\eta_v\rvert^2\right\rangle}$

Table 3.3: The results for the gain, noise and gain-to-noise ratio for the one-step cascade and the two-step cascade in parameters of the acff, given the constraint conditions. For the two-step only the $\omega\to 0$ limit is given, since for large ω the ocff always has a larger *GNR*. E.p.x: Equal production of x, e.t.p.:equal total production, e.s.p.:equal seperate production.

Cons.	**one-step vs acff**	**two-step vs acff**
Equal production of x	Parameter dependent	Parameter dependent
Equal total production	$GNR^{\text{os}} > GNR^{\text{acff}}$	Parameter dependent
Equal separate production		$GNR^{\text{ts}} < GNR^{\text{acff}}$

Table 3.4: Summary of the results for the acff.

is negligible, and the noise contributions $N(\omega)$ in both motifs are similar. The gain in the acff is larger than in the one-step cascade. This seems contradictory, but is a result of the production constraint. In the acff the production of x, p_x, is $\beta \langle s \rangle \langle v \rangle$, such that if $\langle v \rangle$ is very small, the production rate β becomes very to ensure equal production of x. However in the ocff, the production of x is $\beta \langle s \rangle + \gamma \langle v \rangle$ and if $\langle v \rangle$ is nearly zero, production still is possible through the direct pathway $\mathsf{S} \rightarrow \mathsf{X}$. Due to this coincidence coupling between the two pathways, the gain for the acff is larger than for the one-step cascade.

As expected, the acff has a larger *GNR* than the simple two-step cascade if we equalize both p_v and p_x separately. More interesting is the behavior at low frequencies for the two other constraint types. If we equalize the total production the two-step cascade has a larger *GNR* than the acff if

$$\frac{\mu_v \left(\beta \langle v \rangle^{\text{acff}} + \alpha - k_v\right)}{2 \langle s \rangle (\mu_v + k_x)} > \frac{\beta^2 \langle v \rangle^2 \left(\omega^2 + 4\mu_v^2\right)}{\beta^2 \langle s \rangle^2 \left\langle |\eta_v^{\texttt{AND}}|^2 \right\rangle + (\omega^2 + \mu_v^2) \left\langle |\eta_x^{\texttt{AND}}|^2 \right\rangle} \tag{A3.34}$$

$$\frac{\beta \langle s \rangle + \mu_v}{k_x + \mu_v} \frac{\beta \left\langle v^{\text{acff}} \right\rangle + \alpha - k_v}{k \langle v^{\texttt{AND}} \rangle} > 4, \tag{A3.35}$$

where this is possible if $\beta\langle s \rangle \gg k_x$ or $\alpha \gg k_v$. From the constraint condition of total production we have the equality $(k_x + \mu_v) k_v = \mu_v \alpha + \beta \alpha \langle s \rangle$. Taken these relations together, we observe that for $\beta\langle s \rangle \gg k_x$ we require $k_v \gg 1$, while for $\alpha \gg k_v$ we require to $k_x \gg 1$. In both cases the two-step cascade transforms into a one-step cascade because one of the two steps directly tracks the changes upstream.

The other option, where we constrain the production of x, but not v,

$$k_x = \beta \langle s \rangle \frac{\left\langle v^{\text{acff}} \right\rangle}{\langle v^{\text{two}} \rangle}, \tag{A3.36}$$

leads to the following expression for the gain-to-noise ratio at low frequencies for the two-step cascade

$$\mathit{GNR}(\omega = 0) = \frac{\left(\mu_v \beta \left\langle v^{\text{acff}} \right\rangle\right)^2}{\beta \langle s \rangle \frac{\left\langle v^{\text{acff}} \right\rangle}{\langle v^{\text{ts}} \rangle} \left\langle |\eta_v^{\text{ts}}|^2 \right\rangle + \mu_v^2 \left\langle |\eta_x|^2 \right\rangle}. \tag{A3.37}$$

The ratio of the the *GNR*'s is

$$\frac{GNR^{\text{acff}}(\omega=0)}{GNR^{\text{ts}}(\omega=0)} = 4\frac{\left(\beta\langle s\rangle\left\langle v^{\text{acff}}\right\rangle\right)^2\left\langle\left|\eta_v^{\text{ts}}\right|^2\right\rangle+\mu_v^2\langle v^{\text{ts}}\rangle^2\left\langle|\eta_x|^2\right\rangle}{(\beta\langle s\rangle\langle v^{\text{ts}}\rangle)^2\left\langle\left|\eta_v^{\text{acff}}\right|^2\right\rangle+\mu_v^2\langle v^{\text{ts}}\rangle^2\left\langle|\eta_x|^2\right\rangle} \tag{A3.38}$$

$$= 4\frac{\frac{\alpha^2\langle s\rangle^2 k_v}{\mu_v^2}+\frac{\mu_v^2\langle v^{\text{ts}}\rangle^2\langle|\eta_x|^2\rangle}{2\langle s\rangle(\beta\langle s\rangle)^2}}{\frac{\alpha\langle s\rangle^2 k_v^2}{\mu_v^2}+\frac{\mu_v^2\langle v^{\text{ts}}\rangle^2\langle|\eta_x|^2\rangle}{2\langle s\rangle(\beta\langle s\rangle)^2}} \tag{A3.39}$$

$$= 4\frac{\alpha+F}{k_v+F},\ \text{where } F=\frac{k_v\mu_v^2\left\langle|\eta_x|^2\right\rangle}{2\alpha\beta^2\langle s\rangle}. \tag{A3.40}$$

If $k_v \gg \alpha$ the two-step cascade has a larger *GNR*. This reflects again a situation where the two-step cascade has one very fast step and acts effectively as a single one-step cascade.

3.A.4 Comparison of the coherent feed-forward AND and OR motifs

We study if the *GNR* of the acff is larger or smaller than the *GNR* of the ocff. We equalize both production of v and x, such that we have

$$p_v^{\text{ocff}} = p_v^{\text{acff}} \rightarrow \alpha^{\text{ocff}} = \alpha^{\text{acff}} = \alpha, \tag{A3.41}$$

$$p_x^{\text{ocff}} = p_x^{\text{acff}} \rightarrow \beta^{\text{ocff}} + \frac{\alpha\gamma}{\mu_v} = \beta^{\text{acff}}\langle v\rangle, \tag{A3.42}$$

and using these constraints and Eq. A3.18,A3.31, we obtain the ratio

$$\frac{GNR^{\text{acff}}}{GNR^{\text{ocff}}} = \frac{\left(\beta^{\text{ocff}}+\frac{\alpha\gamma}{\mu_v}\right)^2(\omega^2+4\mu_v^2)}{\left(\beta^{\text{ocff}}\right)^2(\omega^2+\mu_v^2)+(\alpha\gamma)^2+2\alpha\beta\gamma\mu_v}\times \tag{A3.43}$$

$$\frac{\gamma^2\left\langle|\eta_v|^2\right\rangle+(\omega^2+\mu_v^2)\left\langle|\eta_x|^2\right\rangle}{\left(\frac{\beta^{\text{ocff}}\mu_v}{\alpha}+\gamma\right)^2\left\langle|\eta_v|^2\right\rangle+(\omega^2+\mu_v^2)\left\langle|\eta_x|^2\right\rangle}. \tag{A3.44}$$

In the limit $\omega\rightarrow\infty$ we obtain

$$\frac{GNR^{\text{acff}}}{GNR^{\text{ocff}}} = \frac{\left(\beta^{\text{ocff}}+\frac{\alpha\gamma}{\mu_v}\right)^2}{\left(\beta^{\text{ocff}}\right)^2} > 1, \tag{A3.45}$$

Cons.	$\omega \to 0$	$\omega \to \infty$
Equal separate production	Parameter dependent	$GNR^{\text{acff}} > GNR^{\text{ocff}}$

Table 3.5: Summary of the results for the comparison between the acff and the ocff.

such that for large ω the acff has a larger *GNR*. In the other limit $\omega = 0$, we obtain

$$\frac{GNR^{\text{acff}}}{GNR^{\text{ocff}}} = \frac{\left(\beta^{\text{ocff}} + \frac{\alpha\gamma}{\mu_v}\right)^2 4\mu_v^2}{\left(\beta^{\text{ocff}}\right)^2 \mu_v^2 + (\alpha\gamma)^2 + 2\alpha\beta\gamma\mu_v} \frac{\gamma^2 \left\langle |\eta_v|^2 \right\rangle + \mu_v^2 \left\langle |\eta_x|^2 \right\rangle}{\left(\frac{\beta^{\text{ocff}}\mu_v}{\alpha} + \gamma\right)^2 \left\langle |\eta_v|^2 \right\rangle + \mu_v^2 \left\langle |\eta_x|^2 \right\rangle} \tag{A3.46}$$

$$= 4 \frac{\gamma^2 \left\langle |\eta_v|^2 \right\rangle + \mu_v^2 \left\langle |\eta_x|^2 \right\rangle}{\left(\frac{\beta^{\text{ocff}}\mu_v}{\alpha} + \gamma\right)^2 \left\langle |\eta_v|^2 \right\rangle + \mu_v^2 \left\langle |\eta_x|^2 \right\rangle}, \tag{A3.47}$$

such that if

$$4\left(\gamma^2 \left\langle |\eta_v|^2 \right\rangle + \mu_v^2 \left\langle |\eta_x|^2 \right\rangle\right) < \left(\frac{\beta^{\text{ocff}}\mu_v}{\alpha} + \gamma\right)^2 \left\langle |\eta_v|^2 \right\rangle + \mu_v^2 \left\langle |\eta_x|^2 \right\rangle, \tag{A3.48}$$

the ocff has a larger GNR. This condition (Eq. A3.48) can be satisfied if in the ocff motif the direct pathway S $\to$ X couples (large β) much stronger to X than the indirect pathway S $\to$ V $\to$ X (small α). Interestingly, the parameter values for which this condition is satisfied, result in an *GNR* for the acff motif with a high-pass filter. This we show by comparing the two extrema for the acff ($\omega \to \infty$ and $\omega = 0$).

$$GNR^{\texttt{AND}}(\omega = 0) = \frac{4\mu_v^2 \left(\beta^{\text{ocff}} + \frac{\alpha\gamma}{\mu_v}\right)^2}{\left(\frac{\beta^{\text{ocff}}\mu_v}{\alpha} + \gamma\right)^2 \left\langle |\eta_v|^2 \right\rangle + \mu_v^2 \left\langle |\eta_x|^2 \right\rangle}, \tag{A3.49}$$

$$\lim_{\omega\to\infty} GNR^{\texttt{AND}}(\omega) = \frac{\left(\beta^{\text{ocff}} + \frac{\alpha\gamma}{\mu_v}\right)^2}{\left\langle |\eta_x|^2 \right\rangle}, \tag{A3.50}$$

where it can be shown that, given the condition in Eq. A3.48 is satisfied, the acff has a larger *GNR* for large frequencies than for small frequencies.

3.A.5 Incoherent feed-forward motif

The network is described by

$$\frac{dv}{dt} = \alpha s - \mu_v v + \eta_v(t), \tag{A3.51}$$

$$\frac{dx}{dt} = \frac{\nu K v}{K+s} - \mu_x x + \eta_x(t). \tag{A3.52}$$

The gain, noise and gain-to-noise (Eqs. 3.9,3.12 and 3.14) in general terms are

$$g^2(\omega) = \frac{(J_{vs}J_{xv} + J_{vv}J_{xs})^2 + J_{xs}^2\omega^2}{(\omega^2+\mu_v^2)(\omega^2+\mu_x^2)}, \tag{A3.53}$$

$$N(\omega) = \frac{J_{xv}^2\left\langle|\eta_v|^2\right\rangle + (\omega^2+\mu_v^2)\left\langle|\eta_x|^2\right\rangle}{(\omega^2+\mu_v^2)(\omega^2+\mu_x^2)}, \tag{A3.54}$$

$$\frac{g^2(\omega)}{N(\omega)} = \frac{(J_{vs}J_{xv} + J_{vv}J_{xs})^2 + J_{xs}^2\omega^2}{J_{xv}^2\left\langle|\eta_v|^2\right\rangle + (\omega^2+\mu_v^2)\left\langle|\eta_x|^2\right\rangle}. \tag{A3.55}$$

First, we study the influence of the topology. For the incoherent feed-forward, four different topologies exist with respect to the regulation. It is known that different topologies have influence on the noise characteristics. Here we show that the *GNR* indeed depends on the specific topology, comparing both a different position for the negative regulation, and the type of integration of the two pathways at the X component.

Instead of negatively regulating X by V or S, we can also negatively regulate V by S, such that we have

$$\frac{dv}{dt} = \frac{\alpha K_V}{K_V + x} - \mu_v v + \eta_v(t), \tag{A3.56}$$

$$\frac{dx}{dt} = \beta s + \gamma v - \mu_x x + \eta_x(t), \tag{A3.57}$$

where we have assumed the signals combine at X following OR strategy. In the previous section, we equalized production and studied the difference in the coupling parameters J_{ij}. Here, instead of equalizing production, we equalize the Jacobian coefficients (as a result the gain is equal). For these coupling parameters, we compare the production of the components (and thus the noise terms). We refer to the topology with superscript ^{V}, for negative

regulation of V, and superscript ^{X} for negative regulation of X.

$$\left|J_{vs}^V\right| = J_{vs}^X \to \alpha^V \frac{K_V}{\left(K_V + \langle s\rangle\right)^2} = \alpha^X, \tag{A3.58}$$

$$J_{xs}^V = J_{xs}^X \to \beta^V = \nu^X \frac{K\langle v\rangle}{\left(K + \langle s\rangle\right)^2}, \tag{A3.59}$$

$$J_{xv}^V = \left|J_{xv}^X\right| \to \gamma^V = \nu^X \frac{K}{K + \langle s\rangle}. \tag{A3.60}$$

And following these equalities, we write for the noise terms

$$\eta_V^V = 2\alpha^V \frac{K_V}{K_V + S} = 2\left(K_V + \langle s\rangle\right)\alpha^X > \eta_V^X, \tag{A3.61}$$

$$\eta_X^V = 2\left(\beta\langle s\rangle + \gamma\langle v\rangle\right) = 2\nu^X \frac{K\langle v\rangle}{K + \langle s\rangle}\left(1 + \frac{\langle s\rangle}{K + \langle s\rangle}\right) > \eta_X^X. \tag{A3.62}$$

For equal coupling constants the noise terms (or equivalently the production terms) in the iff motif are larger for negative regulation on V than on X.

If the pathways combine following `AND` logic we have

$$\frac{dx}{dt} = \gamma v s - \mu_x x + \eta_x(t). \tag{A3.63}$$

Now, we equalize production terms

$$p_V^V = \alpha^V \frac{K_V}{K_V + \langle s\rangle} \quad p_V^X = \alpha^X\langle s\rangle, \tag{A3.64}$$

$$p_X^V = \gamma\langle v\rangle\langle s\rangle \quad p_X^X = \nu \frac{K\langle v\rangle}{K + \langle s\rangle}. \tag{A3.65}$$

and obtain $\alpha^X = \alpha^V K_V\left[\left(K_V + \langle s\rangle\right)\langle s\rangle\right]^{-1}$ and $\gamma = \nu K\left[\left(K + \langle s\rangle\right)\langle s\rangle\right]^{-1}$. We then compare the Jacobian terms

$$\left|J_{vs}^V\right| = \alpha^V \frac{K_V}{\left(K_V + \langle s\rangle\right)^2}; \quad J_{vs}^X = \alpha^X = \alpha^V \frac{K_V}{\left(K_V + \langle s\rangle\right)\langle s\rangle}, \tag{A3.66}$$

$$J_{xs}^V = \gamma\langle v\rangle = \nu \frac{K\langle v\rangle}{\left(\langle s\rangle + K\right)\langle s\rangle}; \quad J_{xs}^X = \nu \frac{K\langle v\rangle}{\left(\langle s\rangle + K\right)^2}, \tag{A3.67}$$

$$J_{xv}^V = \gamma\langle s\rangle = \nu \frac{K}{\left(\langle s\rangle + K\right)}; \quad J_{xv}^X = \nu \frac{K}{\left(\langle s\rangle + K\right)}, \tag{A3.68}$$

and we observe that $J_{vs}^X > \left|J_{vs}^V\right|$, while $J_{xs}^V > J_{xs}^X$. Therefore we conclude that at large frequencies the `AND` integration with negative regulation on V has a larger *GNR* than for negative regulation on X. For small frequencies the *GNR* depends on the specific parameters.

Compared with a one-step cascade for equal production of x (so v is not constrained),

we have for the *GNR* of the incoherent motif

$$p_x^I = p_x^{\text{os}} \to \frac{\nu K \langle v \rangle}{K + \langle s \rangle} = k_x^{\text{os}} \langle s \rangle \to \frac{\nu K}{K + \langle s \rangle} = \frac{k_x^{\text{os}} \mu_v}{\alpha}, \tag{A3.69}$$

leading to

$$\frac{g^2}{N} = \frac{\left(\frac{K k_x^{\text{os}} \mu_v}{K + \langle s \rangle}\right)^2 + \left(\frac{k_x^{\text{os}} \langle s \rangle}{K + \langle s \rangle}\right)^2 \omega^2}{\left(\frac{k_x^{\text{os}} \mu_v}{\alpha}\right)^2 \left\langle |\eta_v|^2 \right\rangle + (\omega^2 + \mu_v^2) \left\langle |\eta_x|^2 \right\rangle} < \frac{k_x^{\text{os}}}{\left\langle |\eta_x|^2 \right\rangle}, \tag{A3.70}$$

such that the incoherent motif always has a smaller *GNR* than the one-step motif for equal production of x.

Compared to the two-step motif and equal total production we distinguish two scenarios. First, we assume equal production of v and obtain

$$p_x^{\text{iff}} = p_x^{\text{two}} \to \frac{\nu K}{K + \langle s \rangle} = \frac{k_x k_v}{\alpha}, \tag{A3.71}$$

and for the *GNR* of the iff cascade

$$\frac{g^2}{N} = \frac{(k_x k_v)^2 \left[\left(\frac{K}{K + \langle s \rangle}\right)^2 + \left(\frac{\langle s \rangle}{(K + \langle s \rangle)\mu_v}\right)^2 \omega^2\right]}{(k_x)^2 \left\langle |\eta_v|^2 \right\rangle + (\omega^2 + \mu_v^2) \left\langle |\eta_x|^2 \right\rangle}. \tag{A3.72}$$

Compared to the *GNR* of a two-step cascade, for small ω the *GNR* of the iff is smaller due to the negative interference in the iff, while for large ω this is larger than the *GNR* of a two-step cascade due to the direct pathway.

In the other scenario only require $p_x^{\text{iff}} = p_x^{\text{two}}$, to obtain

$$\frac{g^2}{N} = \frac{(k_x k_v)^2 \left[\left(\frac{K}{K + \langle s \rangle}\right)^2 + \left(\frac{\langle s \rangle}{(K + \langle s \rangle)\mu_v}\right)^2 \omega^2\right]}{2 \left(\frac{k_x k_v}{\alpha}\right)^2 \alpha \langle s \rangle + (\omega^2 + \mu_v^2) \left\langle |\eta_x|^2 \right\rangle}. \tag{A3.73}$$

Due to the direct pathway the gain-to-noise ratio for the iff is larger for $\omega \to \infty$. For $\omega = 0$

Cons.	$\omega \to 0$	$\omega \to \infty$
Equal total production	$GNR^{\text{iff}} < GNR^{\text{ts}}$	$GNR^{\text{iff}} > GNR^{\text{ts}}$
Equal production of x	Parameter dependent	$GNR^{\text{iff}} > GNR^{\text{ts}}$

Table 3.6: Summary of the results for the comparison between the iff and the two-step cascade.

we obtain

$$\overbrace{\frac{g^2(\omega=0)}{N(\omega=0)}}^{\text{iff}} > \overbrace{\frac{g^2(\omega=0)}{N(\omega=0)}}^{\text{simple 2 step}} \tag{A3.74}$$

$$\frac{(k_x k_v)^2 \left(\frac{K}{K+\langle s\rangle}\right)^2}{2\left(\frac{k_x k_v}{\alpha}\right)^2 \alpha \langle s\rangle + \mu_v^2 \left\langle |\eta_x|^2 \right\rangle} > \frac{(k_v k_x)^2}{2(k_x)^2 k_v \langle s\rangle + m_v^2 \left\langle |\eta_x|^2 \right\rangle} \tag{A3.75}$$

$$\left(\frac{K}{K+\langle s\rangle}\right)^2 \frac{1}{2\left(\frac{1}{\alpha}\right)\langle s\rangle + \frac{\mu_v^2}{(k_x k_v)^2}\left\langle |\eta_x|^2 \right\rangle} > \frac{1}{2\frac{1}{k_v}\langle s\rangle + \frac{\mu_v^2}{(k_x k_v)^2}\left\langle |\eta_x|^2 \right\rangle}, \tag{A3.76}$$

where the above inequality is valid in the following two scenario's: 1) if $k_v < \alpha$, the iff motif has a larger *GNR*, since the signal is transmitted to v with larger gain than in the two-step cascade. 2) If $K \gg \langle s\rangle$, the negative feedback is greatly suppressed and we observe a general two-step cascade instead of an iff motif. If the first condition is not satisfied, the two-step simple cascade has a larger *GNR* for $\omega = 0$. The second condition is required if the difference in the gain due to the first condition is not large enough.

3.A.6 Multimerization

The network for multimerization is described by

$$\frac{dv}{dt} = \alpha s - \mu_v v + \eta_v(t), \tag{A3.77}$$

$$\frac{dx}{dt} = \gamma_v v^n - \mu_x x + \eta_x(t). \tag{A3.78}$$

where n is the number of proteins that jointly activate X. The gain, noise and gain-to-

noise

$$g^2(\omega) = \frac{n^2 {k_s}^{2(n-1)} \alpha^{2n} \gamma_v^2}{{\mu_s}^{2(n-1)} \mu_v^{2(n-1)} (\omega^2 + \mu_v^2)(\omega^2 + \mu_x^2)}, \tag{A3.79}$$

$$N(\omega) = \frac{n^2 {k_s}^{2(n-1)} \alpha^{2(n-1)} \gamma_v^2 \left\langle |\eta_v|^2 \right\rangle + {\mu_s}^{2(n-1)} \mu_v^{2(n-1)} \left(\omega^2 + \mu_v^2\right) \left\langle |\eta_x|^2 \right\rangle}{{\mu_s}^{2(n-1)} \mu_v^{2(n-1)} (\omega^2 + \mu_v^2)(\omega^2 + \mu_x^2)}, \tag{A3.80}$$

$$GNR(\omega) = \frac{n^2 {k_s}^{2(n-1)} \alpha^{2n} \gamma_v^2}{n^2 {k_s}^{2(n-1)} \alpha^{2(n-1)} \gamma_v^2 \left\langle |\eta_v|^2 \right\rangle + {\mu_s}^{2(n-1)} \mu_v^{2(n-1)} (\omega^2 + \mu_v^2) \left\langle |\eta_x|^2 \right\rangle}. \tag{A3.81}$$

For equal production of x (v is by construction in all cascades equal), we have

$$\gamma_{v,n} = \gamma_{n=1} \frac{\mu_s^{n-1} \mu_v^{n-1}}{k_s^{n-1} \alpha^{n-1}}, \tag{A3.82}$$

which after we substitute this in Eq. A3.81, results in Eq. 3.30.

The diamond motif is described by

$$\begin{aligned} \frac{du}{dt} &= \alpha s(t) - \mu_u u(t) + \eta_u(t), \\ \frac{dw}{dt} &= \beta s(t) - \mu_w w(t) + \eta_w(t), \\ \frac{dx}{dt} &= \gamma_u u^2(t) + \gamma_w w^2(t) + 2\gamma_{uw} u(t) w(t) - \mu_x x(t) + \eta_x(t). \end{aligned} \tag{A3.83}$$

For $\mu_w = \mu_u = \mu_v$, $\alpha = C_I \beta$ and $\gamma_u = \gamma_{uw} = C_{II} \gamma_w$, where we have introduced the coefficients C_I and C_{II} to study form of the gain, noise and *GNR* for differences between the intermediates U and W in a general context, the expression for the gain, noise and gain-to-noise are

$$g^2(\omega) = \frac{4 k_s^2 \beta^4 \left(1 + 2C_I + C_I^2 C_{II}\right)^2 \gamma_u^2}{\mu_s^2 \mu_w^2 H(\mu_w, \omega) H(\mu_x, \omega)}, \tag{A3.84}$$

$$N(\omega) = \frac{4 k_s^2 \beta^2 \gamma_u^2 \left(A^2 \left\langle |\eta_u|^2 \right\rangle + (1 + C_I)^2 \left\langle |\eta_w|^2 \right\rangle\right) + \mu_s^2 \mu_w^2 H(\mu_w, \omega) \left\langle |\eta_x|^2 \right\rangle}{\mu_s^2 \mu_w^2 H(\mu_w, \omega) H(\mu_x, \omega)}, \tag{A3.85}$$

$$GNR(\omega) = \frac{4 k_s^2 \beta^4 \left(1 + 2C_I + C_1^2 C_{II}\right)^2 \gamma_u^2}{4 k_s^2 \beta^2 \gamma_u^2 \left(A^2 \left\langle |\eta_u|^2 \right\rangle + (1 + C_I)^2 \left\langle |\eta_w|^2 \right\rangle\right) + \mu_s^2 \mu_w^2 H(\mu_w, \omega) \left\langle |\eta_x|^2 \right\rangle}, \tag{A3.86}$$

where $A = (1 + C_I C_{II})$ and $H(x, y) = x^2 + y^2$.

We take $C_{II} = 1$ and compare this cascade to a cascade with only homodimerization of the component U ($\beta = \gamma_w = \gamma_{uw} = 0$). For readability, we refer to this cascade, as if it has an intermediate component V, and if required, subscripts denote U, V and W.

Equal production at the intermediate level $p_V = p_U + p_W$ and at the level of X ($p_X^{\text{dm}} = p_X^{\text{homo}}$) gives

$$(1 + C_I)\,\beta = \alpha_v \tag{A3.87}$$

$$\begin{aligned}\gamma_{uw} &= \gamma_v \frac{\langle v\rangle^2}{U^2 + 2\,\langle u\rangle\,\langle w\rangle + \langle w\rangle^2} = \gamma_v \frac{(\langle u\rangle + \langle w\rangle)^2}{U^2 + 2\,\langle u\rangle\,\langle w\rangle + \langle W\rangle^2} \\ &= \gamma_v \frac{(\alpha_u + \beta_w)^2}{\alpha_u^2 + 2\alpha_u\beta_w + \beta_w^2} = \gamma_v.\end{aligned} \tag{A3.88}$$

Inserting these equalities in Eq. A3.84, we observe that the gain is equal to Eq. A3.79 for $n = 2$,

$$g^2(\omega) = \frac{4k_s^2\alpha_v^4\gamma_v^2}{{\mu_s}^2\mu_v^2\,(\omega^2 + \mu_v^2)\,(\omega^2 + \mu_x^2)}, \tag{A3.89}$$

$$= \frac{4k_s^2\,(1 + C_I)^4\,\beta_w^4 \left(\frac{\alpha_u^2 + 2\alpha_u\beta_w + \beta_w^2}{(\alpha_u + \beta_w)^2}\right)^2 \gamma_{uw}^2}{{\mu_s}^2\mu_v^2\,(\omega^2 + \mu_v^2)\,(\omega^2 + \mu_x^2)}, \tag{A3.90}$$

$$= \frac{4k_s^2\,(1 + C_I)^4\,\beta_w^4\gamma_{uw}^2}{{\mu_s}^2\mu_v^2\,(\omega^2 + \mu_v^2)\,(\omega^2 + \mu_x^2)}. \tag{A3.91}$$

and the noise is

$$N(\omega) = \frac{4k_s^2\,(1 + C_I)^2\,\beta_w^2\gamma_v^2\left(\left\langle|\eta_u|^2\right\rangle + \left\langle|\eta_w|^2\right\rangle\right) + {\mu_s}^2\mu_v^2\left(\omega^2 + \mu_v^2\right)\left\langle|\eta_x|^2\right\rangle}{{\mu_s}^2\mu_v^2\,(\omega^2 + \mu_v^2)\,(\omega^2 + \mu_x^2)}, \tag{A3.92}$$

where Eq. A3.92 is exactly Eq. A3.85. Therefore, unequal production of u and w is not important, as long as the sum is constrained (and all γ's are equal).

Next we compare to a two-step cascade. We first assume the diamond motif to be completely symmetric for the intermediate components U, W, $\alpha = \beta$, $\mu_u = \mu_w$ and $\gamma_u = \gamma_w = \gamma_{uw}$. We constrain the individual production of each step in the cascade

$$p_v^{\text{two}} = p_u^{\text{dm}} + p_w^{\text{dm}} \to k_v = 2\alpha, \tag{A3.93}$$

$$p_x^{\text{two}} = p_x^{\text{dm}} \to k_x k_v = 4\gamma\frac{\alpha^2\,\langle s\rangle}{\mu_v}. \tag{A3.94}$$

In the limit $\omega \to \infty$ the diamond motif has a larger *GNR* independent of any kinetic

rates, since

$$\frac{GNR^{\text{dm}}}{GNR^{\text{two}}} = \frac{4\mu_v^2}{\mu_v^2} = 4. \tag{A3.95}$$

In the opposite limit, $\omega \to 0$, we have

$$\frac{GNR^{\text{dm}}}{GNR^{\text{two}}} = \frac{(8\langle s\rangle \alpha\gamma_{uw} + 4\mu_w^2}{(8\langle s\rangle \alpha\gamma_{uw} + \langle s\rangle \mu_w^2)} > 1. \tag{A3.96}$$

Thus the symmetric diamond motif has a larger *GNR* for $\omega \to 0$ and $\omega \to \infty$. For equal $\mu_u = m_v$, the *GNR* of the diamond motif is larger for all frequencies.

The above case is for $\alpha = \beta$. Here we use $\alpha = C_I\beta$, but $\mu_v = \mu_w = m_v$. The ratio of the *GNR*'s for $\omega \to \infty$ is simply 4 (Eq. A3.95). For the limit $\omega \to 0$ we have

$$\frac{GNR^{\text{dm}}}{GNR^{\text{two}}} = \frac{4\left((1+C_I)\langle s\rangle \alpha\gamma_{uw} + \mu_w^2\right)}{4(1+C_I)\langle s\rangle \alpha\gamma_{uw} + \mu_w^2} > 1. \tag{A3.97}$$

We do observe that the more asymmetric the diamond motif becomes, the ratio of the *GNR*'s decreases. The performance of the diamond motif then becomes more similar to the two-step cascade.

More interestingly is to have different degradation rates $\mu_v = C_{III}\mu_w$, where C_{III} is an arbitrary constant. The equal production constraints then result in the following two expressions

$$p_v^{\text{two}} = p_v^{\text{dm}} + p_w^{\text{dm}} \to k_v = 2\alpha, p_x^{\text{two}} = p_x^{\text{dm}} \to k_x k_v = m_v\gamma\frac{\alpha^2\langle s\rangle}{C_{III}\mu_w^2}. \tag{A3.98}$$

For high frequencies the ratio of the *GNR* of the two-step and diamond motif is

$$\frac{GNR^{\text{dm}}}{GNR^{\text{two}}} = 16\frac{C_{III}^2\mu_w^2}{(1+C_{III})^2 m_v^2}, \tag{A3.99}$$

which does not lead to a unique conclusion whether which of the two is larger. For $\omega = 0$, we have

$$\frac{GNR^{\text{dm}}}{GNR^{\text{two}}} = \frac{2(1+C_{III})^2\langle s\rangle \alpha\gamma_{uw} + 2C_{III}^2\mu_w^2}{4\left(1+C_{III}^2\right)\langle s\rangle \alpha\gamma_{uw} + C_{III}^2\mu_w^2}. \tag{A3.100}$$

Again, the ratio of the *GNR* depends on the kinetic rates.

3.A.7 Numerical validation

The linearization used in the derivation can change the characteristics of the frequency response, since a linear(ized) system does not change the frequency of the transmitted signal. This may not be the case for a nonlinear system. In this section we show the compari-

son between our analytical results, following the linear noise approximation, for the power spectrum and the result from numerical simulations of the full non-linear network. For the numerical simulations we use the Gillespie algorithm. The negative regulation as present in the incoherent feed-forward motif, or the positive regulation as in the dimerization process are calculated through Hill-like interactions between the components. In the Gillespie simulation we calculated the propensities for every reaction using the coarse grained Hill-expressions for the propensities, such that

$$\mathsf{V} \xrightarrow{r} \mathsf{V} + \mathsf{X}, \tag{A3.101}$$

where r is

$$r = \frac{kK}{S + K}, \tag{A3.102}$$

where here the actual copy number of S is used, and not $\langle s \rangle$, as in the linearized expressions (Eq. A3.55).

The power spectra are calculated using 2^{10} (1024) exponentially distributed frequencies from $\omega = 10^{-1}$ to $\omega = 10^{2}$ and averaged over 16 neighboring frequencies to obtain a single data point. In total we have 64 data points. Fourier transforms and power spectra are directly integrated during runtime. We simulated a minimum of 27 blocks of 5000s. Results are shown in Figs. 3.5-3.9.

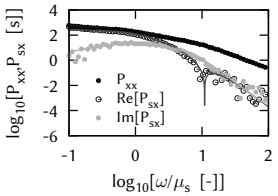


Figure 3.5: The results (symbols) of the Gillespie simulations for the OR-coherent feed-forward motif. Since this motif is linear, we do not expect any deviations between the numerical simulations and the analytical results. We only show the result for one set of parameters: k_S=100,α=11,β=1,γ=10.1,μ_V=10,μ_X=10, μ_S sets the timescale.

3.A.8 Influence of the phase

In Fig. 3.10 we show in some more detail the precise influence of the phase difference between two pathways that combine at a downstream component. A larger phase difference corresponds to a decrease in the gain. However, the phase difference between the pathways is not the only relevant parameter to study the gain. The gain is also greatly reduced if the signal variations are much faster than the lifetime of the components in the cascade. If this is the case, the individual components can not track the variations and start to time-average the variations, thereby losing the specific information in the high frequencies.

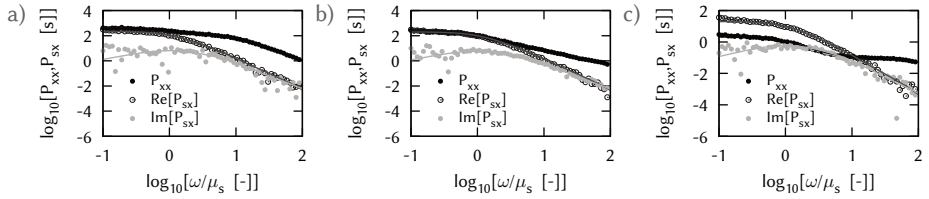


Figure 3.6: The results (symbols) of the Gillespie simulations for the AND-coherent feed-forward motif. Although this motif is non-linear, the results of the simulations are in good agreement with those of the linear analysis (continuous lines). In P_{xx} and $Re[P_{sx}]$ the symbols of the simulations results are on top of the lines in the analytical results. Parameters: **a)** as in Fig. 3.2c red solid, **b)** Fig. 3.2c red dashed, **c)** Fig. 3.2c black solid.

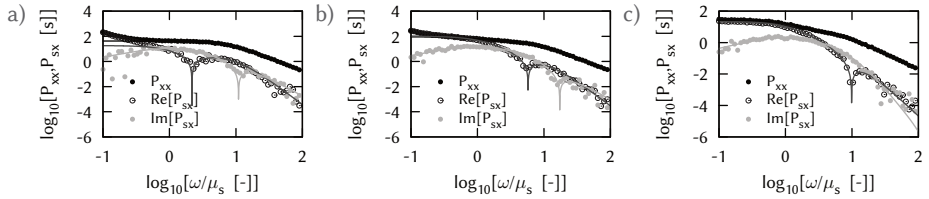


Figure 3.7: The results (symbols) of the Gillespie simulations for the incoherent feed-forward motif. Although this motif is non-linear, the results of the simulations are in good agreement with those of the linear analysis (continuous lines). In P_{xx} and $Re[P_{sx}]$ the symbols of the simulations results are on top of the lines in the analytical results. Parameters: **a)** as in Fig. 3.3c red solid, **b)** Fig. 3.3c red dashed, **c)** Fig. 3.3c black solid.

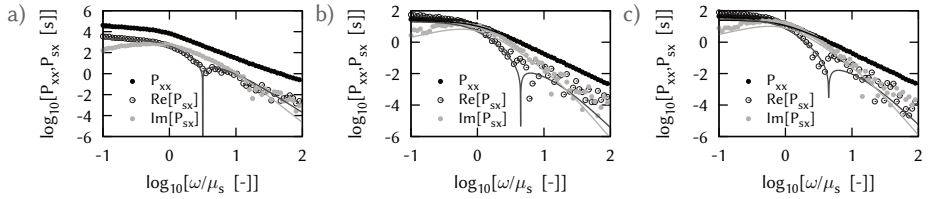


Figure 3.8: The results (symbols) of the Gillespie simulations for the dimerization process. Here the motif is non-linear. We model the propensity function following a Hill-function, $k(V)=k_{max}V^n/(V^n+K^n)$. Next, we take $K \gg V$, such that propensity is approximately $k(V) \approx k_{max}(V/K)^n$ as discussed in the main text. The results of the simulations are in good agreement with those of the linear analysis (continuous lines). In P_{xx} and $Re[P_{sx}]$ the symbols of the simulations results are on top of the lines in the analytical results. Parameters: **a)** two-step process (n=1), **b)** dimer process (n=2), **c)** trimer process (n=3). Parameters are as in Fig. 3.4b.

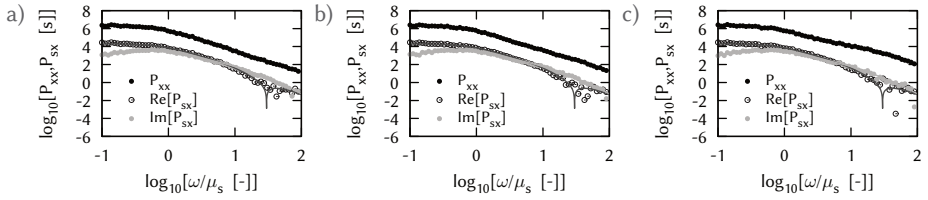


Figure 3.9: The results (symbols) of the Gillespie simulations for the heterodimerization motif, with $\gamma_u=\gamma_w=0$. Although this motif is non-linear, the results of the simulations are in good agreement with the results of the linear analysis (continuous lines).In P_{xx} and $Re[P_{sx}]$ the symbols of the simulations results are on top of the lines in the analytical results. Parameters: **a)** as in Fig. 3.4c red solid, **b)** Fig. 3.4c red dashed, **c)** Fig. 3.4c black solid.

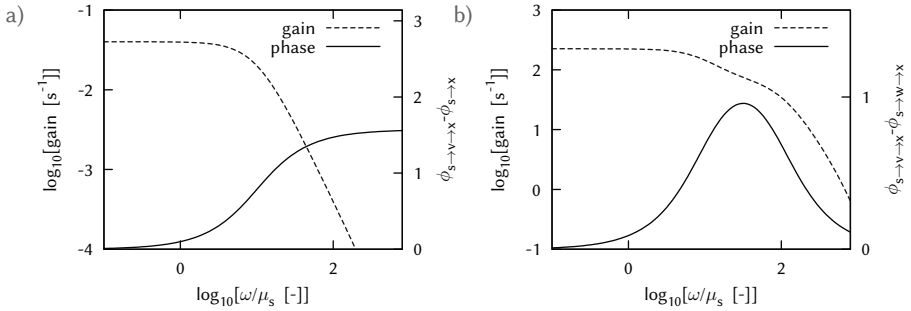


Figure 3.10: In these panels we show the gain (black dashed, right axis) and phase (black solid, left axis). In panel **a** we show the gain and the phase for a ocff-motif. The decrease of the gain corresponds with the increase in the phase difference between the indirect S→V→X and direct pathway S→X. In panel **b** we show a similar plot, but now for the diamond motif with pathways of equal length, but with different degradation rates of the intermediate components. Again we observe that the initial decrease in the gain coincidences with the increase in phase difference. But although the phase difference decrease for larger ω again, the gain continues to decrease, due to the time-averaging of the fluctuations over the lifetime of the intermediate components.

Chapter 4

The Berg-Purcell limit revisited

Biological systems often have to measure at high precision chemical concentrations that are very low. An important question, therefore, is what is the lower bound on the noise in such measurements. Using ideas from theories on diffusion-influenced reactions, we derive an analytical expression for the precision of concentration estimates that are obtained by monitoring the state of a receptor to which a diffusing ligand can bind. The variance in the estimate consists of two terms, one resulting from the intrinsic binding kinetics and the other from the diffusive arrival of ligand at the receptor. The latter term is identical to the fundamental limit derived by Berg and Purcell, but disagrees with the more recent expression of Bialek and Setayeshgar. Comparing the theoretical predictions against results from particle-based simulations shows that the expression of Berg and Purcell and us is highly accurate.

Based on manuscript K. Kaizu, W.H. de Ronde, F. Tostevin, F. Takahashi and P.R. ten Wolde. The Berg-Purcell limit revisited *In preparation.* Simulations are performed by Kazunari Kaizu.

4.1 Introduction

The evidence is accumulating that sensory systems in biology often operate near the fundamental limit set by the noise of counting the signaling molecules. Our visual system can detect single photons [6], some animals can smell single molecules [114], swimming bacteria can respond to the binding and unbinding of only a handful of molecules [83, 115] and eukaryotic cells can be responsive to a difference in $\sim$10 molecules between the front and the back of the cell [7]. Recent experiments suggest that the precision of the embryonic development of the fruitfly *Drosophila* is close to the limit set by the available number of regulatory proteins [8, 116, 117]. This raises the question what is precisely the fundamental limit to the precision of chemical concentration measurements.

In their classic paper, Berg and Purcell considered the scenario in which a cell measures the concentration c of a ligand by monitoring the occupation states of the receptor molecules to which the ligand molecules bind and unbind [83]. A central result is that for a cell that infers the concentration from the time average $\overline{n}$ of the occupation state $n(t)$ of a *single* receptor, where $n(t)$ is zero if the receptor at time t is free and one if it is bound to ligand. The analysis of Berg and Purcell predicts that in the limit that the integration time T is much longer than the correlation time τ_n of the receptor state, the fractional error in the estimate for $\overline{n}$, $n_T = 1/T \int_0^T dt n(t)$, is given by

$$\frac{\delta n}{\overline{n}} = \sqrt{\frac{(1-\overline{n})}{2D\sigma cT}}. \tag{4.1}$$

Here, σ is the receptor-ligand binding cross section, D is the diffusion constant of the ligand, and c is the ligand concentration. The error in the determination of the concentration is related to the error in the measurement of the occupancy via the gain $dc/d\overline{n}$,

$$(\delta c)^2 = \left(\frac{dc}{d\overline{n}}\right)^2 (\delta n)^2, \tag{4.2}$$

yielding the principal result of Berg and Purcell (Eq. [52] in [83]):

$$\frac{\delta c}{c} = \sqrt{\frac{1}{2D\sigma c\,(1-\overline{n})\,T}}. \tag{4.3}$$

This result can be understood intuitively by noting that $4D\sigma c$ is the flux of ligand molecules towards the receptor and $1-\overline{n}$ is the probability that the receptor is free. Importantly, $4D\sigma(1-\overline{n})$ is also the effective binding rate if the binding reaction is diffusion limited, that is, if every receptor-ligand collision is successful. Berg and Purcell argue that their result also holds for reactions that are not deeply in the diffusion-limited regime [83]. They write

> *"Suppose that a molecule that arrives at a vacant binding site sticks with probability α. If it doesn't stick on its first contact, it may soon bump into the site again—and again. If these encounters occur within a time interval short compared to τ_b* [the

> time the molecule is bound], *their result is merely equivalent to a larger value of* α*. As we have no independent definition of the patch radius* s*, we may as well absorb the effective* α *into* s*, writing for the probability that a vacant patch becomes occupied during* dt *simply* $4Dscdt$*."*

This argument seems convincing. On the other hand, after an unsuccessful collision, the ligand molecule may diffuse back into the bulk, and then another ligand molecule may bind to it. Moreover, a ligand molecule that has just dissociated from the receptor, may rapidly rebind to the receptor, possibly a (large) number of times, or it may diffuse away from it into the bulk, after which another molecule may bind to the receptor. It thus remains unclear whether this result also applies to binding reactions that are not diffusion limited.

Bialek and Setayeshgar believed that the analysis of Berg and Purcell relies on a number of assumptions, and they therefore reconsidered this problem [118]. They considered a model in which the ligand molecules can diffuse, bind the receptor upon contact with an intrinsic association rate k_a, and unbind from it with an intrinsic dissociation rate k_d. Invoking the fluctuation-dissipation theorem, they linearized the non-linear reaction-diffusion equation, to obtain the following result for the fractional error in the estimate for the concentration (Eqn. [32] in [118]):

$$\frac{\delta c}{c} = \sqrt{\frac{1}{\pi D \sigma c T} + \frac{2}{k_a c \left(1-\overline{n}\right) T}}. \tag{4.4}$$

The first term arises from the stochastic arrival of the ligand molecules at the receptor by diffusion, while the second term is due to the intrinsic stochasticity of the binding kinetics of the receptor. Indeed, even in the limit that the diffusion constant is infinitely high (and the first term is zero), the concentration can still not be measured with infinite precision, because the receptor still switches between the bound and unbound states with a finite correlation time $\left(k_a c + k_d\right)^{-1}$, leading to noise in the estimate of the receptor occupancy and hence the concentration. This term is absent in Eq. 4.3 since Berg and Purcell assume that the binding reaction is fully diffusion limited, meaning that the intrinsic rates k_a and k_d go to infinity; they argue that the effect of the intrinsic switching kinetics can be captured by renormalizing the cross section.

The first term of Eq. 4.4 should thus be compared with Eq. 4.3: in the limit that the binding reaction is diffusion limited (and the second term in Eq. 4.4 is zero), both theories should yield the same result. However, it is clear that besides the geometrical factor π, which simply comes from the fact that Berg and Purcell assume that the diffusion limited rate is $4\sigma D$ while in the model of Bialek and Setayeshgar it is $4\pi\sigma D$, the expressions differ by a factor $1/(2(1-\overline{n})$. The expression of Bialek and Setayeshgar predicts that in the limit that $\overline{n} \to 1$, the error remains bounded, while the expression of Berg and Purcell suggests that the fractional error diverges; the latter seems reasonable, because if the receptor is occupied most of the time, no new ligand molecules can be counted and hence no new independent measurements can be performed. Only for $\overline{n} = 0.5$ are the two expressions identical (apart from the geometrical factor), but why this would be so, remains unclear.

It is unclear why these two different theories yield different predictions for the funda-

mental limit. Berg and Purcell neglect the intrinsic binding and unbinding dynamics of the receptor, as well as the intricate interplay between binding, unsuccesful ligand-receptor collisions and rapid rebindings, while Bialek and Setayeshgar linearize the non-linear reaction-diffusion equation; given the binary nature of the receptor occupancy, such a linearisation may not be appropriate [68, 119].

In this paper we derive the fundamental limit again, borrowing heavily from the body of work of Agmon, Szabo and coworkers on diffusion-influenced reactions [120]. Like the expression of Bialek and Setayeshgar for the variance in the estimate of the concentration, our expression consists of two terms, one describing the effect of the diffusive transport of the ligand molecules to and from the receptor, the other describing the effect of the intrinsic binding and unbinding kinetics of the receptor. Interestingly, the second term agrees with that of Bialek and Setayeshgar, but the first does not. The first term does, however, agree with the expression of Berg and Purcell (again apart from the geometric factor). We then test these expressions by performing particle-based simulations using our recently developed Green's Function Reaction Dynamics algorithm, which is an exact scheme for simulating reaction-diffusion systems at the particle level [121, 122, 123]. The simulation results agree very well with our expression, leading us to conclude that after 25 years the Berg-Purcell limit still stands as the most accurate expression for the fundamental limit to measuring chemical concentrations. We discuss key assumptions in our analysis, and end by presenting a simple, but intuitive model from which the same fundamental limit can be derived.

4.2 Theory

We consider a single receptor A that is surrounded by a large number N_B of non-interacting ligand molecules B at concentration $c = N_B/V$ in a volume V. We thus consider the pseudo first-order limit, meaning that $N_B \gg N_A = 1$ and $V \to \infty$. Without loss of generality, we may assume that the receptor is static and located at the origin, while the ligand molecules diffuse with diffusion constant D. A ligand molecule can bind a free receptor molecule with an intrinsic association rate k_a when the two come in contact at the contact distance σ, which is the sum of the radii of the two respective molecules. A bound ligand molecule can dissociate from the receptor with an intrinsic dissociation rate k_d. The state of the receptor is denoted by the binary variable $n(t)$, which is one if the receptor is bound to a ligand at time t and zero otherwise. We note that this model is identical to that of Bialek and Setayeshgar for the scenario of a single receptor molecule [118].

Following Berg and Purcell [83] and Bialek and Setayeshgar [118] we imagine that the cell estimates the concentration c from the estimate n_T of the average receptor occupancy $\overline{n}$, which is obtained by integrating $n(t)$ over an integration time T: $n_T = 1/T \int dt n(t)$. The variance in n_T is given by $\sigma^2_{n_T} = \langle n_T^2 \rangle - \langle n_T \rangle^2$, where the angular brackets denote an average over a large number of independent measurements. The variance can be obtained

from the correlation function $C_n(\tau) = \langle (n(\tau) - \langle n \rangle)(n(0) - \langle n \rangle) \rangle$ via (see App. 4.A.1)

$$\sigma^2_{n_T} = \frac{1}{T^2} \int_0^T dt \int_{-t}^{T-t} d\tau C_n(\tau). \tag{4.5}$$

The correlation function defines a correlation time τ_n of the receptor occupancy

$$\tau_n \sigma_n^2 \equiv \int_0^\infty C(\tau)\, d\tau, \tag{4.6}$$

where $\sigma_n^2 = C_n(0) = \langle n^2 \rangle - \langle n \rangle^2$, where for a binary variable this is $\sigma_n^2 = \langle n \rangle (1 - \langle n \rangle) = \overline{n}(1 - \overline{n})$. In the limit that the integration time T is much longer than the correlation time τ_n, the variance in our estimate n_T for $\overline{n}$ is given by

$$\sigma^2_{n_T} = \frac{2\sigma_n^2 \tau_n}{T} = \frac{P_n(\omega = 0)}{T} = \frac{2\mathrm{Re}\left[\hat{C}_n(s = 0)\right]}{T}, \tag{4.7}$$

where $P_n(\omega)$ is the power spectrum of the receptor state n and $\hat{C}_n(s)$ is the Laplace transform of the correlation function $C_n(t)$. The error in the estimate for the concentration $\overline{n}$ can then be obtained from Eq. 4.7 and Eq. 4.2, where $\sigma^2_{n_T} = (\delta n)^2 = \langle n_T^2 \rangle - \langle n_T \rangle^2$ and the gain $dc/d\overline{n} = c/(\overline{n} - \overline{n}^2)$. We will now proceed to compute the Laplace transform of the correlation function, $\hat{C}_n(s)$.

The correlation function of any binary switching process is given by

$$C_n(\tau) = p_*^0 \left(p_{*|*}(\tau) - p_*^0 \right), \tag{4.8}$$

where $p_*^0 \equiv \langle n \rangle$ is the equilibrium probability for the bound state and the probability the receptor is bound $(*)$ at $t = \tau$ given it was bound at $t = 0$ is $p_{*|*}(\tau) = \langle n(\tau) n(0) \rangle / \langle n \rangle$. To obtain the correlation function, we thus need $p_{*|*}(\tau)$. It is convenient to focus on the conjugate "survival probability"

$$\mathscr{S}_{\mathrm{rev}}(t|*) = 1 - p_{*|*}(t), \tag{4.9}$$

which is the probability that the receptor is free at time t given that it was bound at $t = 0$. Following Agmon and Szabo [120], we use the subscript "rev" to indicate that we consider a *reversible* reaction, meaning that in between $t = 0$ and t the receptor may bind and unbind ligand a number of times. The probability that a receptor-ligand pair dissociates between t' and $t' + dt'$ to form an unbound pair at contact is $k_d[1 - \mathscr{S}_{\mathrm{rev}}(t'|*)]dt'$, while the probability that the receptor with now a ligand molecule at contact is still unbound at time $t > t'$ is $\mathscr{S}_{\mathrm{rad}}(t - t'|\sigma)$; the subscript "rad" means that we now consider an *irreversible* reaction with $k_d = 0$, which can be described by solving the diffusion equation using a "radiation"

boundary condition [120]. Hence, $\mathscr{S}_{\text{rev}}(t|*)$ is given by [120]

$$\mathscr{S}_{\text{rev}}\left(t|*\right) = k_d \int_0^t \left[1 - \mathscr{S}_{\text{rev}}\left(t'|*\right)\right] \mathscr{S}_{\text{rad}}\left(t - t'|\sigma\right) dt'. \tag{4.10}$$

We emphasize that up to this point no approximation has been made. The question now is what is $\mathscr{S}_{\text{rad}}(t|\sigma)$, which is the quantity needed to solve Eq. 4.10. In general, it is not possible to obtain an exact analytical expression for this quantity. Imagine a bound receptor-ligand pair that is surrounded by an equilibrium, i.e. uniform, distribution of ligand particles. If this receptor-ligand pair dissociates to form a receptor-ligand pair at contact surrounded by an equilibrium distribution of ligand molecules, then the probability that the receptor is still unbound at a later time t is given by the survival probability [120]

$$\mathscr{S}_{\text{rad}}\left(t|\sigma\right) = \mathscr{S}_{\text{rad}}\left(t|\text{eq}\right) S_{\text{rad}}\left(t|\sigma\right), \tag{4.11}$$

where $\mathscr{S}_{\text{rad}}(t|\text{eq})$ is the probability that a receptor which initially is free and surrounded by an equilibrium distribution of ligand molecules remains free until at least a later time t, while $S_{\text{rad}}(t|\sigma)$ is the probability that a free receptor that initially is surrounded by only one *single* ligand molecule at contact is still unbound at a later time t. Now, the ligand molecule at contact may either rebind the receptor or diffuse away from it. If it rebinds the receptor, then after the next dissociation event, the probability that the receptor will remain free for at least another time t will again be given by Eq. 4.11. The problem arises when the ligand molecule at contact instead diffuses away from the receptor and another ligand molecule binds the receptor: If this second ligand molecule dissociates from the receptor before the first ligand has relaxed to equilibrium, then the assumption of Eq. 4.11 breaks down. Indeed, the process of receptor binding generates non-trivial spatio-temporal correlations between the positions of the ligand molecules, which depend on the history of the association and dissociation events. This impedes an exact solution of the problem. However, if the dissociation rate k_d is low then it becomes reasonable to assume that after *each* dissociation event, the unbound receptor-ligand pair at contact is surrounded by an equilibrium distribution of B particles [120], in which case the survival probability is given by Eq. 4.11. This is the crucial assumption that we will make in our analysis.

With the assumption of Eq. 4.11, Eq. 4.10 can now be solved in the Laplace domain. For a pseudo-first-order irreversible reaction with a static target [124], $\mathscr{S}_{\text{rad}}(t|\text{eq})$ of Eq. 4.11 is given by

$$\mathscr{S}_{\text{rad}}\left(t|\text{eq}\right) = e^{-c\int_0^t k_{\text{rad}}(t')dt'}, \tag{4.12}$$

where $k_{\text{rad}}(t)$ is the time-dependent rate coefficient. Moreover, $S_{\text{rad}}(t|\sigma)$ in Eq. 4.11 is via detailed balance (and the backward Smoluchovski equation) related to $k_{\text{rad}}(t)$: $k_{\text{rad}}(t) = k_a S_{\text{rad}}(t|\sigma)$ [120]. Together, these relations yield a simple expression for the Laplace transform of $\mathscr{S}_{\text{rad}}(t|\sigma)$ in terms of the Laplace transform $\hat{\mathscr{S}}_{\text{rad}}(s|\text{eq})$ of $\mathscr{S}_{\text{rad}}(t|\text{eq})$, as described in App. 4.A.2. Substituting this in the solution of Eq. 4.10 in the Laplace domain allows us to obtain the following expression for the Laplace transform of the correlation function in

terms of $\hat{\mathscr{S}}_{\text{rad}}(s|\text{eq})$:

$$\hat{C}(s) = \sigma_n^2 \frac{\langle n \rangle \hat{\mathscr{S}}_{\text{rad}}(s|\text{eq})}{1 - (1 - \langle n \rangle) s\hat{\mathscr{S}}_{\text{rad}}(s|\text{eq})}, \tag{4.13}$$

$$= \sigma_n^2 \frac{ck_a\tau_c\hat{\mathscr{S}}_{\text{rad}}(s|\text{eq})}{1 - k_d\tau_c s\hat{\mathscr{S}}_{\text{rad}}(s|\text{eq})}, \tag{4.14}$$

where $\langle n \rangle = k_a c/\tau_c$ and $\tau_c = (k_a c + k_d)^{-1}$ is the correlation time of the intrinsic receptor switching dynamics, i.e. the correlation time of the receptor occupancy when receptor-ligand association is reaction-limited and the effect of diffusion can be neglected.

To obtain an analytically closed form for the correlation function, we require an expression for $\hat{\mathscr{S}}_{\text{rad}}(s|\text{eq})$. In the limit that the concentration c is small, we can expand and Laplace transform Eq. 4.12 to obtain (see Eq. A4.50, [120])

$$s\hat{\mathscr{S}}_{\text{rad}}(s|\text{eq}) \approx \left(1 + c\hat{k}_{\text{rad}}(s)\right)^{-1}. \tag{4.15}$$

Noting that $\hat{k}_{\text{rad}}(s) = k_a\hat{k}_{\text{abs}}(s)/(k_a + sk_{\text{abs}}(s))$, where $\hat{k}_{\text{abs}}(s) = 4\pi\sigma D(1 + \sigma\sqrt{s/D})$ is the Laplace transform of the time-dependent diffusion-limited rate constant $k_{\text{abs}}(t)$, and substituting the above expression in Eq. 4.13 yields

$$\hat{C}(s) = \sigma_n^2 \frac{\tau_c'(s)}{s\tau_c'(s) + 1}, \tag{4.16}$$

where $\tau_c'(s)$ is the intrinsic correlation time τ_c renormalized by the concentration fluctuations [118]:

$$\tau_c'(s) = \tau_c(1 + \Sigma(s)); \ \Sigma(s) = \frac{k_a}{k_D\left(1 + \sqrt{s\tau_m}\right)}, \tag{4.17}$$

with the diffusion-limited rate constant $k_D = 4\pi\sigma D$ and the molecular time scale $\tau_m = \sigma^2/D$. The correlation time τ_n of the receptor state is given by $\tau_n = \hat{C}(s=0) = \tau_c'(s=0)$:

$$\tau_n = \frac{1}{k_{\text{on}}c + k_{\text{off}}}, \tag{4.18}$$

where k_{on} and k_{off} are the renormalized association and dissociation rates

$$k_{\text{on}} = \left(\frac{1}{k_a} + \frac{1}{k_D}\right)^{-1} = \frac{k_a k_D}{k_a + k_D}, \tag{4.19}$$

$$k_{\text{off}} = \left(\frac{1}{k_d} + \frac{K_{\text{eq}}}{k_D}\right)^{-1} = \frac{k_d k_D}{k_a + k_D}. \tag{4.20}$$

and $K_{\text{eq}} = k_a/k_d$ is the equilibrium constant.

The error in the estimate of the concentration can be obtained by combining Eq. 4.16

with Eq. 4.7 and Eq. 4.2, yielding our principal result

$$\frac{\delta c}{c} = \sqrt{\frac{1}{2\pi\sigma D c\,(1-\overline{n})\,T} + \frac{2}{k_a c\,(1-\overline{n})\,T}}. \tag{4.21}$$

The first term describes the error in our estimate that stems from the stochastic diffusive arrival of the ligand molecules, while the second term describes the error in our estimate that results from the intrinsic binding dynamics of the receptor. If the receptor-ligand association reaction is fully reaction-limited, meaning that $k_a, k_d \to 0$ or $D \to \infty$, then the first term is zero: the error in our estimate of the concentration is then limited by the intrinsic binding and unbinding dynamics of the receptor. Conversely, if the reaction is diffusion limited, meaning that $k_a, k_d \to \infty$ or $D \to 0$, then the second term is zero, and the precision of our concentration estimate is limited by the diffusive arrival and departure of the ligand molecules to and from the receptor. This is considered to be the fundamental limit to the accuracy of measuring chemical concentrations: the noise in our concentration measurement has a floor that is set by the physics of diffusion, independent of the kinetic parameters k_a and k_d.

It is clear that the second term in Eq. 4.21 is identical to that in the expression of Bialek and Setayeshgar, Eq. 4.4. Yet, the first term, that determines the fundamental limit, is different: the expression of Bialek and Setayeshgar misses a factor $2/(1-\overline{n})$. The Berg-Purcell expression does contain this factor, and indeed, apart from a geometrical factor, their expression is identical to ours in the limit that the reaction is fully diffusion limited (and the second term in Eq. 4.21 is zero). Intuitively, the factor $1-\overline{n}$ makes sense: the rate of counting molecules is determined by the rate at which molecules arrive, which is $4\pi\sigma D$, times the probability $1-\overline{n}$ that the receptor is free and capable of actually binding and detecting the ligand molecules. Our analysis thus suggests that the Berg-Purcell limit is the most accurate expression for the precision by which concentrations can be measured.

4.3 Numerical Results

To test our theory, we have performed particle-based simulations. A key quantity of our theory is $\hat{C}(s)$, Eq. 4.16, since the precision of our concentration estimate directly follows from this quantity and the gain $dc/d\overline{n}$ (see Eq. 4.7 and Eq. 4.2). We will therefore compare the power spectrum, $P_n(\omega) = 2\mathrm{Re}[\hat{C}_n(s=i\omega)]$ with $\hat{C}_n(s)$ given by Eq. 4.16, to that obtained from simulations. The particle-based simulations have been performed using our Green's Function Reaction Dynamics (GFRD) algorithm [121, 122, 123], which is an *exact* scheme for simulating reaction diffusion systems at the particle level. The model of the simulations is almost identical to that studied in our theory. It consists of a static single receptor in the center of the simulation box with linear dimension L, surrounded by ligand molecules that diffuse with diffusion constant D. A ligand molecule that is in contact with a free receptor at the contact distance σ can associate with the receptor with an intrinsic association rate k_a and then dissociate from it with an intrinsic dissociation rate k_d; after dissociation, the ligand molecule is put at contact again.

Figure 4.1 shows the power spectrum as obtained from the simulations (black line) together with the prediction of our theory (grey line; Eq. 4.16), for two different concentrations, $c = 0.1\mu$M (panel a) and $c = 100\mu$M (panel b). The first concentration is in the biologically relevant regime $1\text{nM} \lesssim c \lesssim 1\mu\text{M}$, while the second is much larger than that. Each panel also shows power spectra for switching processes with uncorrelated and exponentially distributed waiting times, one with the intrinsic correlation time $\tau_c = (k_a c + k_d)^{-1}$ (red dashed line) and one that has the same effective correlation time as that of our theory, $\tau_n = \tau_c'(s = 0) = (k_{\text{on}}c + k_{\text{off}})^{-1}$ (red solid line); since the variance of any binary switching process is $\sigma_n^2 = \overline{n}(1 - \overline{n})$, the power spectrum of the second exponential switching process at zero frequency, $P_n(\omega = 0) = 2\sigma_n^2\tau_n$, is also equal to that of our theory. It is seen that the power spectrum predicted by our theory, by Eq. 4.16, nicely interpolates between the respective spectra of the two exponential switching processes. Interestingly, not only in the high-frequency regime, but also in the low-frequency regime, the simulation results agree very well with the prediction of Eq. 4.16, both at low and high concentration. In fact, for the biologically relevant concentration, the agreement between our theory and simulation is very good over essentially the full frequency range. A deviation is seen only in the intermediate frequency regime for the high concentration, outside the biologically relevant range.

The high frequency regime of the power spectrum corresponds to the intrinsic switching dynamics of the receptor. In this regime, diffusion hardly plays any role and the receptor dynamics is dominated by the binding of ligand molecules that are essentially in contact with the receptor, which occurs with an exponentially distributed waiting time with mean $(k_a c)^{-1}$, and the unbinding of ligand, which is a Poisson process with mean waiting time k_d^{-1}. Fig. 4.1 shows that this regime is well described by our theory.

The intermediate frequency regime of the power spectrum starts at $\omega_m = 1/\tau_m = D/\sigma^2$. It corresponds to the regime in which a ligand molecule, after dissociation from the receptor, manages to diffuse away from the receptor over a few molecular distances σ, but then rebinds the receptor before another ligand molecule from the bulk does. Fig. 4.1 shows that our theory describes this regime at low concentrations quite well, but at high concentrations a deviation is visible. This can be traced back to the break down of our key assumption, Eq. 4.11, which we discuss in detail in the next section.

The low frequency regime of the power spectrum corresponds to the regime in which after receptor dissociation the ligand molecule diffuses into the bulk and, most likely, another molecule from the bulk binds the receptor. The agreement between the power spectrum of the simulations and that of the full theory, and on the other hand that of an exponential switching process, suggests that in this regime receptor binding and unbinding has become memoryless. This seems plausible: when the dissociated molecule enters the bulk, it will eventually loose memory of where it came from; also the other ligand molecules will bind the receptor in a memoryless fashion. However, this idea is only correct if the system is bounded. Indeed, it is well known that in an unbounded system, the correlation function of a reversible reaction exhibits an algebraic tail [125, 126]. The simulation box of our system is, like a living cell, finite, and the collisions of the ligand molecules with the cell boundaries randomize their trajectories on time scales larger than L^2/D; consequently, the correlation

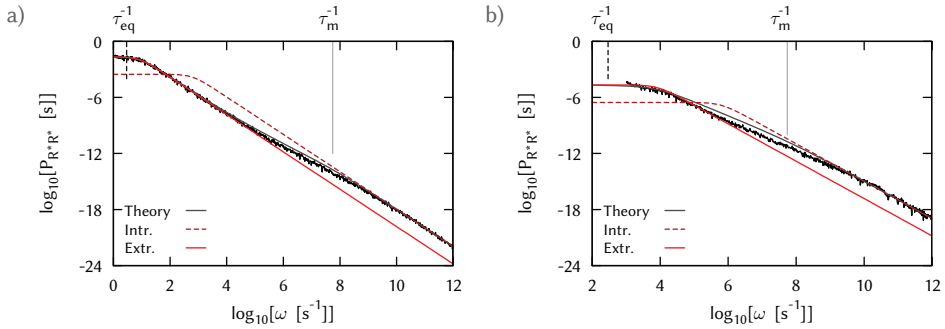


Figure 4.1: The power spectrum of the receptor state $P_n(\omega)$. **a)** c=0.1 μM, **b)** c=100 μM. The simulation results (black) agree very well with the theoretical prediction of Eq. 4.16 (gray line). At high frequencies $\omega > 1/\tau_m = D/\sigma^2$, the receptor dynamics is that of an exponential switching process with the intrinsic binding and unbinding rates $k_a c$ and k_d, respectively, as shown by the dashed red line; in this regime, the ligand remains close to the receptor and the effect of diffusion is negligible. At low frequencies, the receptor dynamics is to a very good approximation given by a random telegraph process with effective on and off rates $k_{on} c$ and k_{off}, respectively, as shown by the solid red line. Note that the correlation time of that process, $\tau_n = (k_{on} c + k_{off})^{-1}$, is by definition equal to that of our theory, Eq. 4.16; this means that also $P_n(\omega=0) = 2\sigma_n^2 \tau_n$, which determines the error in our estimate of the receptor occupancy and the concentration, Eq. 4.7, is the same. Above each panel, two time scales are shown: $\tau_{eq} = L^2/D$ (the time to cross the simulation box) and $\tau_m = \sigma^2/D$ (the time to diffuse a molecular diameter). We observe that τ_m marks the transition from the regime in which the receptor dynamics is dominated by the intrinsic binding kinetics to that which also depends on the diffusion of the ligand molecules towards and away from the receptor. At high concentration (panel b), the simulation results differ slightly from the prediction of Eq. 4.16 in the intermediate regime. We believe this is due to interference of bulk molecules with the rebinding of the dissociated ligand molecule. Parameters: $\bar{n}$=0.9, D=1 $\mu m^2\, s^{-1}$, σ=10 nm. **a)** L=584 nm, k_d=61 s; **b)** L=58 nm, k_d=17 hr.

function is exponential at long times. The correlation function of our theory does exhibit an algebraic tail, which is the remant of the factor $S_{\mathrm{rad}}(t|\sigma)$ in Eq. 4.11, which is the probability that a free receptor with a single molecule at contact (and no other molecules present) is still unbound at a later time t *in an unbounded system.* However, in Eq. 4.11 we assume that the other ligand molecules, forming the "bulk", do have a uniform distribution. These molecules will bind the receptor in a memoryless fashion. Moreover, at long times, they will dominate the binding of ligand to receptor. In our theory, this crossover time can be estimated from

$$-\frac{d\mathscr{S}_{\mathrm{rad}}(t|\sigma)}{dt} = -\overbrace{S_{\mathrm{rad}}(t|\sigma)\frac{d\mathscr{S}_{\mathrm{rad}}(t|\mathrm{eq})}{dt}}^{\text{bulk}-\text{binding}} - \overbrace{\mathscr{S}_{\mathrm{rad}}(t|\mathrm{eq})\frac{dS_{\mathrm{rad}}(t|\sigma)}{dt}}^{\text{rebinding}}, \tag{4.22}$$

which is the propensity function for receptor binding, i.e. the probability that a receptor with a ligand molecule at contact and surrounded by a uniform distribution of ligand molecules, binds a ligand molecule for the first time at a later time t. The first term is the probability that this ligand molecule is one from the bulk, while the second gives the probability

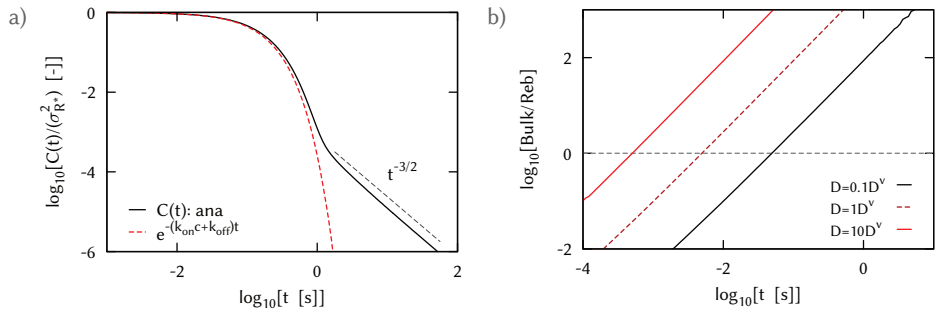


Figure 4.2: a) The correlation function C(t) for Eq. 4.16 (black solid) and an exponential switching process with waiting times k_{on},k_{off} (red dashed). The correlation time for both processes is the same. Parameters: D=1 $\mu m^2\,s^{-1}$, n=0.5, c=0.1 μM, σ=10 nm, k_a=5520 $\mu M^{-1}\,s^{-1}$ and k_d= 613 s. **b)** The ratio of the change in $\mathscr{I}_{rad}$ for a bulk first binding and for a rebinding of a dissociated B particle. The bulk dominates the change in the survival probability already for very small t. Indeed for D=1 $\mu m^2\,s^{-1}$, the association timescales are $t_a=1/k_a c\approx 0.02$ s, $t_{on}=1/k_{on}c\approx 0.15$ s Parameters: as in panel **a)**, D^v=1 $\mu m^2\,s^{-1}$, k_a=552 $\mu M^{-1}\,s^{-1}$.

that this is the ligand molecule that was in contact, diffused away, but then came back to rebind the receptor. Fig. 4.2b shows that only at very short times, rebindings dominate the bulk bindings. For long times, receptor binding is completely dominated by the binding of molecules from the bulk, which in our theory, as in the simulations, bind the receptor in a memoryless fashion. The algebraic tail of the correlation function in our theory is thus very small, as shown in Fig. 4.2a. Indeed, at long times $t \gg \tau_m$ the receptor dynamics is to a very good approximation a random telegraph process with an on rate $k_{\text{on}}c$ and an off rate k_{off}.

The most important point of the power spectrum is at zero-frequency, $P_n(\omega = 0)$: that determines the correlation time of the receptor and hence the error in our estimate of the average receptor occupancy $\overline{n}$ and the average concentration c, see Eq. 4.7 and Eq. 4.2. The simulation results of Fig. 4.1 show that Eq. 4.16 captures this limit very well, not only at the low concentration of 100nM, but, perhaps surprisingly, also at the high concentration of 100μM. Fig. 4.3 shows $P_n(\omega = 0)$ as a function of the average receptor occupancy $\overline{n}$ at a concentration of 1nM. It is seen that the agreement between theory and simulations is very good. Fig. 4.3 also shows the prediction of Bialek and Setayeshgar for $P_n(\omega = 0)$ [118]. Only for $\overline{n} = 0.5$ does their prediction agree with our prediction and simulation results. Moreover, while their analysis predicts that $P_n(\omega = 0) = P_{1-n}(\omega = 0)$, our results show that the dependence of $P_n(\omega)$ on n is non-symmetric, which reflects the fact that when the receptor is free, more binding events can be counted, leading to a more accurate estimate of the concentration. Since the Berg-Purcell formula in Eq. 4.3 directly follows from our expression for $P_n(\omega = 0) = 2\text{Re}[\hat{C}(s = 0)]$, via Eq. 4.16, Eq. 4.7 and Eq. 4.2, we conclude that the Berg-Purcell limit provides an accurate upper bound on the precision by which chemical concentrations can be measured.

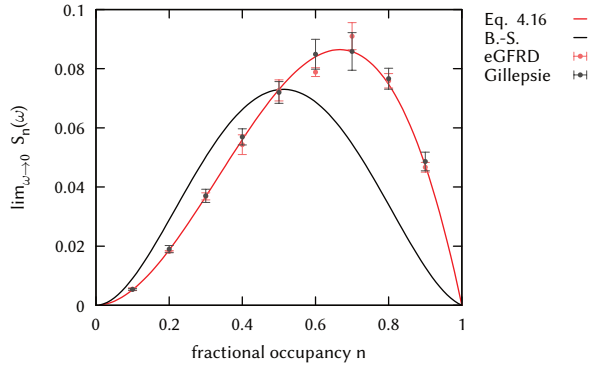


Figure 4.3: The power spectrum in the zero-frequency limit, $P_n(\omega{=}0)$, as a function of the average receptor occupancy $\bar{n}$, for c=100 nM. The receptor occupancy is varied by changing k_d. It is seen that the agreement between our theoretical prediction of Eq. 4.16 and the simulation results is very good (red line). In contrast, the prediction of Bialek and Setayeshgar [118] (black line) differs markedly from our results. While their analysis predicts that $P_n(\omega){=}P_{1-n}(\omega)$, our results show that $P_n(\omega)$ is non-symmetric in n. This reflects the fact that when the receptor occupancy is low more binding events can be counted. Other parameters: L=584 nm,D=1 $\mu m^2 s^{-1}$,$k_a{=}10k_D{=}3.76{\times}10^2\ \mu M^{-1}s^{-1}$, σ=5 nm.

4.4 Validity assumption under biological conditions

The key assumption of our theory is that of Eq. 4.11, which states that after dissociation the unbound receptor-ligand pair is surrounded by a uniform distribution of ligand molecules. This assumption breaks down when a) after receptor dissociation, the rebinding of the ligand molecule to the receptor is preempted by the receptor binding of another, second, ligand molecule and b) this second ligand molecule dissociates from the receptor before the first has diffused into the bulk. We argue that under biologically relevant conditions both conditions a) and b) are not satisfied and that therefore the key assumption of our analysis holds.

Firstly, rebinding trajectories of ligand molecules that have just dissociated from the receptor are so short that the likelihood that a molecule from the bulk interferes with such a rebinding event is negligible, especially when the concentrations are low. In this case, condition a) is not met and a dissociated ligand molecule rebinds the receptor before it diffuses into the bulk as often as when it would be the only ligand molecule present in the system. To make this intuition quantitative, we note that if the unbound receptor-ligand pair at contact is surrounded by an equilibrium distribution of ligand molecules, then the probability that a ligand molecule from the bulk does not interfere with the receptor rebinding of the ligand molecule at contact, is

$$1 - p_{\text{int}} = -\int_0^\infty \frac{dS_{\text{rad}}\left(t|\sigma\right)}{dt} \mathscr{S}_{\text{rad}}\left(t|\text{eq}\right) dt + S_{\text{rad}}\left(\infty|\sigma\right), \tag{4.23}$$

where the first term is the integral of the second term in Eq. 4.22 — it is the probability that the ligand molecule which has just dissociated from the receptor rebinds the receptor before

a ligand molecule from the bulk does; the second term $S_{\text{rad}}(\infty|\sigma)$ is the probability that the ligand molecule at contact (with no other ligand molecules present) escapes into the bulk — when the ligand at contact escapes into the bulk, then, by definition, no bulk molecule can interfere with its rebinding. Combining the above expression with Eq. 4.22 shows that the probability of rebinding interference is

$$p_{\text{int}} = -\int_0^\infty dt d\mathscr{S}_{\text{rad}}(t|\text{eq})/dt\left[S_{\text{rad}}(t|\sigma) - S_{\text{rad}}(\infty|\sigma)\right], \tag{4.24}$$

which is indeed the probability that a molecule from the bulk binds the receptor before the ligand molecule which started at contact and that would have rebound the receptor if there were no other ligand molecules, does. Fig. 4.4 shows that for biologically relevant concentrations and diffusion constants, the probability of rebinding interference P_{int} is very small, which means that the central assumption, Eq. 4.11, will hold.

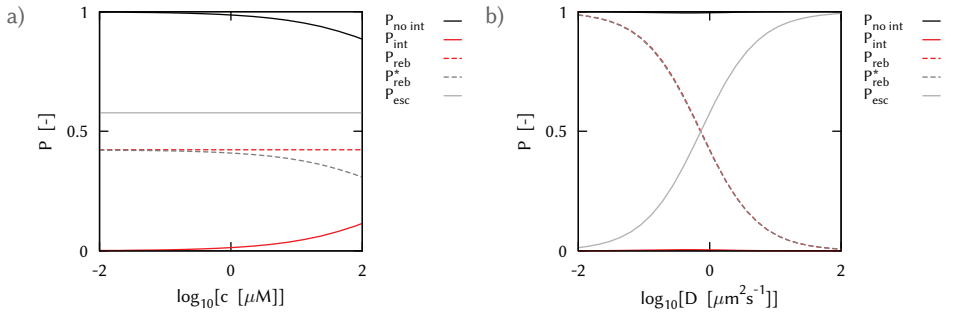


Figure 4.4: Probability p_{int} (red line) that a ligand molecule from the bulk interferes with the receptor rebinding of a dissociated ligand molecule as a function of the concentration c (panel a) and diffusion constant D (panel b). The probability $1-p_{\text{int}}$ of no interference (black line) is given by Eq. 4.23 and consists of two terms: the probability p^*_{reb} that the dissociated ligand molecule rebinds the receptor before a molecule from the bulk does (dashed grey line) and the probability $p_{\text{esc}}=S_{\text{rad}}(\infty|\sigma)=k_D/(k_a+k_D)$ that it escapes into the bulk (solid grey line; second term). It is seen that for biologically relevant concentrations $\text{nM}\lesssim c\lesssim\mu\text{M}$ and diffusion constants $0.1\,\mu\text{m}^2\text{s}^{-1}\lesssim D\lesssim 10\,\mu\text{m}^2\text{s}^{-1}$ bulk molecules hardly interfere with the receptor rebinding of dissociated ligand molecules. This is the motivation for the simplified model of Fig. 4.5, in which $\mathscr{S}_{\text{rad}}(t|\text{eq})$ in Eq. 4.23 is taken to be unity, and the probability of rebinding is approximated as $p_{\text{reb}}=1-p_{\text{esc}}=S_{\text{rad}}(\infty|\sigma)=k_a/(k_a+k_D)$ (dashed red line); note that the correspondence between p_{reb} and the full expression p^*_{reb} in the biologically relevant regime. Parameters are $D=1\,\mu\text{m}^2\text{s}^{-1}$ (panel **a**), $c=1\,\mu\text{M}$ (panel **a**) and $k_a=55\,\mu\text{M}^{-1}\text{s}^{-1}$, $\sigma=10\,\text{nm}$.

Yet, even when occassionally rebinding interferences do occur, and condition a) is met, then Eq. 4.11 is still likely to hold, because condition b) is not met. For a biologically relevant but high concentration $c = 1\,\mu\text{M}$ and high effective association rate $k_{\text{on}} = 1\,\mu\text{M}^{-1}\text{s}^{-1}$, and for $\overline{n} = 1/2$, $D = 1\,\mu\text{m}^2\text{s}^{-1}$ and $\sigma = 10\,\text{nm}$, the time a ligand molecule is bound to the receptor is $t_d = k_d^{-1} \approx 1\,\text{s}$. The time for a dissociated ligand molecule to equilibrate in a cell of linear dimension L is upper bounded by L^2/D, which is on the order of a second if $L =$

$1\,\mu$m and $D = 1\,\mu\text{m}^2\text{s}^{-1}$. We thus expect that for cells of micron size and concentrations that are in the nM to μM range, also condition b) is not met, which means that even when rebinding interferences do arise, Eq. 4.11 still holds.

4.5 A simple coarse-grained model

Our simulation results show that in the relevant low frequency regime, the receptor dynamics is to a very good approximation that of a random telegraph process with on rate $k_{\text{on}}c$ and off rate k_{off}. This suggests the following simple picture, inspired by our earlier analysis of the effect of diffusion of gene regulatory proteins on noise in gene expression [121].

This simplified model is based on the observations that rebindings are fast and the probability of rebinding interference is negligible. When molecules from the bulk do not interfere with the rebinding of a dissociated ligand molecule, we can simply estimate the number of rounds N_{reb} the ligand molecule rebinds to and dissociates from the receptor before it diffuses into the bulk as: $N_{\text{reb}} = (1 - p_{\text{reb}}) \sum_{i=0}^{\infty} i p_{\text{reb}}^i = p_{\text{reb}}/(1 - p_{\text{reb}})$, where p_{reb} is the probability that a ligand molecule which has just dissociated from the receptor rebinds the receptor before it diffuses into the bulk in the absence of any rebinding interference; this is the first term on the right-hand side of Fig. 4.4 with $\mathscr{S}_{\text{rad}}(t|\text{eq}) = 1$, yielding $p_{\text{reb}} = 1 - S_{\text{rad}}(\infty|\sigma) = k_a/(k_a + k_D)$, where $k_D = 4\pi\sigma D$ is the diffusion-limited association rate. This gives $N_{\text{reb}} = k_a/k_D$. When the rebinding trajectories are short compared to the time $t_d = k_d^{-1}$ the ligand molecule is bound to the receptor, we can estimate the total time t_{on} the receptor is bound to the receptor before it diffuses into the bulk simply as $t_{\text{on}} = t_d(1 + N_{\text{reb}})$ (see Fig. 4.5). The effective rate of dissociation is then $k_{\text{off}} = 1/t_{\text{on}} = k_d/(1 + N_{\text{reb}}) = k_d k_D/(k_a + k_D)$, which is precisely the effective dissociation rate predicted by our theory, Eq. 4.20. This argument shows that, because the rebindings are so fast, their non-trivial dynamics can be captured by a simple rescaling of the dissociation rate by the average number of rebindings. Since receptor binding is an equilibrium process, when we rescale the dissociaton rate, we must also renormalize the association rate, to obey the detailed-balance condition $\overline{n}/(1 - \overline{n}) = k_a/k_d = k_{\text{on}}/k_{\text{off}}$. It follows that also the association rate k_a must be scaled with $1+N_{\text{reb}}$, yielding $k_{\text{on}} = k_a/(1+N_{\text{reb}}) = k_a k_D/(k_a+k_D)$, which, indeed, is the effective association rate of our theory, Eq. 4.19.

Another interpretation of the binding dynamics is that the effective dissociation rate k_{off} is the intrinsic dissociation rate k_d times the probability $1 - p_{\text{reb}} = S_{\text{rad}}(\infty|\sigma) = k_D/(k_a + k_D)$ that the molecule subsequently escapes into the bulk, while the effective association rate of a ligand molecule is the rate at which the molecule arrives from the bulk at the receptor, which is the diffusion limited rate k_D, times the probability that it subsequently binds the receptor, which is $p_{\text{reb}} = k_a/(k_a + k_D)$. Also this interpration hinges on the idea that molecules from the bulk do not interfere with the future dynamics of a ligand molecule that is in contact with the receptor, be it receptor binding if the molecule came from the bulk, or receptor rebinding or diffusing into the bulk if it came from the receptor.

Once we have captured the effect of diffusion in the renormalized binding rates, the

receptor binding dynamics is simply given by the linear equation

$$\frac{dn\left(t\right)}{dt} = k_{\text{on}}\left(1 - \left(n\left(t\right)\right)\right) - k_{\text{off}} n\left(t\right), \tag{4.25}$$

from which the power spectrum $P_n(\omega)$ and the error in our estimate of $\overline{n}$ and c can be obtained straightforwardly. This indeed yields the result of Eq. 4.21.

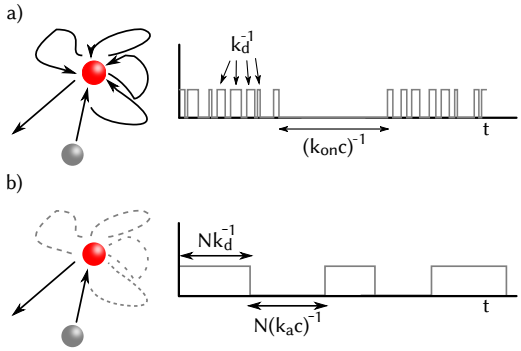


Figure 4.5: Cartoon of the coarse-grained model. Panel a) shows a schematic of a typical time trace of the receptor occupation state n(t). Panel b) shows a time trace in the coarse-grained model. Here, N is the number of rebindings N_{reb} plus one.

4.6 Discussion

Using results from the theory on reversible diffusion-influenced reactions of Amgon and Szabo [120] we have derived the physical limits to the precision of measuring chemical concentrations via the binding of a diffusing ligand to a receptor. Our approach does not only take ligand diffusion into account, but also the intrinsic binding kinetics at the receptor. Our principal result, Eq. 4.21, agrees with that of Bialek and Setayeshgar [118] in the sense that the variance of the concentration estimate consists of two terms, one describing the effect of the diffusive transport of ligand to and from the receptor and the other the effect of the intrinsic binding kinetics given that the ligand is in contact. However, while the second term is identical to that of Bialek and Setayeshgar [118], the first is not. But this term is identical to that of Berg and Purcell; it differs from that of Bialek and Setayeshgar in that it contains an additional factor $1/(2(1-\overline{n}))$, which takes into account that molecules can only be counted when the receptor is free.

A comparison of the power spectrum of the receptor occupancy as predicted by our theory against results from particle-based simulations reveals that our theory is very accurate under biologically relevant conditions. The key assumption of our theory is that after a ligand has dissociated from the receptor, the unbound receptor-ligand pair is surrounded by an equilibrium distribution of ligand molecules. The success of this assumption lies in the fact that rebinding trajectories are very short compared to the time the molecule is bound and

that biologically relevant concentrations are low, which means that ligand molecules from the bulk do not interfere with the rebinding trajectories of dissociated ligand molecules. Under these conditions, the rebindings can be integrated out, yielding effective association and dissociation rates.

Our theory also gives a fairly accurate prediction for the power spectrum over the full frequency range, even though our theory is for an unbounded domain while the simulation box is finite. However, the central assumption of our theory that the unbound receptor-ligand pair at contact is surrounded by a uniform distribution of ligand, is more likely to hold in a finite domain, since encounters of ligand molecules with the wall tend to randomize their positions. One may nonetheless wonder how accurately our description is for the binding of ligand to a receptor in an unbounded domain—this question might be relevant for the binding of *extracellular* ligand to receptors at the outside surface of a cell, although even the extracellular environment of most, if not all, organisms is not an empty unbounded space. It is known that concentration fluctuations in an unbounded domain lead to a long algebraic tail in the correlation function, the magnitude of which is underestimated by our theory [125, 126]. Yet, in the limit that the concentration is low the correlation time, which is the integral of the normalized correlation function, is accurately predicted by our description [125, 126]. When the integration time is long compared to the correlation time, only the integral and not the precise shape of the correlation function determines the error in our concentration estimate, Eq. 4.7, and we thus expect that our expression for the fundamental limit also holds in this scenario.

Finally, our observations have important implications for the modeling of intracellular biochemical networks. They underscore our earlier observation [121] that when the cell does not exhibit macroscopic concentration gradients on cellular length scales, the effect of diffusion can often be captured in a well-stirred model, which can then be simulated using the Gillespie algorithm [75] instead of a much more computationally demanding particle-based algorithm [121, 123]. In such a well-stirred model, the rapid rebindings are integrated out and association and dissociation occur in a memoryless fashion, with exponentially distributed waiting times with mean k_{on}^{-1} and k_{off}^{-1}, respectively. This is a simplification — at short times the association-time distribution is algebraic due to the rapid rebindings while only at times that are long compared to the time to travel the cell diameter the association-time distribution is memoryless — but it is an accurate one: the high-frequency noise from the rapid rebindings is typically filtered by the network downstream; only the low-frequency noise obeying exponential statistics is significantly propagated downstream. In fact, this model has the correct zero-frequency limit of the power spectrum, which determines how accurate the receptor can be read out and how much noise is propagated down the signaling pathway. However, we emphasize that this approach of integrating out the rebindings cannot be used when the rebindings can change the future dynamics of the network, as in the MAPK pathway where a single rebinding event can cause irreversible modification of a reactant [123]. Indeed, in such systems that employ multi-site protein modification, rebindings can qualitatively change the average macroscopic behavior of the system [123]. These systems have to be treated as reaction-diffusion systems and modeled at the particle level, even when they are uniform at the cellular scale.

4.A Supplementary Information

4.A.1 The correlation function

The time-average of an observable $A(t)$ [127] is defined as

$$A_T = \frac{1}{T}\int_0^T dt A(t), \tag{A4.1}$$

and the variance of the time-averaged mean $\sigma^2_{A_T}$ is

$$\sigma^2_{A_T}(T) = \left\langle A_T^2 \right\rangle - \left(\frac{1}{T}\int \langle A(t)\rangle dt\right)^2 \tag{A4.2}$$

$$= \frac{1}{T^2}\int_0^T\int_0^T dt dt' \left\langle A(t) A\left(t'\right)\right\rangle - \langle A_T\rangle^2 \tag{A4.3}$$

$$= \frac{1}{T^2}\int_0^T dt \int_{-t}^{T-t} d\tau \left\langle A(0) A\left(t'-t\right)\right\rangle - \langle A_T\rangle^2 \tag{A4.4}$$

$$= \frac{1}{T^2}\int_0^T dt \int_{-t}^{T-t} d\tau \langle A(0) A(\tau)\rangle - \langle A\rangle^2, \tag{A4.5}$$

where for Eq. A4.4 the process is assumed to be stationary and we define $\tau = t' - t$. The correlation function for this observable $A(t)$ is defined as

$$C\left(t,t'\right) = \left\langle \left(A\left(t'\right) - \langle A\rangle\right)\left(A(t) - \langle A\rangle\right)\right\rangle. \tag{A4.6}$$

$$C(\tau) = \langle (A(\tau) - \langle A\rangle)(A(0) - \langle A\rangle)\rangle \tag{A4.7}$$

$$= \langle A(0) A(\tau)\rangle - \langle A\rangle^2. \tag{A4.8}$$

Substitution of Eq. A4.8 into Eq. A4.5 leads to

$$\sigma^2_{A_T}(T) = \frac{1}{T^2}\int_0^T dt \int_{-t}^{T-t} d\tau C(\tau) + \frac{1}{T}\int_0^T \langle A\rangle^2 - \langle A\rangle^2 \tag{A4.9}$$

$$= \frac{1}{T^2}\int_0^T dt \int_{-t}^{T-t} d\tau C(\tau). \tag{A4.10}$$

In the limit of small T and large T a solution for Eq. A4.10 can be obtained

$$\sigma^2_{A_T}(0) = C(0) = \sigma^2_A. \tag{A4.11}$$

$$\lim_{T\to\infty} \approx \lim_{T\gg\tau_A} = \sigma^2_{A_T}(T \gg \tau_A) = \frac{1}{T^2}\int_0^T dt \int_{-\infty}^{\infty} d\tau C(\tau)$$

$$= \frac{2\sigma^2_A \tau_A}{T}, \tag{A4.12}$$

where we used the fact that $\lim_{\tau \gg \tau_A} C(\tau) = 0$. We introduced the correlation time τ_A in Eq. A4.12 which is commonly defined as

$$\tau_A \equiv \frac{1}{\sigma_A^2} \int_0^\infty C(\tau)\, d\tau. \tag{A4.13}$$

As an example, for a process $A(t)$ with exponential waiting times, $C(\tau) \propto e^{-k\tau}$ and $\int C(\tau)\, d\tau \propto k^{-1} \sim \tau_A$.

The power spectrum ($P_A(\omega)$) and correlation function are related through the Fourier Transform

$$C_A(\tau) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} d\omega P_A(\omega)\, e^{i\omega\tau}, \tag{A4.14}$$

$$P_A(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} d\tau C_A(\tau)\, e^{-i\omega\tau}, \tag{A4.15}$$

such that

$$C_A(0) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} d\omega P_A(\omega) = \sigma_A^2, \tag{A4.16}$$

$$P_A(0) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} d\tau C_A(\tau) \equiv 2\sigma_A^2 \tau_A = T\sigma_{A_T}^2, \tag{A4.17}$$

where τ_A is the correlation time (see Eq. A4.13).

Closely related to the correlation time is the relaxation function $\mathcal{A}(\tau)$ for a small perturbation close to equilibrium

$$\mathcal{A}(\tau) \equiv \frac{\langle A(\tau)\rangle_{\text{pert}} - \langle A\rangle}{\langle A(0)\rangle_{\text{pert}} - \langle A\rangle}, \tag{A4.18}$$

where it can be shown [128], following the fluctuation-dissipation theory, that the relaxation function is equal to the normalized correlation function

$$\mathcal{A}(\tau) = \frac{C(\tau)}{C(0)}. \tag{A4.19}$$

In other words, the average relaxation from an external perturbation close to equilibrium to the equilibrium distribution is indistinguishable from fluctuations in the equilibrium distribution. We thus have

$$\int d\tau \mathcal{A}(\tau) = \tau_{\mathcal{A}}. \tag{A4.20}$$

Next, we note the relation between the Laplace transform and the Fourier transform

$$\mathcal{F}\left(C_A\left(\tau\right)\right) = P_A\left(\omega\right) \tag{A4.21}$$
$$= \mathrm{Re}\left[\mathcal{L}\left(C_A\left(t\right)\right)\right] + \mathrm{Re}\left[\mathcal{L}\left(C_A\left(t\right)\right)^\dagger\right] \tag{A4.22}$$
$$= \mathrm{Re}\left[\hat{C}_A\left(s = i\omega\right)\right] + \mathrm{Re}\left[\hat{C}_A\left(s = -i\omega\right)\right] \tag{A4.23}$$
$$= 2\mathrm{Re}\left[\hat{C}_A\left(s = i\omega\right)\right]. \tag{A4.24}$$

The correlation time is therefore related to the Laplace transform of the correlation function by

$$\sigma_A^2 \tau_A = P_A\left(0\right) \tag{A4.25}$$
$$= \mathrm{Re}\left[\hat{C}\left(s = i\omega\right)\right]_{\omega=0} + \mathrm{Re}\left[\hat{C}\left(s = -i\omega\right)\right]_{\omega=0} \tag{A4.26}$$
$$= 2\mathrm{Re}\left[\hat{C}\left(s = i\omega\right)\right]_{\omega=0}. \tag{A4.27}$$

Lastly, we provide a straightforward relation between the integral of an arbitrary function and the Laplace transform of this function

$$\hat{A}\left(s = 0\right) = \int_0^\infty A\left(t\right). \tag{A4.28}$$

4.A.2 Derivation of Eq. 4.13

In this section we derive the correlation function and time for a binary switching process, where a single R particle is surrounded by multiple B particles. For a binary switching process

$$R + B \underset{k_d}{\overset{k_a}{\rightleftharpoons}} R^* \quad \text{with } \left(R, R^* = 0, 1\right), \tag{A4.29}$$

we can write the correlation function for the state n of A^* as

$$C_n\left(\tau\right) = p_*^0 \left(p_{*|*}\left(\tau\right) - p_*^0\right), \tag{A4.30}$$

where p_*^0 is the equilibrium probability for the complex to exist and $p_{*|*}\left(\tau\right)$ is the probability that the complex A^* exist at time $t = \tau$, given that it was present at time $t = 0$. For every two-state process we write

$$p_{*|*}\left(\tau\right) = 1 - p_{0|*}\left(\tau\right) \tag{A4.31}$$
$$= 1 - \mathscr{S}_{\mathrm{rev}}\left(\tau|*\right), \tag{A4.32}$$

where $\mathscr{S}_{\text{rev}}(\tau|*)$ is the survival probability for the particle R - the probability that A is free at $t=\tau$ given that it was bound at $t=0$. In Laplace space

$$\hat{p}_{*|*}(s) = s^{-1} - \hat{\mathscr{S}}_{\text{rev}}(s|*), \tag{A4.33}$$

such that we have

$$\hat{C}_n(s) = p_*^0\left(s^{-1} - \hat{\mathscr{S}}_{\text{rev}}(s|*) - \frac{p_*^0}{s}\right). \tag{A4.34}$$

The initial value theorem states that the $s\to\infty$ limit of $\hat{C}_n$ in the Laplace domain is equal to the $t\to 0$ limit of the $C(t)$ in the time domain, thus

$$\lim_{s\to\infty} s\hat{C}_n(s) = C_n(0) = \sigma_n^2. \tag{A4.35}$$

For a binary process the variance is

$$\sigma_n^2 = p_*^0\left(1-p_*^0\right) = \overline{n}(1-\overline{n}). \tag{A4.36}$$

From Eq. A4.34 and Eq. A4.36 it follows that $\lim_{s\to\infty} s\hat{\mathscr{S}}_{\text{rev}}(s|*) = 0$. This of course is true by definition of the survival probability, but it can be used as a sanity check in future expressions.

The Laplace transform of Eq. 4.10 leads to

$$\mathcal{L}\left[\mathscr{S}_{\text{rev}}(t|*) = k_d\int_0^t \left[1-\mathscr{S}_{\text{rev}}\left(t'|*\right)\right]\mathscr{S}_{\text{rad}}\left(t-t'|\sigma\right)dt'\right] \tag{A4.37}$$

$$\Rightarrow s\hat{\mathscr{S}}_{\text{rev}}(s|*) = \frac{k_d\hat{\mathscr{S}}_{\text{rad}}(s|\sigma)}{1+k_d\hat{\mathscr{S}}_{\text{rad}}(s|\sigma)}. \tag{A4.38}$$

For non-interacting B particles, we follow Agmon and Szabo [120] and approximate the survival probability as

$$\mathscr{S}_{\text{rad}}(t|\sigma) = \mathscr{S}_{\text{rad}}(t|\text{eq})\,S_{\text{rad}}(t|\sigma), \tag{A4.39}$$

where $S_{\text{rad}}(t|\sigma)$ is the survival probability for the geminate R-B pair at a distance a and $\mathscr{S}_{\text{rad}}(t|\text{eq})$ is the survival probability for R in a sea of equilibrated B particles. In the main text we elaborate on the assumptions underlying this approximation.

We rewrite the survival probability for a particle R surrounded by an equilibrated distribution of B-particles to the survival probability at contact. Using the detailed balance requirement for the process and the backward Smoluchovski equation we write [120]

$$k_{\text{rad}}(t) = k_a S_{\text{rad}}(t|\sigma). \tag{A4.40}$$

Using the exact relation between the survival probability and the rate coefficient (Eq. A4.41)

$$\mathscr{S}_{\text{rad}}\left(t|\text{eq}\right) = e^{-c\int_0^t k_{\text{rad}}(t')dt'}, \tag{A4.41}$$

and Eq. A4.40 we obtain for the time-derivative of $\mathscr{S}_{\text{rad}}\left(t|\text{eq}\right)$

$$\begin{aligned}
\frac{d\mathscr{S}_{\text{rad}}\left(t|\text{eq}\right)}{dt} &= -ck_{\text{rad}}\left(t\right)\mathscr{S}_{\text{rad}}\left(t|\text{eq}\right) && \text{(A4.42)}\\
&= -ck_a S_{\text{rad}}\left(t|\sigma\right)\mathscr{S}_{\text{rad}}\left(t|\text{eq}\right) && \text{(A4.43)}\\
&= -ck_a\mathscr{S}_{\text{rad}}\left(t|\sigma\right). && \text{(A4.44)}
\end{aligned}$$

The Laplace transform of Eq. A4.44 is

$$s\hat{\mathscr{S}}_{\text{rad}}\left(s|\text{eq}\right) - 1 = -ck_a\hat{\mathscr{S}}_{\text{rad}}\left(s|\sigma\right). \tag{A4.45}$$

We substitute the Laplace transform of Eq. A4.44 in Eq. A4.38 and combined with Eq. A4.33 we obtain

$$1 - s\hat{\mathscr{S}}_{\text{rev}}\left(s|*\right) = \frac{cK_{\text{eq}}}{1 + cK_{\text{eq}} - s\hat{\mathscr{S}}_{\text{rad}}\left(s|\text{eq}\right)}, \tag{A4.46}$$

where K_{eq} is the equilibrium constant k_a/k_d. Substituting this result in Eq. A4.34, we derive

$$\begin{aligned}
\hat{C}_n\left(s\right) &= \frac{\overline{n}}{s}\left(1 - \overline{n} - s\hat{\mathscr{S}}_{\text{rev}}\left(s|*\right)\right)\\
&= \frac{\overline{n}}{s}\left(1 - \overline{n} + \left(\frac{cK_{\text{eq}}}{1 + cK_{\text{eq}} - s\hat{\mathscr{S}}_{\text{rad}}\left(s|\text{eq}\right)} - 1\right)\right)\\
&= \frac{\sigma_n^2}{s}\left(1 - \frac{1 - s\hat{\mathscr{S}}_{\text{rad}}\left(s|\text{eq}\right)}{\left(1 - \overline{n}\right)\left(1 + cK_{\text{eq}} - s\hat{\mathscr{S}}_{\text{rad}}\left(s|\text{eq}\right)\right)}\right)\\
&= \sigma_n^2\frac{\overline{n}\hat{\mathscr{S}}_{\text{rad}}\left(s|\text{eq}\right)}{1 - \left(1 - \overline{n}\right)s\hat{\mathscr{S}}_{\text{rad}}\left(s|\text{eq}\right)}.
\end{aligned} \tag{A4.47}$$

To continue an expression for $\hat{\mathscr{S}}_{\text{rad}}\left(s|\text{eq}\right)$ is required. A general expression for the Laplace transform of Eq. A4.41 is not available. For small c, however, the Laplace transform can be approximated

$$\begin{aligned}
\mathscr{S}_{\text{rad}}\left(t|\text{eq}\right) &= e^{-c\int_0^t k_{\text{rad}}(t')dt'} && \text{(A4.48)}\\
&\approx 1 - c\int_0^t k_{\text{rad}}\left(t'\right)dt' + \frac{1}{2}\left(c\int_0^t k_{\text{rad}}\left(t'\right)dt'\right)^2 + \ldots && \text{(A4.49)}
\end{aligned}$$

This approximation in general is valid if the exponent is small, which is likely at short times, or if either a very small concentration c is present or if $k_a \ll 1, k_D \ll 1$ such that $k_{\text{on}} \ll 1$,

since $\lim_{t\to\infty} k_{\text{rad}} \approx k_{\text{on}}$. We note here that for longer times t, the expansion becomes less accurate. Using the assumption of a small concentration we Laplace transform Eq. A4.49

$$\hat{\mathscr{S}}_{\text{rad}}(s|\text{eq}) = s^{-1} - s^{-1}c\hat{k}_{\text{rad}}(s) + \dots \tag{A4.50}$$

$$\approx s^{-1}\left(1 + c\hat{k}_{\text{rad}}(s)\right)^{-1}, \tag{A4.51}$$

where a linearization of Eq. A4.51 is equal to Eq. A4.50. We substitute $\hat{k}_{\text{rad}}$ [120]

$$\hat{k}_{\text{rad}} = \frac{k_a k_D}{s}\frac{1+\tau(s)}{k_a + k_D(1+\tau(s))}, \tag{A4.52}$$

in Eq. A4.51

$$\lim_{c\to 0}\hat{\mathscr{S}}_{\text{rad}}(s|\text{eq}) \approx \frac{k_a + k_D(1+\tau(s))}{s(k_a + k_D(1+\tau(s))) + k_a c k_D(1+\tau(s))}, \tag{A4.53}$$

with $\tau(s) = \sigma\sqrt{s/D}$

We substitute Eq. A4.53 in Eq. A4.47 and take the limit $s = 0$, since this corresponds $2\text{Re}\left[\hat{C}_n(s=0)\right] = 2\sigma_n^2\tau_n$ (see Eq. A4.25)

$$\lim_{c\to 0}\hat{C}_n(s=0) = \sigma_n^2\overline{n}\hat{\mathscr{S}}_{\text{rad}}(0|\text{eq}) = \sigma_n^2\overline{n}\frac{k_a + k_D}{k_a c k_D} \tag{A4.54}$$

$$= \sigma_n^2\frac{k_a + k_D}{(k_a c + k_d)k_D} = \frac{P_A(\omega=0)}{2}, \tag{A4.55}$$

and the correlation time τ_n is

$$\tau_n = \frac{k_a + k_D}{(k_a c + k_d)k_D}. \tag{A4.56}$$

Chapter 5

Protein Logic

In recent years both experimentalists and theorists have begun to appreciate that cellular decision making can be performed at the level of a single molecule. Receptor proteins that bind two types of ligand can act as logic gates, initiating different cellular responses depending on which combination of ligands is present. Here we investigate the versatility of receptor function by applying a well known statistical mechanical model to receptors which bind either one or two ligands, and their associated dimers. Notably, we find that a single heterodimer can realize any of the 16 possible logic gates, including the XOR gate, by variation of biochemical parameters. We then investigate versatility by recombination, asking whether a set of only three receptors with fixed parameters can encode the four functionally unique logic gates (OR, AND, ANDNOT, and XOR) simply by forming the possible dimer combinations. An exhaustive search reveals that the simplest set (two single-ligand receptors and one double-ligand receptor) can realize several different groups of three gates, a result for which the previous analysis of single receptors and dimers provides a clear interpretation. Both results underscore the surprising functional freedom readily available to cells at the single-molecule level.

Based on manuscript W.H. de Ronde, P.R. ten Wolde and A.M. Mugler. Protein Logic *In preparation.*

5.1 Introduction

Cells depend on cues from their environment to initiate behaviors, including growth, division, differentiation, and death. For this purpose, cells have evolved surface receptors that detect environmental signals and relay them to their interior. Intracellular proteins respond to the states of these receptors, triggering a network of biochemical interactions that ultimately leads to a specific action. Although a particular environmental signal often elicits a particular cellular response, it is well established that signals can also act in combination [51, 129, 130, 131]. In this case the response triggered when two signals are present can be distinct from the responses triggered by each signal alone. The cell thereby acts as a logic gate, integrating two inputs to produce a single output. The performance of logical computations by cells has been studied in depth both experimentally, e.g. for engineered circuits [132, 133] and molecules [134, 135, 136], and theoretically, e.g. for biochemical networks [137] and gene regulation [138, 139, 140].

Remarkably, it is becoming clear that such logical computations can be encoded entirely within a single receptor molecule [84]. That is, while some receptors respond quite specifically to a single type of signaling molecule (or *ligand*), many respond to multiple ligands [141, 142, 143, 144]. Thus, the local concentrations of two or more ligands can act as the logical input, while the activity state of the receptor acts as the logical output. Moreover, many receptors exist in the form of dimers or larger oligomers. For example, G Protein-Coupled Receptors (GPCR) and ErbB receptors can each form dimers consisting of receptors of the same type (*homodimers*) or a receptor of each type (*heterodimers*) [141, 145, 146]. One might intuit (and indeed we will see) that a heterodimer, in particular, is well poised to act as a logical integrator of two ligand signals. Both of these mechanisms — multi-ligand specificity and dimerization — allow for the possibility that a cell encodes decisions directly at the detection level, even before further intracellular processing leads to the ultimate phenotypic response.

The focus of this study is to elucidate the functional capabilities available to a cell directly at the level of its receptors. We use a statistical mechanical model to investigate the versatility of logical functions attainable by single receptor monomers or dimers. Versatility is explored in two contexts. First, for a given monomer or dimer, we identify the set of logical functions that can be explored by variation of its biochemical parameters. Notably, we find that all 16 possible logic gates can be performed by a single dimer, a result that we support analytically. Second, motivated by the fact that receptors can diffuse, bind, and unbind on the cell surface [147], we ask whether a minimal set of receptor monomers with fixed parameters can perform a maximally distinct set of logical functions simply by forming the possible dimer combinations. We find several parameter settings at which dimers realize three out of four unique functions, a result for which the first analysis provides a clear interpretation. These results make clear that single receptor molecules can encode diverse functions that cells can exploit both on long evolutionary timescales, via parameter exploration, or on the short timescales of cell signaling, via diffusive recombination.

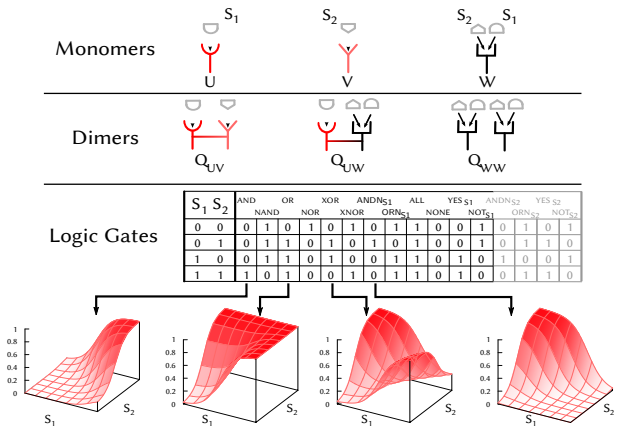


Figure 5.1: Setup. Top row: receptor monomers can bind ligand 1 (U), ligand 2 (V), or both competitively (W). Middle row: by dimerizing, receptors gain the capacity to function as logic gates. Bottom row: the table defines the 16 possible gates in terms of binary input and output; below, for the four functionally unique gates, we plot the continuous analogs given by the statistical mechanical model.

5.2 Methods

We study receptor function by appealing to an equilibrium statistical mechanical model. Statistical mechanical models have been used quite fruitfully in the study of many molecular biology problems, including receptor activity and gene regulation [148]. In the case of receptors, several models are well known. All assume that a receptor can exist in either an active (A) or an inactive (I) state, and that binding of a ligand changes the receptor bias for each state. In the Koshland-Nemethy-Filmer (KNF) model, ligand binding directly activates the receptor [149]. That is, the bias is complete: a ligand-bound receptor is active, and an unbound receptor is inactive. This condition is relaxed in the Monod-Wyman-Changeux (MWC) model, in which ligand-bound receptors can be in either state, but coupled receptors switch between states in synchrony [150]. Finally, in the conformational spread (CS) model [151], both conditions are relaxed: a ligand-bound receptor can be in either state, and coupled receptors can be in different states. Because we are interested in the minimal model that can capture the ability to perform logic gates, we adopt the MWC model; the KNF model prohibits certain logic gates by construction, while the CS model allows excess parametric freedom (clearly, what can be achieved by the MWC model can be achieved by the CS model). Furthermore, the MWC model has been shown to agree with experiments on receptors [87, 152].

The input in our model is the pair of concentrations $[S_1]$ and $[S_2]$ of two different ligands. The output is the probability for a receptor monomer or dimer to be in its active state. We consider three monomer types and the associated dimers (Fig. 5.1): a monomer that binds ligand 1 (U), one that binds ligand 2 (V), and one that binds both ligands (W). In the last case, ligand binding is competitive: there is only one binding pocket, so only one ligand type can

bind at a time. Noncompetitive binding, in which both ligand types can bind simultaneously, is captured by the Q_{UV} dimer (Fig. 5.1).

The probability p^A for a receptor to be in the active state is computed from the partition functions, which enumerate all possible ways a receptor can be in either the active (Z^A) or inactive (Z^I) state:

$$p^A = \frac{Z^A}{Z^I + Z^A}. \tag{5.1}$$

The explicit forms of the partition functions under the MWC model are presented as each monomer and dimer is discussed in the Results section. For intuition, we provide an example here: the partition functions for monomer W are

$$Z^A = \omega_0 \left(1 + q_1^A + q_2^A\right), \tag{5.2}$$

$$Z^I = 1 + q_1^I + q_2^I, \tag{5.3}$$

where $\omega_0 = e^{-E_0/k_BT}$ is the Boltzmann probability corresponding to the energy difference E_0 between the active and inactive state, and $q_i^j = [S_i]/K_i^j$ is the ratio of the concentration of ligand $i \in \{1, 2\}$ to the dissociation constant in activity state $j \in \{A, I\}$. In Eq. 5.2, the three terms correspond to the receptor being active when no ligand is bound, when ligand 1 is bound, and when ligand 2 is bound, respectively. The same holds for Eq. 5.3 with the receptor being inactive.

The dependence of p_A on $[S_1]$ and $[S_2]$ defines the receptor's function (Fig. 5.1, bottom row). Functions are categorized based on the idealized behavior prescribed by the 16 possible two-input binary logic gates (Fig. 5.1). Mathematically, the function approaches binary logic when the output is either minimal ($p^A \to 0$) or maximal ($p^A \to 1$) in each of the four states defined by each input being absent ($[S_i] = 0$) or present ($[S_i] > 0$). Numerically, when varying parameters to asses whether a receptor can realize a particular logic gate, we use an evolutionary algorithm based on the Wright-Fisher model [80, 81]. Briefly, the difference between p^A and the binary logic gate is evaluated over a grid of points and used to define a "fitness". Parameters for which p^A better approximates the logic gate yield a higher fitness, permitting optimization. A detailed description of the optimization procedure is provided in App. 5.A.1.

5.3 Results

First, we identify the logic gates that each receptor monomer and dimer can perform by parameter variation. Here, several derived analytic constraints support the numerical results. Then, we investigate the extent to which a set of monomers can access distinct logic gates by dimerizing. Several combinations of distinct gates are possible, a finding for which the first results provide a clear interpretation.

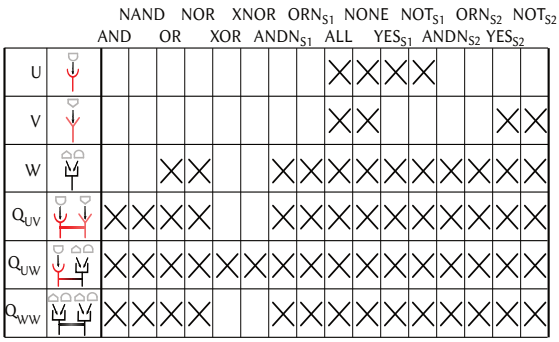


Figure 5.2: Functional versatility by parameter variation. For all monomers and dimers, we show the possible functions attainable by varying parameters. Attainability is assessed by numerical optimization and interpreted based on analytic constraints derived in the text.

5.3.1 Functions accessible by parameter variation

Figure 5.2 shows the set of logic gates that each monomer and dimer can perform, as determined by numerical optimization of model parameters. The most striking feature is that one of the dimers can perform all 16 possible gates. This and the other numerical results in Fig. 5.2 can be understood intuitively by appealing to analytic results derived from the underlying model, which we will describe in turn for each monomer and dimer.

Monomers

The first two monomers, receptors U and V, respond to only one input each. Therefore, they are trivially constrained to gates which depend on neither input (`ALL`, `NONE`) or on only one input ($\texttt{YES}_{S_i}$, $\texttt{NOT}_{S_i}$). Receptor W, on the other hand, allows competitive binding of both inputs, and can therefore realize several nontrivial gates.

At this point it is useful to observe that the gates exist in antagonistic pairs (a gate and its inverse), shown consecutively in Fig. 5.2: (`AND`, `NAND`), (`OR`, `NOR`), etc. Any receptor that can perform one member of a pair can perform the other, simply by inverting certain parameter values. Furthermore, several gates are equivalent under reversal of the two ligands (those with subscripts in Fig. 5.2): ($\texttt{ANDN}_{S_1}$, $\texttt{ANDN}_{S_2}$), ($\texttt{ORN}_{S_1}$, $\texttt{ORN}_{S_2}$), etc. Again, any receptor that can perform one of these can perform the other, simply by switching certain parameter values (corresponding to exchanging the effect of S_1 and S_2). Eliminating these redundancies, we arrive at four unique gates that respond nontrivially to both inputs:

$$\texttt{AND}, \quad \texttt{OR}, \quad \texttt{ANDN}_{S_1}, \quad \texttt{XOR}. \tag{5.4}$$

We will consider only these four unique gates from this point on.

The third monomer, receptor W, whose partition functions are given in Eqs. 5.2-5.3, can realize two of the four unique gates: `OR` and $\texttt{ANDN}_{S_1}$. The `OR`-gate follows straightforwardly from the situation where both ligands activate the receptor individually; their combination will then activate it as well. The $\texttt{ANDN}_{S_1}$-gate can be formed if ligand 1 binds more strongly

than ligand 2 ($q_1^j \gg q_2^j$), but ligand 1 only weakly biases the receptor toward the active state ($q_1^A \sim q_1^I$), while ligand 2 strongly biases it ($q_2^A \gg q_2^I$). In this scenario, a receptor that is inactive in the absence of both ligands ($\omega_0 \ll 1$) will only be active in the presence of ligand 2 *and not* 1.

Receptor W cannot realize the other two unique gates, `AND` and `XOR`. Both gates require a cooperative effect when both ligands are present: in the `AND`-gate, neither ligand activates the receptor individually, but both activate it together; in the `XOR`-gate, each ligand activates the receptor individually, but both suppress activation together. Such cooperative effects are not possible with competitive binding. As we will see next, dimerization is required to perform these gates.

Dimers

The three monomers admit six possible dimer combinations — three homodimers and three heterodimers. The homodimers Q_{UU} and Q_{VV} respond to only one input each and are therefore trivially constrained like monomers U and V. Moreover, heterodimers Q_{UW} and Q_{WV} are equivalent upon ligand exchange and can therefore realize equivalent sets of logic gates upon parameter variation. This leaves three dimers that can realize unique sets of logic gates upon parameter variation: Q_{UV}, Q_{UW}, and Q_{WW}.

The first dimer, receptor Q_{UV}, is the simplest heterodimer: it is formed by combining monomer U, which responds only to ligand 1, and monomer V, which responds only to ligand 2. Unlike receptor W, which is limited to competitive binding, the dimeric receptor Q_{UV} has two binding pockets and therefore allows noncompetitive (i.e. cooperative) binding. Accordingly, its partition functions extend those of receptor W (Eqs. 5.2-5.3) to include a cooperative term:

$$Z^A = \omega_0(1 + q_1^A + q_2^A + \overbrace{\omega_{12} q_1^A q_2^A}^{\text{cooperative}}), \tag{5.5}$$

$$Z^I = 1 + q_1^I + q_2^I + \omega_{12} q_1^I q_2^I, \tag{5.6}$$

The cooperative term contains an additional Boltzmann factor $\omega_{12} = e^{-E_{12}/k_B T}$ corresponding to the cooperative binding energy E_{12}, which could originate from, e.g., a conformational change of the receptor upon binding of one ligand that opens the binding pocket for the other ligand. For example, in the actin regulatory protein N-WASP, the binding affinity of each of its inputs, Cdc42 and PIP2, is increased by a factor of $\sim$400 when the other input is bound [130].

Receptor Q_{UV} can realize three of the four unique gates: `OR`, $\texttt{ANDN}_{S_1}$, and `AND`. The `OR` and $\texttt{ANDN}_{S_1}$ gates follow straightforwardly from the fact that Q_{UV} reduces to W for no cooperativity ($\omega_{12} = 0$), and receptor W can realize these gates as previously discussed. The `AND`-gate is formed when the receptor is inactive in the presence of each ligand alone but, due to the cooperative interaction, is active in the presence of both ligands together. Receptor Q_{UV} cannot realize the `XOR`-gate: if the receptor is activated by either one of the two ligands, it must also be activated by both ligands together. The cooperative interaction enhances the effect that each ligand individually has on the activation of the receptor, but it

cannot reverse it.

The intuition behind why receptor Q_{UV} can realize the AND-gate can be quantified by considering the constraints that an AND-gate places on the partition functions:

$$\begin{array}{ccc|l} S_1 & S_2 & p^A & \\ \hline 0 & 0 & 0 & \omega_0 \ll 1 \\ 1 & 0 & 0 & \omega_0 \left(1 + q_1^A\right) \ll 1 + q_1^I \\ 0 & 1 & 0 & \omega_0 \left(1 + q_2^A\right) \ll 1 + q_2^I \\ 1 & 1 & 1 & \omega_0 \left(1 + q_1^A + q_2^A + \omega_{12} q_1^A q_2^A\right) \gg 1 + q_1^I + q_2^I + \omega_{12} q_1^I q_2^I \end{array} \tag{5.7}$$

Here, we have recognized that a low output requires $Z^A \ll Z^I$ (see Eq. 5.1); therefore, the first three lines reflect that in an AND-gate the output is low in the first three input conditions. Similarly, a high output requires $Z^A \gg Z^I$, which is reflected in the last line. Receptor Q_{UV} can realize the AND-gate precisely because the constraints in Eq. 5.7 can be met simultaneously. For example, taking for illustration the simplifying case of intermediate cooperativity ($\omega_{12} \gtrsim 1$) and symmetric, saturating ligand concentrations ($q_1^j = q_2^j \gg 1$), Eq. 5.7 reduces to

$$1 \ll 1/\sqrt{\omega_0} \ll q_1^A/q_1^I \ll 1/\omega_0. \tag{5.8}$$

Indeed, we see that the AND-gate requires a bias upon ligand binding that is too weak to activate the receptor individually ($q_1^A/q_1^I \ll 1/\omega_0$), but strong enough to activate the receptor cooperatively ($q_1^A/q_1^I \gg 1/\sqrt{\omega_0}$). The strength of the cooperativity influences the quantitative properties of the AND-gate: an increase in ω_{12} shifts the transition region of the gate to smaller ligand concentrations, as indeed observed in studies of the AND-like N-WASP protein [130].

The second dimer, receptor Q_{UW}, is also a heterodimer: it is formed by combining monomer U, which responds only to ligand 1, and monomer W, which responds competitively to both ligands. The partition functions for this receptor are

$$Z^A = \omega_0\Big(1 + q_{1,U}^A + q_{1,W}^A + q_2^A + \omega_{11} q_{1,U}^A q_{1,W}^A + \omega_{12} q_{1,U}^A q_2^A\Big), \tag{5.9}$$

$$Z^I = 1 + q_{1,U}^I + q_{1,W}^I + q_2^I + \omega_{11} q_{1,U}^I q_{1,W}^I + \omega_{12} q_{1,U}^I q_2^I. \tag{5.10}$$

Here, since ligand 1 can bind to either monomer U or W, we distinguish these cases with the second subscript on q_1^j. There are now two cooperative terms, corresponding to the cases where monomers U and W bind, respectively, ligands 1 and 1 (ω_{11}), or ligands 1 and 2 (ω_{12}). Eqs. 5.9-5.10 make clear that receptor Q_{UW} reduces to receptor W (Eqs. 5.2-5.3) in the limit $q_{1,U}^j \to 0$, and to receptor Q_{UV} (Eqs. 5.5-5.6) in the limit $q_{1,W}^j \to 0$.

Receptor Q_{UW} can realize all four unique gates (and therefore all 16 possible gates; Fig. 5.2). The OR, ANDN_{S_1}, and AND gates follow straightforwardly from the fact that receptor Q_{UV}, which can realize these gates, is a limiting case. The XOR-gate is less trivial. Below we offer an intuitive argument for why receptor Q_{UW} can realize an XOR-gate, and in App. 5.A.2 we prove analytically that the output can be a nonmonotonic function of the two inputs for this receptor, which is required for an XOR-gate.

The XOR-gate is formed when each ligand individually activates the receptor by binding to monomer W, but when both ligands are present, ligand 1 is outcompeted and thus binds to monomer U, in turn suppressing activation. It is instructive here to describe this process in more detail. Suppose that ligand 1 promotes activation when bound to W but suppresses activation when bound to U. Further, suppose that ligand 1 binds more strongly to W than to U, such that in the presence of ligand 1 alone, the receptor is active. Now suppose that ligand 2 promotes activation when bound to W. Since ligand 2 can only bind to W, in the presence of ligand 2 alone, the receptor is also active. Finally, suppose that ligand 2 "interferes" with ligand 1, i.e. binds more strongly to W than ligand 1 does. Then, in the presence of both ligands, ligand 2 binds to W, leaving ligand 1 to bind to U. If U suppresses activation more strongly than W promotes activation, then in the presence of both ligands, the receptor is inactive. The resulting logic is the XOR-gate.

The third dimer, receptor Q_{WW}, is a homodimer: it is formed by combining two W monomers, each of which responds competitively to both ligands. The partition functions for this receptor are

$$Z^A = \omega_0\Big(1 + 2q_1^A + 2q_2^A + \omega_{11}q_1^Aq_1^A + 2\omega_{12}q_1^Aq_2^A + \omega_{22}q_2^Aq_2^A\Big), \tag{5.11}$$

$$Z^A = 1 + 2q_1^I + 2q_2^I + \omega_{11}q_1^Iq_1^I + 2\omega_{12}q_1^Iq_2^I + \omega_{22}q_2^Iq_2^I. \tag{5.12}$$

Here, the factors of two account for the fact that each ligand can be bound to either of two symmetric monomers. There are now three cooperative terms, corresponding to the cases where both monomers bind ligand 1 (ω_{11}), both bind ligand 2 (ω_{22}), or one binds ligand 1 and the other binds ligand 2 (ω_{12}).

Receptor Q_{WW} can realize three of the four unique gates: OR, ANDN_{S_1}, and AND. The OR and ANDN_{S_1} gates follow straightforwardly from the fact that each monomer alone can realize these gates as previously discussed. The AND-gate relies on strong suppression of cooperation between monomers if they are bound to the same ligand type (i.e. $\omega_{11} \to 0, \omega_{22} \to 0$); this suppression prevents activation when only one ligand is present. In fact, this limit reduces Eqs. 5.11-5.12 to Eqs. 5.5-5.6 (up to factors of 2), meaning the AND-gate constraint, Eq. 5.8, also holds here under the same conditions for which it was derived. Receptor Q_{WW} cannot realize the XOR-gate: because the individual monomers are identical, no negative interference is possible, as it is for receptor Q_{UW}.

5.3.2 Functions accessible by recombination

In the previous section we identified the logic gates accessible by individual receptors via variation of intrinsic biochemical parameters. In this section we ask a second question. Recognizing that receptors bind and unbind to the cell membrane, diffuse within the membrane,

and bind to each other (Fig. 5.3a), we here seek the logic gates that a set of monomers can realize — at fixed parameters — simply by forming the possible dimer combinations. This question is key to functional control at the single-molecule level: if diverse logic gates can be realized by a small set of monomers, cellular function could be strongly tuned by simply expressing a particular pair of monomers in abundance over the others. Alternatively, a signaling mechanism such as covalent modification could determine when a particular dimer was active, such that even if all dimer combinations were represented simultaneously on the membrane, only one would dominate the functional response of the cell. This question is also critically related to the challenge that all cells face: to encode reliable responses using limited resources (here, a limited set of monomers) and on short timescales (here, set by expression and diffusion).

The three monomers we study form four functional dimers: Q_{UV}, Q_{UW}, Q_{WV}, and Q_{WW} (the dimers Q_{UU} and Q_{VV} respond to only one input each and are neglected; additionally we neglect the monomer W itself although it is functional because its behavior is captured by Q_{WW}). This fact leads to the enticing question of whether there exist parameters at which the four dimers perform the four unique logic gates (Eq. 5.4). Such a finding would be highly nontrivial: all monomers are present in at least two dimers, and therefore the performance of a particular logic gate by one dimer places heavy constraints on the parameters of the other dimers.

An exhaustive search, in which we numerically optimize for each of the $4! = 24$ dimer-to-logic gate mappings in turn (App. 5.A.1), suggests that no parameter set exists at which all four unique logic gates are performed. Moreover, replacing any subset of gates with the corresponding inverse gates and repeating gives the same result in each of the $2^4 = 16$ cases. Interestingly, the result seems to be due to the fact that the parameters which support the `XOR`-gate in receptor Q_{UW} (or its counterpart Q_{WV}) prohibit the `AND`-gate in any of the other receptors. Next we support this numerical observation with an intuitive argument.

Suppose that receptor Q_{UW} performs an `XOR`-gate. As described in the previous section, the `XOR`-gate requires that when ligand 1 is present alone, it activates the receptor by binding to monomer W. Since the `AND`-gate requires the opposite behavior, namely that the receptor is inactive when ligand 1 is present alone, then the `AND`-gate cannot be formed by any receptor in which ligand 1 only binds to W. This group includes receptors Q_{WW} and Q_{WV}, leaving only receptor Q_{UV}. Then, as also described in the previous section, the `XOR`-gate requires that ligand 1 suppresses activation when bound to monomer U. Since the `AND`-gate requires that the receptor is active when both ligands are present, in receptor Q_{UV} this suppression would have to be overpowered by activation via ligand 2 binding to V. However, if this were the case, the receptor would surely be active in the presence of ligand 2 alone, which is inconsistent with the behavior of an `AND`-gate. These arguments make clear that if receptor Q_{UW} performs an `XOR`-gate, no other receptor can form an `AND`-gate. The same arguments, but with the ligands exchanged, hold if receptor Q_{WV} performs the `XOR`-gate instead of receptor Q_{UW}. Since receptors Q_{UW} and Q_{WV} are the only receptors that can perform the `XOR`-gate, we conclude that the `XOR` and `AND` gates are not mutually accessible by recombination of monomers U, V, and W at fixed parameters.

Even though all four unique logic gates cannot be performed at fixed parameters, we

do find six parameter sets at which unique groups of three logic gates are performed by three of the dimers. We denote these parameter sets as φ_k, for $k \in \{1, 2, \ldots, 6\}$, and show the logic gates and the dimers that perform them in Fig. 5.3b. We stress that this result is still nontrivial: two of the groups are performed by receptors Q_{WW}, Q_{UW}, and Q_{WV}, which all contain monomer W; additionally, two groups are performed by receptors Q_{UW}, Q_{WV}, and Q_{UV}, in which each monomer is represented in two of the three dimers. Due to the high degree of monomer overlap in both cases, one might have expected the three dimers to be constrained to similar functionality at fixed parameters; instead, we find that three unique logic gates can be formed. Further, Fig. 5.3b reveals that all four logic gates are represented among the six groups (but, as expected, never XOR and AND in the same group). Finally, the optimal solutions shown in Fig. 5.3b are robust to parametric perturbation: as shown in App. 5.A.3, for all φ_k, most random perturbations in which each parameter changes by an average of $\sim 20\%$ change the fitness of none of the three logic gates by more than 10%. All of these features underscore the functional versatility available to cells by dimeric recombination.

In the remainder of this section, we provide for parameter sets φ_1 and φ_2 the intuition behind how the three logic gates in Fig. 5.3b are performed by the corresponding receptors. In App. 5.A.4, we provide similar intuition for parameter sets φ_3, φ_4, φ_5, and φ_6. Furthermore, in App. 5.A.4, we argue why the groups observed in Fig. 5.3b (and their counterparts obtained upon ligand exchange) are the *only* groups of three unique logic gates that one expects to observe under this model.

Parameter set φ_1 (Fig. 5.3b, first row) corresponds to a case where each ligand only weakly promotes activation in the presence of monomer W. This feature allows receptor Q_{WW} to remain inactive when each ligand is present individually but become activated when both ligands are present together, forming the AND-gate. Furthermore, ligand-bound U both promotes activation and strongly enhances the binding of ligand 2 to W. This feature allows receptor Q_{UW} to perform the OR-gate: when ligand 1 is present alone, it promotes activation by binding to U; when ligand 2 is present in abundance and ligand 1 is present only in a small amount (and thus still in the "off" state), the small amount of ligand 1 is nonetheless sufficient to promote activation via enhanced binding of ligand 2 to W; and when both ligands are present in abundance, the two effects combine, resulting in activation. Finally, ligand-bound V both suppresses activation and strongly enhances the binding of ligand 1 to W. This feature allows receptor Q_{WV} to perform the $\texttt{ANDN}_{S_2}$ gate: when ligand 2 is present it suppresses activation via V, independent of ligand 1; but when ligand 1 and not (very much of) ligand 2 is present, the small amount of ligand 2 strongly enhances binding of ligand 1 to W, thus promoting activation.

Parameter set φ_2 (Fig. 5.3b, second row) corresponds to a case where ligand-bound W promotes activation. This feature is sufficient for receptor Q_{WW} to perform the OR-gate. Furthermore, ligand 2 binds more strongly to V than to W, and ligand-bound V suppresses activation more strongly than ligand-bound W promotes activation. These features allow receptor Q_{WV} to perform the $\texttt{ANDN}_{S_2}$ gate, since only in the presence of ligand 1 and not 2 will activation be promoted via W and not suppressed via V. Finally, (i) ligand 1 binds more strongly to W than to U, (ii) ligand 2 binds more strongly to W than ligand 1 does, and

(iii) ligand-bound U suppresses activation more strongly than ligand-bound W promotes activation. These are the precise features that allow receptor Q_{UW} to perform the XOR-gate, as outlined in detail in the previous section.

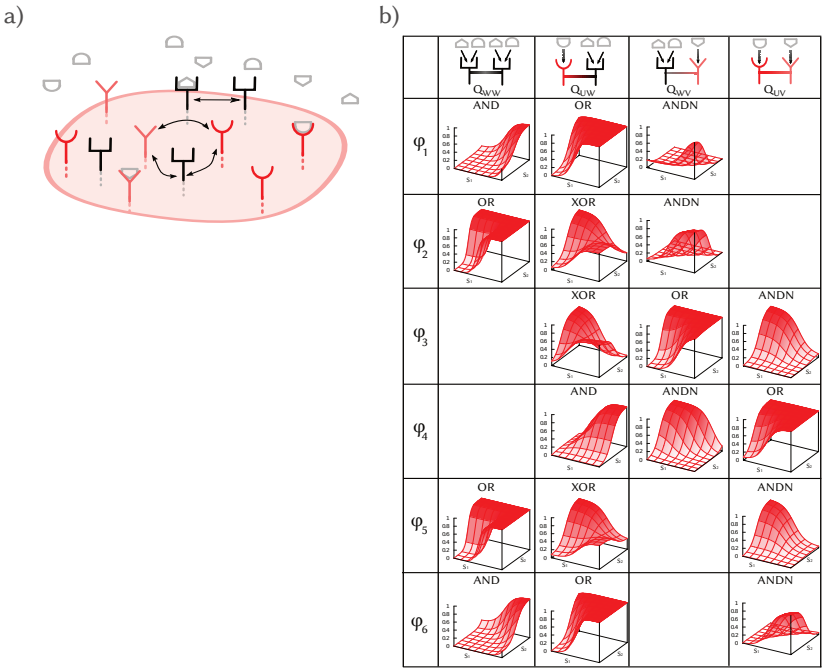


Figure 5.3: Functional versatility by recombination. **a)** Given the three monomer types, four functional units can be formed by receptors diffusing and dimerizing on the membrane (arrows). **b)** Four parameter sets φ_i are shown at which three of the four dimers perform functionally unique logic gates.

5.4 Discussion

We have used a statistical mechanical model to investigate the versatility of receptor function in two contexts: (i) the ability of a single receptor to access logical functions by parameter variation, and (ii) the ability — at fixed parameters — for a set of receptor monomers to access logical functions by dimerizing. The first context is arguably important on evolutionary timescales, on which mutations and environmental pressures act to change a cell's intrinsic biochemical parameters. The second context is more critical at short timescales, e.g. for cell signaling, during which gene expression and diffusion can potentially change cellular function at the molecular level by favoring the dimerization of particular receptors over others.

In the first context, we find that a single heterodimer, receptor Q_{UW}, can realize all 16 possible logic gates by parameter variation. Our analysis reveals that such complete func-

tional freedom, while perhaps surprising, is in fact quite intuitive for this receptor. In particular, receptor Q_{UW} performs the most challenging function, the XOR-gate, by exploiting an interference between the two ligands (i.e. when both ligands are present, one outcompetes the other for the activating binding pocket, ultimately causing suppression). Such a non-monotonic response requires competitive binding and asymmetric activation biases, both of which are possible by heterodimerization.

In the second context, we find that the simplest combination of monomers that yields four functional dimers cannot in fact perform the four unique logic gates at fixed parameters, an observation we explain by arguing that the AND-gate and the XOR-gate are not mutually accessible. Nonetheless, an exhaustive numerical search reveals that several distinct of groups of three unique gates are performable, a result that is nontrivial given the high degree of overlap among dimers' parameter spaces. We offer intuitive explanations for the emergence of these groups, and further, we argue that these groups are exhaustive. The ability to perform diverse functions with a limited set of simple components is of critical importance to the question of how cells encode reliable responses with limited resources.

We have adopted a minimal model (the MWC model) to describe a minimal set of components, and we have explored the functional capabilities available under these conditions. We are further encouraged by the fact that the MWC model has been shown to agree with experiments on receptors [87, 152]. Nonetheless, several extensions to the model or the study itself are natural choices for further exploration. First, the conformational spread (CS) model [151] generalizes the MWC model, and thus it would allow for more functional freedom in logic gate construction. However, it is always a concern that generalizing one's model can reduce the fraction of functional parameter space simply by increasing the total volume of parameter space.

Second, it would be straightforward to introduce one or more additional monomers when considering recombination. For example, introducing an additional monomer that binds a single ligand might in fact admit parameter sets at which all four unique logic gates are performed, at the expense of increasing the number of individual components that the cell must produce. The impact of such a finding, however, would be reduced by the fact that more than four functional dimer combinations would be possible. Third, it would also be straightforward to consider more complex dimers (or higher oligomers), such as $Q_{W_1W_2}$, in which each pocket binds both ligands competitively, but with asymmetric parameters. Of course, such increasing complexity would only be justified in the context of correspondingly detailed biological examples.

It is well established that receptors are responsive to multiple ligands. Interestingly, recent experiments have even exploited this fact to synthetically construct proteins that perform certain logic gates [84]. At the same time, receptor dynamics on the membrane — diffusion, binding, unbinding, dimerization — suggest that receptors can as functional signaling units by recombination. Indeed, experiments have shown that for many receptors, such as ErbB and GPCR, monomers combine to form different dimers that have different functionality [141, 145, 146]. It is our hope that this study will contribute to a predictive framework in which experiments like these can be interpreted and extended. The findings we report — that a single receptor can function as any logic gate and that a limited set

of monomers can access diverse logic gates by dimerizing — speak to the large degree of functional control available to cells at the level of individual receptor molecules.

5.5 Acknowledgements

I would like to thank Jose Alvarado for a critical reading of this chapter.

5.A Supplementary Information

5.A.1 Optimization details

In this section we describe in detail the numerical optimization algorithm. In Table. 5.1 we list the optimization parameters.

Single receptor

We use an evolutionary algorithm based upon the Wright-Fisher algorithm to obtain the parameters that correspond to a specific logic gate for every receptor type (Fig. 5.1). The Wright-Fisher algorithm models the evolution of a population. Evolution occurs in discrete, synchronous steps, and the population size remains constant. At each step, each member of the population produces offspring in proportion to its fitness. Then, mutations occur, and the mutated offspring comprise the population for the next step.

In our case, for a given receptor, we have a "population" of R initial parameter points φ_r. Each point has fitness f_r, and the total fitness for the receptor is $F = \sum_r f_r$. At each step, R new points ("offspring") are drawn from the distribution $p_r = f_r/F$, which weights each point by its fitness. Each new point is then "mutated" by multiplying a randomly selected parameter by the factor $(1+\delta)$, where δ is drawn uniform randomly from the range $[-\Delta : \Delta]$; we take $\Delta = 0.3$.

We define fitness as the agreement between the real-valued output of the statistical mechanical model, p^A and the binary output of a specific ideal logic gate. The ideal logic gate is prescribed by the goal function $G\left([\mathrm{S}_1], [\mathrm{S}_2]\right)$, which takes the value 0 or 1 depending on whether each input is "off" ($[S_i] < [S^*]$) or "on" ($[S_i] > [S^*]$), where $[S^*]$ sets the scale of ligand concentrations to which the receptor responds. We compare p^A and G over an $N \times N$ grid of input values, spaced logarithmically over the ranges of $[S_1]$ and $[S_2]$, which we take to be $[S_i]/[S^*] \in [10^{-2}\ 10^2]$. The fitness is thus

$$f_r = -\sum_{n,n'=1}^{N} \left| p^A\left([S_1^n], [S_2^{n'}]\right) - G\left([S_1^n], [S_2^{n'}]\right) \right|. \tag{A5.1}$$

The results in this article are obtained for $N = 4$. Taking $N = 2$ leads to suboptimal results, while taking different values of $N > 2$ yields similar results to $N = 4$. Similar results are also obtained for a fitness function with $N = 2$ and an additional central point at $[S_1] = [S_2] = [S^*]$, at which G is the average of the truth table for the gate.

The optimization parameters, as well as the bounds within which the model parameters are initialized and constrained during optimization, are given in Table. 5.1.

Multiple receptors

In the case where M different receptors (e.g. $\mathrm{Q_{UV}}$, $\mathrm{Q_{WV}}$, $\mathrm{Q_{UW}}$) are combined to act in different combinations as unique logic gates, the optimization algorithm follows a specific order in the optimization. A straightforward extension of the model for a single receptor is the optimization of three (or four) gates simultaneously, and taking as fitness $\mathcal{F}$ the summed

Model input	**Range**
$[\mathrm{S_i}]\,/[S^*]$	$\left[10^{-2} - 10^{2}\right]$
Model parameter	**Bounds**
$K^j_{i,k}/[S^*]$	$\left[10^{-4} - 10^{3}\right]$
ω_0	$\left[10^{-3} - 10^{3}\right]$
$\omega_{ii'}$	$\left[10^{-2} - 10^{2}\right]$
Optimization parameter	**Value**
Δ	0.3
N	4
R	50000
Steps	1000

Table 5.1: Overview of parameters. During optimization, model parameters are initialized and constrained within the indicated bounds

fitness of every gate F_m:

$$\mathcal{F} = \sum_{m=1}^{M} F_m. \tag{A5.2}$$

However, this optimization is not capable of optimizing all the gates independently. Instead, the algorithm optimizes either one (or two) gates, but then cannot optimize the third gate. To optimize the third gate, the already optimized gates decrease (temporarily) in fitness. This decrease is larger than the increase in fitness for the third gate and the algorithm finds suboptimal peaks in this rugged fitness landscape.

Instead of optimizing all gates simultaneously, we optimize gates in order. For the homodimer construction (Q_{WW}, Q_{UW}, Q_{WV}), we first optimize Q_{WW}, then Q_{UW}, where we only change the parameters of U, and then Q_{WV}, only changing V. The achieved results greatly outperform the results where we optimize all three gates simultaneously.

For the heterodimer construction (Q_{UW}, Q_{WV}, Q_{UV}), we again start by optimizing gate Q_{UW}, then the two gates Q_{UW} and Q_{WV} simultaneously, and finally Q_{UW}, Q_{WV}, and Q_{UV}. Again this procedure gives much better results than simultaneous optimization of all three gates.

For the construction with Q_{WW}, Q_{UW}, Q_{UV}), we start by optimizing gate Q_{WW}, then the two gates Q_{WW} and Q_{UW} simultaneously, and finally Q_{WW}, Q_{UW}, and Q_{UV}.

5.A.2 Formal proof for a XOR-gate for receptor Q_{UW}

The probability to be active $p^A\left([S_1], [S_2]\right)$ in an XOR-gate is a nonmonotonic function of $[S_1]$ and $[S_2]$ simultaneously. More specifically, for constant $[S_2] = [S_2^c]$, $p^A\left([S_1], [S_2^c]\right)$ is either monotonically increasing or decreasing, depending on the value $[S_2^c]$: for small $[S_2^c]$, p^A is monotonically increasing, while for large $[S_2^c]$, p^A is monotonically decreasing.

A positive derivative $\partial p^A/\partial[S_2]$ reflects a monotonically increasing function, while a negative derivative reflects a monotonically decreasing function. Therefore, in an XOR-gate, the derivative of the probability with respect to $[S_2]$ at constant $[S_1^c]$ should change sign as function of $[S_1^c]$. Again due to symmetry, the derivative of the probability with respect to $[S_1]$ at constant $[S_2^c]$ should change sign as function of $[S_2^c]$.

We will prove that the XOR-gate is possible for the Q_{UW} receptor even with $\omega_{11} = \omega_{21} = 1$. Recalling Eq. 1, the derivative can be written $\partial p^A/\partial[S_2] = f/(Z^A + Z^I)^2$, where the numerator

$$f\left([S_1^c], [S_2]\right) = \frac{\partial Z^A}{\partial [S_2]} Z^I - Z^A \frac{\partial Z^I}{\partial [S_2]} \tag{A5.3}$$

alone determines the sign. We therefore must show that f changes sign as function of $[S_1^c]$.

Specifically, the XOR-gate requires

$$f > 0 \text{ for } [\mathrm{S}_1^c] < [S_1^c]^*, \tag{A5.4}$$
$$f < 0 \text{ for } [\mathrm{S}_1^c] > [S_1^c]^*, \tag{A5.5}$$

for some $[S_1^c]^*$ and all $[S_2]$.

The partition functions Z^A and Z^I for the $\mathrm{Q_{UW}}$ receptor are given by Eqs. 9-10 with $q_i^j = [\mathrm{S_i}]/K_i^j$. Performing the derivatives in Eq. A5.3 reveals that f is a third order polynomial in $[\mathrm{S}_1^c]$ in which all dependence on $[S_2]$ drops out. Only one root is potentially positive:

$$[\mathrm{S}_1^c]^* = \frac{K_{1,W}^I K_{1,W}^A \left(K_2^I - K_2^A\right)}{K_{1,W}^I K_2^A - K_{1,W}^A K_2^I}. \tag{A5.6}$$

To satisfy Eqs. A5.4-A5.5, we require that the zeroth order term (the intercept) is positive and that the leading order term is negative; enforcing these conditions yields

$$K_2^I - K_2^A > 0, \tag{A5.7}$$
$$K_{1,W}^I K_2^A - K_{1,W}^A K_2^I > 0, \tag{A5.8}$$

which are in fact the precise conditions that maintain positivity of the root (Eq. A5.6). Parameters that satisfy these conditions enable the sign of $\partial p^A/\partial[S_2]$ to depend on constant $[S_1^c]$, which is one of the two conditions necessary to perform the XOR-gate. Notably, Eq. A5.7 directly shows that the binding of ligand 2 to the W monomer in $\mathrm{Q_{UW}}$ in the active state is less likely than binding in the inactive state.

The second requirement is that the sign of $\partial p^A/\partial[S_1]$ depends on constant $[S_2^c]$. Specifically, as in Eqs. A5.4-A5.5, the XOR-gate requires that the numerator $g([\mathrm{S}_1], [\mathrm{S}_2^c])$ of the derivative satisfies

$$g > 0 \text{ for } [\mathrm{S}_2^c] < [S_2^c]^*, \tag{A5.9}$$
$$g < 0 \text{ for } [\mathrm{S}_2^c] > [S_2^c]^*. \tag{A5.10}$$

for some $[S_2^c]^*$ and all $[S_1]$. Performing the derivative reveals that g is a second order polynomial whose coefficients depend on $[S_1]$. To satisfy Eqs. A5.9-A5.10, we again require that the intercept is positive and that the leading order term is negative; enforcing these conditions yields

$$h\left([S_1], K_{1,U}^I, K_{1,U}^A, K_{1,W}^I, K_{1,W}^A\right) > 0, \tag{A5.11}$$
$$K_{1,U}^I - K_{1,U}^A < 0, \tag{A5.12}$$

where the function h results straightforwardly from the derivative but is unwieldy, such that we do not reproduce it here. The roots of the polynomial $[S_2^c]^*$ are similarly unwieldy, but noting positivity requirements ($[S_2^c]^* > 0$, $K_{1,W}^I > 0$, $K_{1,U}^I > 0$, $K_2^I > 0$), parameter

regimes can be derived that satisfy both Eqs. A5.7-A5.8 and Eqs. A5.11-A5.12 simultaneously. As an example we present one possible regime here:

$$\frac{K^A_{1,W}}{K^I_{1,W}} < \frac{K^A_2}{K^I_2} < 1, \tag{A5.13}$$

$$0 < \frac{K^A_{1,W}}{K^I_{1,W}} \leq \frac{K^I_{1,U}}{K^I_{1,U} + K^I_{1,W}}, \tag{A5.14}$$

$$K^A_{1,W} + K^A_{1,U} > K^I_{1,U} + K^I_{1,W}. \tag{A5.15}$$

Eq. A5.13 states that the W monomer is activated by $[S_1]$ and $[S_2]$, and that activation by $[S_1]$ is stronger than activation by $[S_2]$. Note that for small concentrations of either $[S_1]$ or $[S_2]$, W is inactive. The two more interesting constraints are in Eq. A5.14 and Eq. A5.15. Eq. A5.15 states that $[S_1]$ bound to monomer U deactivates the receptor ($K^A_{1,U} > K^I_{1,U}$), since, from Eq. A5.13, we have seen that $K^A_{1,W} < K^I_{1,W}$. More importantly the deactivation of U by binding $[S_1]$ is stronger than the activation of W by $[S_1]$, and following Eq. A5.13, it is thus also stronger than activation of W by $[S_2]$. This is precisely the interference interaction as described in the main text. The last constraint, Eq. A5.14, provides the required binding strength of $[S_1]$ to U and W to satisfy all constraints. In the main text we argue that W should preferably bind $[S_2]$, such that in the presence of both ligand $[S_2]$ binds to W and $[S_1]$ binds to U with as result that Q_{UW} is inactive.

Here we have shown that the Q_{UW} receptor is capable of the nonmonotonic derivatives required by the XOR-gate. This capability is necessary but not sufficient to perform the gate, as an ideal logic gate requires that the output be maximally high and low at the appropriate input values. Our numerical results, however, indeed confirm that the Q_{UW} receptor can perform the XOR-gate.

5.A.3 Parameter sensitivity

In this section we discuss the sensitivity to parameter variation of the results at the six parameter sets φ_k shown in Fig. 5.3. Robustness against parameter fluctuations generally is considered an important quality of biochemical systems, due to stochastic nature of intra- and extracellular processes. If the observed logic gates only function within a very narrow parameter regime, this could lead to unreliable functioning.

Parameters are varied according to

$$\varphi^z_{\text{new}} = \varphi^z_{\text{old}} \left(1 + n^z\right) \tag{A5.16}$$

where n^z is the zth component of a uniformly distributed random vector $\mathbf{n}$ with norm $|\mathbf{n}| = \eta$. Under this implementation, η sets the average (root mean square) factor by which each parameter changes via $\langle \delta\varphi^z/\varphi^z \rangle = \eta/\sqrt{Z}$, where Z is the number of parameters. We sample 10^6 different vectors $\mathbf{n}$.

Sensitivity is measured by computing the fraction of new parameter sets for which, for

each individual gate m, the relative change in fitness is less than a factor λ:

$$\frac{\left|F_m^{\text{new}} - F_m^{\text{old}}\right|}{F_m^{\text{old}}} < \lambda \quad \forall\, m. \tag{A5.17}$$

Figure 5.4 reveals that for all φ_k, most random perturbations in which each parameter changes by an average of $\langle \delta\varphi^z/\varphi^z \rangle \sim 20\%$ change the fitness of none of the three logic gates by more than $\lambda = 10\%$.

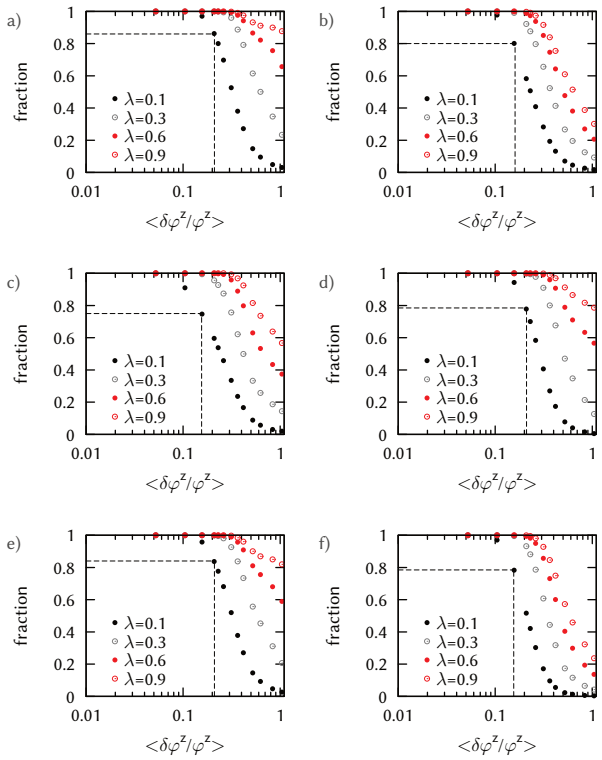


Figure 5.4: Robustness to parameter variation for the six parameter sets at which dimers can form three unique logic gates (Fig. 5.3): **a)** φ_1, **b)** φ_2, **c)** φ_3, **d)** φ_4, **e)** φ_5, and **f)** φ_6. An increase in $\langle \delta\varphi^z/\varphi^z \rangle$ reflects a larger range of parameter fluctuations and an increase in λ reflects a loosening of the robustness constraint. The dashed black lines indicate that a significant fraction of random perturbations in which each parameter changes by an average of $\langle \delta\varphi^z/\varphi^z \rangle \sim 20\%$ change the fitness of none of the three logic gates by more than $\lambda = 10\%$.

5.A.4 Figure 5.3b is exhaustive

Here, we argue why the groups observed in Fig. 5.3b (and their counterparts obtained upon ligand exchange) are the only groups of three unique logic gates that one expects to observe

under this model. The overall logic is presented first, with the arguments subsequently given in subsections.

There are 4 ways to choose a group of three from the four functional dimers Q_{WW}, Q_{UW}, Q_{WV}, and Q_{UV} to perform the three unique logic gates, namely: $\{Q_{WW}, Q_{UW}, Q_{WV}\}$, $\{Q_{UW}, Q_{WV}, Q_{UV}\}$, $\{Q_{WW}, Q_{UW}, Q_{UV}\}$, and $\{Q_{WW}, Q_{WV}, Q_{UV}\}$. The last two groups are symmetric upon ligand exchange; we therefore consider only the first three groups.

The first group is $\{Q_{WW}, Q_{UW}, Q_{WV}\}$. As shown in the main text, receptor Q_{WW} is capable of performing an AND-gate, an OR-gate, or an ANDN-gate, but not an XOR-gate (Fig. 5.2). If receptor Q_{WW} performs an AND-gate, receptor Q_{UW} can perform an ANDN-gate or an OR-gate, but not an XOR-gate (Argument 1). Receptor Q_{WV} then performs the OR-gate or the ANDN-gate, respectively (it also cannot perform an XOR-gate by the same argument). These two possibilities are represented by parameter set φ_1 (Fig. 5.3b) and its counterpart upon ligand exchange. If receptor Q_{WW} performs an OR-gate, receptor Q_{UW} can perform an ANDN-gate or an XOR-gate, but not an AND-gate (Argument 2). Receptor Q_{WV} then performs the XOR-gate or the ANDN-gate, respectively (it also cannot perform an AND-gate by the same argument). These two possibilities are represented by parameter set φ_2 (Fig. 5.3b) and its counterpart upon ligand exchange. If receptor Q_{WW} performs an ANDN-gate, three unique gates cannot be performed (Argument 3). Therefore, this group is exhaustively represented by φ_1 and φ_2.

The second group is $\{Q_{UW}, Q_{WV}, Q_{UV}\}$. As shown in the main text, receptor Q_{UV} is capable of performing an ANDN-gate, an OR-gate, or an AND-gate, but not an XOR-gate (Fig. 5.2). If receptor Q_{UV} performs an ANDN-gate, receptor Q_{UW} can perform an XOR-gate or an OR-gate, but not an AND-gate (Argument 4). Receptor Q_{WV} then performs the OR-gate or the XOR-gate, respectively (it also cannot perform an AND-gate by the same argument). These two possibilities are represented by parameter set φ_3 (Fig. 5.3b) and its counterpart upon ligand exchange. If receptor Q_{UV} performs an OR-gate, receptor Q_{UW} can perform an AND-gate or an ANDN-gate, but not an XOR-gate (Argument 5). Receptor Q_{WV} then performs the ANDN-gate or the AND-gate, respectively (it also cannot perform an XOR-gate by the same argument). These two possibilities are represented by parameter set φ_4 (Fig. 5.3b) and its counterpart upon ligand exchange. If receptor Q_{UV} performs an AND-gate, three unique gates cannot be performed (Argument 6). Therefore, this group is exhaustively represented by φ_3 and φ_4.

The third group is $\{Q_{WW}, Q_{UW}, Q_{UV}\}$. We note that this group is different from the first two groups, since it does not contain the two receptors Q_{UW} and Q_{UV} which are symmetric upon ligand exchange. Q_{UW} and Q_{WV} is not present in the receptors Q_{UW}, Q_{UV}. receptor Q_{UV}. To obtain a set which includes the XOR-gate, this gate has to be accessed by receptor Q_{UW}. As shown in the main text, receptor Q_{WW} is capable of performing an AND-gate, an OR-gate, or an ANDN-gate, but not an XOR-gate (Fig. 5.2). If receptor Q_{WW} performs an AND-gate, receptor Q_{UW} can perform an ANDN-gate or an OR-gate, but not an XOR-gate (Argument 1). If receptor Q_{UW} performs an ANDN-gate, receptor Q_{UV} cannot perform an OR-gate (Argument 7); since receptor Q_{UV} also cannot perform an XOR-gate (Fig. 5.2), three unique gates cannot be performed. Therefore, receptor Q_{UW} must perform an OR-gate, leaving receptor Q_{UV} to perform an ANDN-gate. This possibility is represented by parameter set φ_6 (Fig. 5.3b).

If receptor Q_{WW} performs an OR-gate, receptor Q_{UW} can perform an ANDN-gate or an XOR-gate, but not an AND-gate (Argument 2). If receptor Q_{UW} performs as an ANDN-gate, receptor Q_{UV} cannot perform an AND-gate (Argument 8); since receptor Q_{UV} also cannot perform an XOR-gate, three unique gates cannot be performed. Therefore, receptor Q_{UW} must perform an XOR-gate, leaving receptor Q_{UV} to perform an ANDN-gate. This possibility is represented by parameter set φ_6 (Fig. 5.3b). If receptor Q_{WW} performs as an ANDN-gate, three unique gates can not be performed (Argument 9). Therefore, this group is exhaustively represented by φ_5 and φ_6.

This completes the logic arguing that the groups observed in Fig. 5.3b are exhaustive.

Argument 1

If receptor Q_{WW} performs an AND-gate, ligand 2 alone does not promote activation when binding to monomer W. Therefore, because ligand 2 does not bind to monomer U, the receptor Q_{UW} is always inactive if ligand 2 is present alone. This behavior is inconsistent with the logic of the XOR-gate.

Argument 2

If receptor Q_{WW} performs an OR-gate, ligand 2 alone promotes activation when binding to monomer W. Therefore, because ligand 2 does not bind to monomer U, the receptor Q_{UW} is always active if ligand 2 is present alone. This behavior is inconsistent with the logic of the AND-gate.

Argument 3

If receptor Q_{WW} performs an ANDN-gate, receptors Q_{UW} and Q_{WV} can each perform neither an XOR-gate nor an AND-gate, thereby preventing the group $\{Q_{WW}, Q_{UW}, Q_{WV}\}$ from performing three unique gates. The reason is straightforward: if receptor Q_{WW} performs an ANDN-gate, one ligand must suppress activation via W while the other ligand promotes activation via W. This feature immediately excludes the XOR-gate since, as described in the main text, an XOR-gate requires both ligands to promote activation via W. This feature also excludes an AND-gate since, as also described in the main text, an AND-gate requires either that activation via W is promoted only weakly or that both ligands suppress activation via W. In the first case, activation of receptor Q_{UW} (or Q_{WV}) is only achieved cooperatively when both ligands are present. In the second case, activation is achieved with both ligands present via U (or V) due to an interference effect similar to that underlying the XOR-gate (see discussion of parameter set φ_4 above).

Argument 4

If receptor Q_{UV} performs an ANDN_{S_1}-gate, Q_{UW} cannot function as an AND-gate. To function as an AND-gate (see Eq. 9), this requires that $\omega_0 q_{1,U}^A q_2^A \gg q_{1,U}^I q_2^I$, while $\omega_0 q_{1,U}^A \ll q_{1,U}^I$ and $\omega_0 q_2^A \ll q_2^I$. This conditions cannot be satisfied simultaneously.

Argument 5

If receptor Q_{UV} performs an OR-gate, ligand 1 activates the receptor via U. However, for receptor Q_{UW} to perform the XOR-gate, ligand 1 must suppress activation via U, as described

in the main text.

Argument 6

If Q_{UV} functions as an AND-gate U is activated by S_1 and V is activated by S_2, but both activation biases alone are insufficient to activate the receptor. This excludes the formation of a XOR-gate for either the Q_{UW} or Q_{WV}. As we have discussed in the previous section, the XOR-gate is obtained by the deactivation of U (V) by ligand S_1 (S_2). However, it is possible that Q_{UW} is a OR-gate, while Q_{WV} is a ANDN-gate. The Q_{UW}-OR-gate requires that W is activated by S_2 and S_1, since monomer U is not active in the presence of S_1. The Q_{WV}-ANDN-gate requires that W is strongly deactivated by S_1. These two conditions on W are mutually exclusive.

Argument 7

If receptors Q_{WW} and Q_{UW} perform an AND-gate and an ANDN-gate, respectively, the ANDN-gate must be $\texttt{ANDN}_{S_1}$, not $\texttt{ANDN}_{S_2}$. The reason is that the AND-gate requires ligand 2 to promote activation via W, while the $\texttt{ANDN}_{S_2}$ gate requires ligand 2 to suppress activation via W. Then, if Q_{UW} indeed performs the $\texttt{ANDN}_{S_1}$ gate, receptor Q_{UV} cannot perform an OR-gate. The reason is that the AND and $\texttt{ANDN}_{S_1}$ gates require ligand 1 to suppress activation via U and not via W, while the OR-gate requires ligand 1 to promote activation via U. deactivated by S_1. Therefor receptor Q_{UV} can not function as an OR-gate.

Argument 8

If receptors Q_{WW} and Q_{UW} perform an OR-gate and an ANDN-gate, respectively, the ANDN-gate must be $\texttt{ANDN}_{S_1}$, not $\texttt{ANDN}_{S_2}$. The reason is that the OR-gate requires ligand 2 to promote activation via W, while the $\texttt{ANDN}_{S_2}$ gate requires ligand 2 to suppress activation via W. Then, if Q_{UW} indeed performs the $\texttt{ANDN}_{S_1}$ gate, receptor Q_{UV} cannot perform an AND-gate. The reason is that the OR and $\texttt{ANDN}_{S_1}$ gates require ligand 1 to suppress activation via U and not via W, while the AND-gate requires ligand 1 to promote activation via U. de-activates U. As a

Argument 9

If receptor Q_{WW} performs an ANDN-gate, receptor Q_{UW} cannot perform a XOR-gate gate, since this requires that both ligands activate W. If receptor Q_{WW} performs an $\texttt{ANDN}_{S_1}$ gate, receptor Q_{UW} cannot perform an AND-gate, since Q_{UW} is active if only ligand 2 is present. If receptor Q_{WW} performs an $\texttt{ANDN}_{S_1}$ gate, receptor Q_{UW} can perform an OR-gate if ligand 1 activates U more strongly than it deactivates W. However, receptor Q_{UV} is then always active if ligand 1 is present, and this is inconsistent with the logic of the AND-gate. If receptor Q_{WW} performs an $\texttt{ANDN}_{S_2}$ gate, receptor Q_{UW} cannot perform an AND-gate, since Q_{UW} is active if only ligand 1 is present (ligand 1 activates U) or Q_{UW} is never active (ligand 1 deactivates U more strongly than it activates W). If receptor Q_{WW} performs an $\texttt{ANDN}_{S_2}$ gate, receptor Q_{UW} can perform an OR-gate, if (i) ligand 1 activates U and (ii) in the presence of small ligand 1 and an abundance of ligand 2 the receptor Q_{UW} is active. However, receptor Q_{UV} is then always active if ligand 1 is present and this is inconsistent with the logic of the AND-gate.

Chapter 6

Reliability of frequency- and amplitude-decoding in gene regulation

In biochemical signaling, information is often encoded in oscillatory signals. However, the advantages of such a coding strategy over an amplitude encoding scheme of constant signals remain unclear. Here we study the dynamics of a simple model gene promoter in response to oscillating and constant transcription factor signals. We find that in biologically-relevant parameter regimes an oscillating input can produce a more constant protein level than a constant input. Our results suggest that oscillating signals may be used to minimize noise in gene regulation.

Based on manuscript F. Tostevin, W.H. de Ronde and P.R. ten Wolde (**2012**). Reliability of frequency- and amplitude-decoding in gene regulation *Physical Review Letters* **108**.
doi:10.1103/PhysRevLett.108.108104

6.1 Introduction

Cells are constantly exposed to a range of environmental stimuli to which they must respond reliably. In recent years, it has become increasingly clear that cells use complex encoding strategies to represent information about the environment in the temporal dynamics of intracellular components [153]. In particular, oscillatory or pulsatile signals are commonly found in signaling and gene regulatory networks [154]. Perhaps the best known example is the phenomenon of calcium oscillations [155]. Oscillatory dynamics have also been observed at the level of gene regulation in nuclear localization of signaling proteins [156] and transcription factors [157, 158, 159] and in transcription factor expression [160]. Yet the advantages of such a coding strategy for signal transmission remain unclear.

A number of possible advantages for oscillatory signals have been suggested. Oscillatory signals minimize the prolonged exposure to high levels of calcium, which can be toxic for cells [161]. In systems with cooperativity [162] an oscillating signal effectively reduces the signal threshold for response activation. Pulsed signals also provide a way of controlling the relative expression of different genes [159]. However, these studies have ignored the impact of biochemical noise on the reliability of signal transmission. Encoding of signals in protein oscillations may play a direct role in ensuring accuracy in intracellular signaling.

Variability in the cellular response to an external signal will arise from the temporal pattern of network activation and from inevitable biochemical noise in the reactions making up the processing network, both of which will depend on the coding strategy employed. A reliable response requires minimization of such variability. Encoding of stimuli into oscillatory signals can reduce the impact of noise in the input signal and during signal propagation [163]. However, it remains unclear whether oscillatory signals can also be *decoded* with a similar fidelity to constant signals – one might expect that the inherent variability of an oscillatory signal would inevitably lead to a more variable response. In this chapter we investigate whether oscillatory signals can be reliably decoded in a simple model of gene regulation. Surprisingly, we find that in biologically-relevant parameter regimes an oscillating input can lead to a more constant output protein level than a constant input. This effect arises from differences in promoter state fluctuations, which it has recently been shown can be a dominant source of noise in vivo [164].

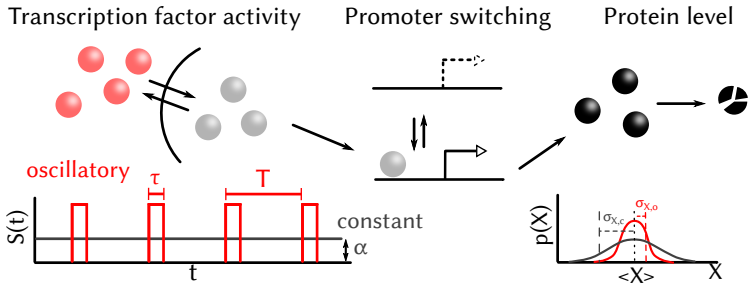


Figure 6.1: A gene promoter driven by constant and oscillating signals, s(t). We compare the variances of X for the two input signals, $\sigma^2_{X,c}$ and $\sigma^2_{X,o}$, at the same mean level $\langle X \rangle$.

6.2 Model

We focus on the simple regulatory motif (Fig. 6.1) of a single gene promoter which can be in an active (P^*) or inactive (P) state. The input to the system is the activity $s(t)$ of a transcription factor which enhances activation of the promoter. A protein X, whose level $X(t)$ constitutes the network output, is transcribed from the active promoter; proteins also spontaneously decay. The system therefore consists of the reactions

$$\mathsf{P} \underset{\lambda}{\overset{\kappa s(t)}{\rightleftharpoons}} \mathsf{P}^*, \tag{6.1}$$

$$\mathsf{P}^* \xrightarrow{\rho} \mathsf{P}^* + \mathsf{X}, \tag{6.2}$$

$$\mathsf{X} \xrightarrow{\mu} \emptyset. \tag{6.3}$$

We will consider two forms for $s(t)$. In the first instance we take $s(t) = \alpha$ to be constant. In the second, $s(t)$ is oscillatory, reflecting the action of a periodic upstream signal or the inherently oscillatory dynamics of the transcription factor. Since transcription factor pulses often resemble distinct sharp peaks [157, 158, 159] we take $s(t)$ to be a binary switching process, with $s(t) = 1$ for $nT \leq t < nT + \tau$ ($n \in \mathbb{Z}$) and $s(t) = 0$ otherwise (see Fig. 6.1). This form also has the advantage of making the dynamics analytically tractable. We characterize the system response in terms of the means $\langle X \rangle, \langle P^* \rangle$ and variances $\sigma^2_X, \sigma^2_{P^*}$; for the oscillating input we define the stationary mean $\langle X \rangle = T^{-1} \int_0^T \mathbb{E}\left[X(t)\right] dt$, where $\mathbb{E}\left[\right]$ denotes averaging over network realizations with the same input, and the stationary variance $\sigma^2_X = T^{-1} \int_0^T \mathbb{E}\left[X(t)^2\right] dt - \langle X \rangle^2$. These are calculated from the chemical master equation [68], and verified by stochastic simulations [74].

When driven by a constant signal the promoter simply undergoes random switching with constant rates. Hence the probability of the promoter being active at any time is $\langle P^* \rangle_{\mathrm{c}} = \alpha\kappa/(\alpha\kappa + \lambda)$. For an oscillating input the average activity,

$$\langle P^* \rangle_{\mathrm{o}} = \frac{\kappa}{\kappa + \lambda} \left[\frac{\tau}{T} + \frac{\kappa(1 - e^{-(\kappa+\lambda)\tau})(1 - e^{-\lambda(T-\tau)})}{\lambda T(\kappa + \lambda)(1 - e^{-\kappa\tau - \lambda T})} \right], \tag{6.4}$$

contains a term from the expected promoter activity when $s = 1$ multiplied by the fraction of time, τ/T, for which $s(t) = 1$, and a correction due to the fact that the promoter response timescales when $s(t) = 1$ and $s(t) = 0$ ($[\kappa + \lambda]^{-1}$ and λ^{-1}, respectively) differ. Since promoter switching is a two-state process the variance in promoter activity is determined by the mean, $\sigma^2_{P^*} = \langle P^* \rangle(1 - \langle P^* \rangle)$. The mean protein level also has the same form for both input signals, $\langle X \rangle = \rho\langle P^* \rangle/\mu$. However, differences in the timing of protein production will mean that the variance in the protein level differs between the two signals.

For a constant signal the variance

$$\sigma^2_{X,\mathrm{c}} = \langle X\rangle_\mathrm{c}\left[1+\frac{\rho\lambda}{(\alpha\kappa+\lambda)(\alpha\kappa+\lambda+\mu)}\right] \tag{6.5}$$

consists of an intrinsic Poissonian term due to randomness in protein production, and an extrinsic contribution from fluctuations in the promoter state. The variance given an oscillatory input $\sigma^2_{X,\mathrm{o}}$ can similarly be derived, but the full expression is unwieldy and thus not presented here. In the following we will compare $\sigma^2_{X,\mathrm{o}}$ with $\sigma^2_{X,\mathrm{c}}$ at the same mean response level, achieved by choosing the level of the constant signal $\alpha = \alpha(\tau, T, \kappa, \lambda)$ such that $\langle P^*\rangle_\mathrm{c} = \langle P^*\rangle_\mathrm{o} = \langle P^*\rangle$, and hence also $\langle X\rangle_\mathrm{c} = \langle X\rangle_\mathrm{o} = \langle X\rangle$.

6.3 Results

It is instructive to first consider cases in which the two input signals lead to similar distributions for X. First, if $\rho, \mu \to \infty$ with $X_\infty = \rho/\mu$ held constant, whenever the promoter is inactive the protein level is $X(t) = 0$, while when the promoter is active the protein level will be Poisson-distributed with mean X_∞. The variance in X is then given by $\sigma^2_{X,\mathrm{c}} = \sigma^2_{X,\mathrm{o}} = \langle X\rangle[1 + X_\infty - \langle X\rangle]$. In this limit of fast protein dynamics the precise pattern of promoter switching does not affect the variance in protein expression. At the other extreme, if either promoter switching or protein production is slow compared to the oscillating input ($\kappa, \lambda \ll T^{-1}$ or $\rho, \mu \ll T^{-1}$), the slow reactions effectively integrate over the temporal variation of the input. Since the network dynamics is too slow to reliably respond to the oscillating signal, the protein response is equivalent to that for the constant input.

Differences appear between the variances for the oscillatory and constant signals in the biologically most important regime of intermediate parameter values, where the promoter is able to respond to the oscillating input signal, protein production can react to switching of the promoter, and the protein lifetime is sufficiently long that patterns of promoter activity are important. Figures 6.2a-c show that there exist regions in which either the constant or oscillating signal leads to smaller fluctuations σ^2_X over large parameter ranges. Here we consider parameter ranges representative of eukaryotic cells, in which promoter switching occurs on a timescale of minutes to hours [164, 165] and protein lifetimes are in the range of a few to hundreds of hours [166]. The timescale of the input signal is chosen to be representative of NF-κB oscillations [157]. However, we emphasize that out results are general and would apply equally to prokaryotic cells in an appropriate parameter regime.

To understand the noise properties for the two input signals we consider the network response in the frequency domain. The dynamics of $X(t)$ can be described by the Langevin equation

$$\frac{dX}{dt} = \rho P^*(t) - \mu X(t) + \eta(t), \tag{6.6}$$

where $\eta(t)$ represents Gaussian white noise with [97]

$$\langle \eta(t) \rangle = 0 \text{ and } \left\langle \eta(t)\eta(t') \right\rangle = \left(\rho \mathbb{E}[P^*(t)] + \mu \mathbb{E}[X(t)]\right) \delta\left(t - t'\right).$$

Using the spectral addition rule [96] it is straightforward to calculate from Eq. 6.6 the power spectrum of fluctuations in $X(t)$, $S_X(\omega) = \left[\rho^2 S_P(\omega) + S_\eta(\omega)\right] / (\omega^2 + \mu^2)$, in terms of the spectra of promoter fluctuations $S_P(\omega)$ and of intrinsic noise in the production and decay of X, $S_\eta(\omega)$. The variance σ_X^2 can be found by integrating $S_X(\omega)$ over all frequencies, and can therefore be written as the sum of intrinsic and extrinsic terms, $\sigma_X^2 = \sigma_{\text{ex}}^2 + \sigma_{\text{in}}^2$ with

$$\sigma_{\text{ex}}^2 = \int_0^\infty g^2(\omega) S_P(\omega) d\omega, \tag{6.7}$$

$$\sigma_{\text{in}}^2 = \int_0^\infty \frac{S_\eta(\omega)}{\omega^2 + \mu^2} d\omega, \tag{6.8}$$

where $g^2(\omega) = \rho^2/(\omega^2 + \mu^2)$ is a signal-independent gain factor. The intrinsic noise $\sigma_{\text{in}}^2 = \langle X \rangle$ is also independent of the input signal. Any differences between $\sigma_{X,\text{o}}^2$ and $\sigma_{X,\text{c}}^2$ must, therefore, arise from differences in $S_P(\omega)$.

For a constant input, $S_{P,\text{c}}(\omega) = 2\sigma_{P*}^2 \tau_P / (\omega^2 \tau_P^2 + 1)$ (Fig. 6.2d, black line) has a simple Lorentzian form due to random promoter switching, with $\tau_P = (\alpha\kappa + \lambda)^{-1}$ the switching correlation time. A general expression for $S_{P,\text{o}}(\omega)$ with an oscillating input is more difficult to calculate. However, simulations show (Fig. 6.2d, red line) that there are two significant components. First, there are sharp peaks at frequencies corresponding to the signal period $\omega_T = 2\pi/T$ and multiples thereof, reflecting systematic changes in $\mathbb{E}[P^*(t)]$ due to the periodicity of $s(t)$. Second, $S_{P,\text{o}}(\omega)$ includes an approximately Lorentzian noise background associated with random (Poissonian) switching of the promoter when $s(t) = 1$ and delays in deactivation when $s(t) = 0$.

The protein lifetime $\tau_{\text{X}} = \mu^{-1}$ relative to the signal period T is particularly important in determining which signal minimizes the output noise. The constant signal typically leads to smaller fluctuations when $\tau_{\text{X}} < T$ (see Fig. 6.2a). In this regime, for the oscillatory signal σ_{ex}^2 is dominated by contributions from the peaks of $S_{P,\text{o}}(\omega)$ appearing at $\omega = n\omega_T$, since $\mu \gtrsim \omega_T$ and the gain $g^2(\omega)$ is large for frequencies $\omega < \mu$ (Fig. 6.2d, dashed line). Consequently, $X(t)$ features large production bursts when the signal is "on" with most proteins decaying before the next input pulse, while for a constant signal production and decay are more regularly distributed (see Fig. 6.2b).

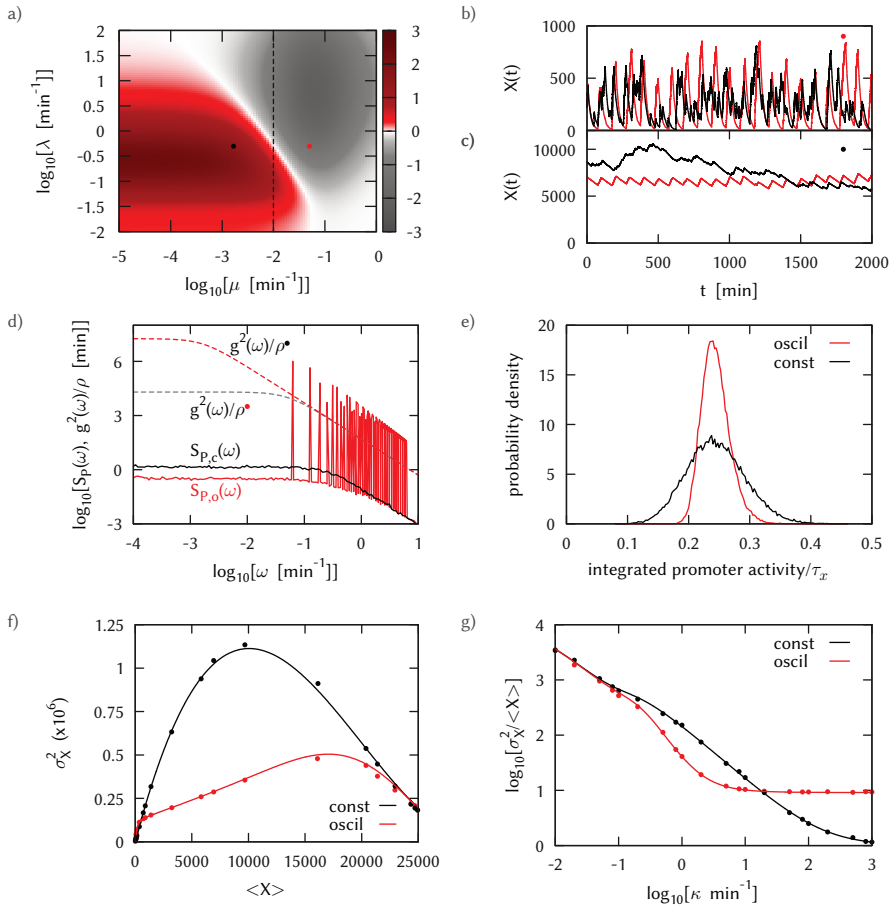


Figure 6.2: **a)** Relative noise for oscillating and constant signals, $\log_2(\sigma^2_{X,o}/\sigma^2_{X,c})$, as the protein degradation rate μ and promoter deactivation rate λ are varied. In red (black) regions the oscillatory (constant) signal leads to lower noise. Other parameters are: τ=25 min, T=100 min, κ=1 min^{-1}, ρ=50 min^{-1}. The dotted line indicates T^{-1}. **b,c)** Typical time-series of X(t) for the parameter combinations denoted in **a)** by red ∘ **b)**: λ=0.2 min^{-1},μ=0.05 min^{-1}) and black ∘ **c)**: λ=0.2 min^{-1},μ=0.00167 min^{-1}). Unless otherwise noted, parameters as in **c)** are used in panels **(d-g)**. **d)** The promoter power spectrum $S_P(\omega)$ and gain $g^2(\omega)=\rho^2/(\omega^2+\mu^2)$. For both inputs there is a noise background due to randomness in promoter switching. Peaks appear in $S_{P,o}(\omega)$ at $\omega=n\omega_T$ due to the periodicity of the oscillating input signal. **e)** Distribution of the fraction of time the promoter is active in an interval τ_X, calculated from stochastic simulations. The oscillatory signal leads to more reproducible promoter activity on the timescale τ_X. **f)** Variance against mean output level as T is varied with τ held constant. The oscillating signal can achieve a lower variance over nearly the full range of output levels. In **f)** and **g)**, lines show exact analytic results, points show results of stochastic simulations. **g)** Fano factor $\sigma^2_X/\langle X\rangle$ as κ is varied with constant K=λ/κ=0.2. For slow switching $\kappa\tau\lesssim 1$ the two signals become equivalent. For extremely fast switching the constant signal minimizes σ^2_X. At intermediate switching rates, 0.1 min$^{-1}\lesssim\kappa\lesssim$18 min^{-1}, the oscillating signal allows for smaller σ^2_X.

For long protein lifetimes $\tau_{\mathrm{X}} > T$, $\sigma^2_{X,\mathrm{o}}$ is typically smaller than $\sigma^2_{X,\mathrm{c}}$ (Fig. 6.2a). In this regime the impact of production bursts is reduced because proteins produced during many previous signal periods contribute to $X(t)$. Since $\mu < \omega_T$, the region $\omega \lesssim \mu$ where $g^2(\omega)$ is largest does not reach the first peak of $S_{P,\mathrm{o}}(\omega)$ at ω_T (Fig. 6.2d, solid black line); hence σ^2_{ex} is dominated by promoter switching noise at low frequencies $\omega < \omega_T$, for which $S_{P,\mathrm{o}}(\omega) < S_{P,\mathrm{c}}(\omega)$. The large amplitude changes of the oscillatory signal strongly bias the promoter to be active during a signal pulse and to be inactive between pulses, which in turn greatly reduces the probability of observing very long periods of promoter (in)activity. The elimination of such slow promoter fluctuations, which lead to the largest fluctuations in $X(t)$, means that on the timescale of the protein lifetime, promoter activity becomes more reproducible and the production of proteins becomes less variable when driven by an oscillatory signal (Fig. 6.2e). Furthermore, Fig. 6.2f shows that in this regime $\sigma^2_{X,\mathrm{o}} \leq \sigma^2_{X,\mathrm{c}}$ over nearly the full range of expression levels as the signal period T is varied, indicating that this result does not require fine tuning of the reaction rates to the oscillation timescale.

Output noise σ^2_X tends to decrease as promoter switching is made faster (Fig. 6.2g): increasing the promoter switching rate reduces $S_P(\omega)$ at low frequencies, shifting power instead to high frequencies where $g^2(\omega) \sim \omega^{-2}$ is small. Interestingly, for extremely fast switching the constant signal is able to achieve a smaller variance. In this limit noise in promoter switching becomes negligible, and Eq. 6.1 reduces to

$$\emptyset \xrightarrow{\rho' s(t)} \mathsf{X} \xrightarrow{\mu} \emptyset. \tag{6.9}$$

With a constant signal the effective production rate of X is $\rho'_{\mathrm{c}} = \rho\langle P^* \rangle/\alpha$ and $\sigma^2_{X,\mathrm{c}} = \langle X \rangle = \sigma^2_{\mathrm{in}}; \sigma^2_{\mathrm{ex}}$ (Eq. 6.7) can be made arbitrarily small by shifting all promoter fluctuations to extremely high frequencies. With an oscillatory input the effective production rate becomes $\rho'_{\mathrm{o}} = \rho/(1+K)$, where $K = \lambda/\kappa$. The resulting variance,

$$\sigma^2_{X,\mathrm{o}} = \langle X \rangle \left[1 + \frac{\rho'_o}{\mu} \left(1 - \frac{\tau}{T} - \frac{(1-e^{-\mu\tau})(1-e^{-\mu(T-\tau)})}{\mu\tau(1-e^{-\mu T})} \right) \right], \tag{6.10}$$

includes a (positive) extrinsic contribution. While the promoter switching noise background in $S_{P,\mathrm{o}}(\omega)$ vanishes, as with a constant signal, the peaks due to the periodicity of $s(t)$ remain. Hence for an oscillatory input there will always be some overlap between $S_{P,\mathrm{o}}(\omega)$ and $g^2(\omega)$, and a non-zero extrinsic noise σ^2_{ex}. With the reaction rates representative of eukaryotic transcription used in Fig. 6.2g, the cross-over at which $\sigma^2_{X,o} = \sigma^2_{X,c}$ is $\kappa \approx 18\,\mathrm{min}^{-1}$. Experimentally-determined rates of promoter activation, however, are typically $0.01 - 1\,\mathrm{min}^{-1}$ [164, 165], which suggests that under biologically-relevant conditions oscillatory signals can lead to more robust expression.

Thus far we have assumed that the signals $s(t)$ are deterministic. In reality, these signals will also be noisy as they are themselves generated by stochastic biochemical processes. An important question is whether oscillatory signals can still be decoded more reliably once noise in the input stimulus is taken into account.

We first consider the effect of transcription factor copy-number fluctuations around a

constant mean by simulating $s(t)$ as a birth-death process. We find that σ_X^2 at fixed $\langle X \rangle$ increases monotonically as the variance or correlation time of $s(t)$ are increased. Therefore, without additional non-exponential temporal correlations in the input signal, noise is unable to reduce σ_X^2 below that achieved with $s(t) = \alpha$.

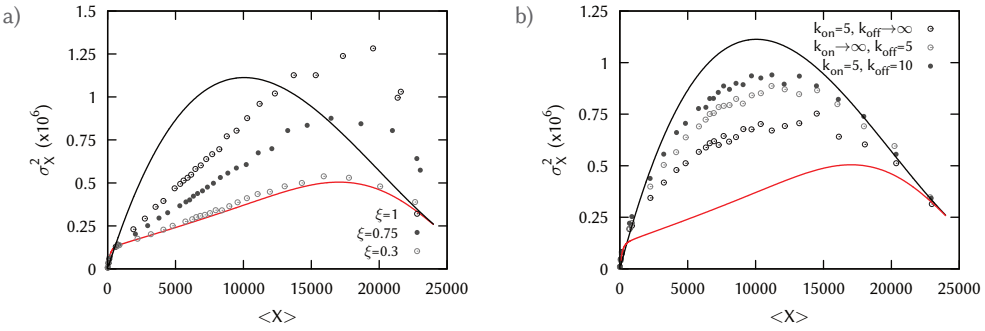


Figure 6.3: Effects of noise in **a)** oscillation amplitude, and **b)** the duration of "on" and "off" periods. Even with significant variability in amplitude or timing, the oscillating signal leads to a smaller σ_X^2 than a constant signal. Parameters are as in Fig. 6.2**f**.

For an oscillatory input signal there are two principal types of noise: fluctuations in the amplitude and timing of signal pulses. First we perform simulations in which on each occasion that the oscillatory signal switches "on" the amplitude $s(t) = a$ is sampled from a log-normal distribution with mean $\overline{a} = 1$ and width parameterized by $\xi^2 = \ln[\overline{a^2}]$. Figure 6.3a shows that σ_X^2 is largely unchanged until the noise in the pulse amplitude becomes large. Even with $\xi = 1$ ($\sigma_a/\overline{a} \approx 1.3$) the output noise is typically smaller than for a constant, noiseless, input signal. Intuitively, in the regime where an oscillating signal leads to a smaller σ_X^2, $K \ll 1$; the promoter activity is driven to saturation during a signal pulse and hence amplitude fluctuations have little effect until there is a significant probability that $a \sim K$. The decoding of large-amplitude oscillatory signals is therefore highly robust to noise in the oscillation amplitude.

We next investigate by simulations the effect of noise in the duration of signal pulses and inter-pulse intervals. The duration of each "on" period is chosen independently from a gamma distribution with mean τ and shape parameter k_{on}. The resulting variance in "on" durations is $\sigma_{t_{\text{on}}}^2 = \tau^2/k_{\text{on}}$. "Off" periods are similarly sampled from a distribution with mean $T - \tau$ and parameter k_{off}. Figure 6.3b shows that the noisy oscillation can lead to protein level fluctuations which are similar to or smaller than $\sigma_{X,\text{c}}^2$ even when variability in signal timing is significant ($k_{\text{on}} = 5$ gives $\sigma_{t_{\text{on}}}/\tau \approx 0.45$). Experimentally-observed fluctuations in oscillation periods or peak widths vary between different systems but can be $20 - 30\%$ [156, 167], suggesting that in vivo oscillations can be sufficiently precise as to reduce output noise compared to a constant-amplitude signal.

The simple model of gene expression considered here neglects mRNA dynamics and processing. Such processes can affect the propagation of promoter-state fluctuations in two ways. First, mRNA dynamics will integrate over promoter fluctuations on timescales shorter than the mRNA lifetime (typically tens to hundreds of minutes in eukaryotic cells [16, 20]).

However, differences between $\sigma^2_{X,c}$ and $\sigma^2_{X,o}$ are primarily due to promoter-state fluctuations on timescales comparable to or longer than the slowest timescale of variations in the protein level, which is typically determined by the protein life-time (ten to a hundred hours [166]). Promoter fluctuations on these long timescales can not be filtered by the mRNA dynamics, and hence even with mRNA dynamics taken into account an oscillatory input will lead to more robust expression since an oscillatory signal suppresses fluctuations on timescales longer than the signal period (see Fig. 6.2d). Second, a (random) delay between transcription initiation and protein synthesis will be introduced. However, such delays will only significantly affect the number of proteins produced on the timescale τ_X if the width of the delay distribution itself becomes comparable to the protein lifetime, which seems unrealistic. Hence we conclude that mRNA dynamics will have little effect on our results.

It is believed that biochemical networks employ frequency-encoding schemes in which stimuli are represented in the frequency of oscillations of signaling molecules [155, 159]. Our results suggest that frequency-encoding may allow for more reliable signaling than amplitude-encoding schemes because oscillatory signals can be decoded more reliably.

6.4 Acknowledgements

I thank Andrew Mugler for comments on the chapter.

Chapter 7

Amplitude multiplexing of biochemical signals

In this chapter we show that living cells can multiplex biochemical signals, *i.e.*, transmit multiple signals through the same signaling pathway simultaneously, and yet respond to them very specifically. We demonstrate how two binary input signals can be encoded in the concentration of a common signaling protein, which is then decoded such that each of the two output signals provides reliable information about one corresponding input. Under biologically relevant conditions the network can reach the maximum amount of information that can be transmitted, which is 2 bits.

Based on manuscript W.H. de Ronde, F. Tostevin and P.R. ten Wolde (**2011**). Multiplexing Biochemical Signals *Physical Review Letters* **107**.
doi:10.1103/PhysRevLett.107.048101

7.1 Introduction

Cells continually have to respond to a myriad of signals. One strategy for transmitting distinct stimuli is to use distinct signal transduction networks . It is, however, increasingly recognized that components are often shared between pathways [168]. Moreover, cells can transmit different signals through one and the same pathway, and yet respond to them specifically. In rat cells, for instance, neuronal growth factor and epidermal growth factor stimuli are transmitted through the same MAPK pathway, yet give rise to different cell fates, differentiation and proliferation respectively [91]. These observations suggest that cells are able to transmit multiple messages through the same signal transduction network, just as many telephone calls can be transmitted via a single wire. Indeed, the intriguing question arises whether biochemical networks, like electronic circuits, can *multiplex* signals: can multiple input signals be combined (encoded) *simultaneously* in the dynamics of a common signaling pathway, and then decoded such that cells can respond specifically to each signal (see Fig. 7.1)?

An open question in biology is how cells transduce multiple signals via pathways that share components, since sharing components may lead to unwanted crosstalk between the different signals. In recent years, several mechanisms for ensuring signaling specificity have been proposed. One is spatial insulation, where the shared components are incorporated into distinct macromolecular complexes on scaffold proteins [168, 169], leading effectively to *independent* communication channels for the transmission of the respective signals. Other proposals are based on the temporal dynamics of the system, such as cross-pathway inhibition [170] and kinetic insulation [171]. With these mechanisms the system cannot be decomposed into independent pathways for the transmission of the respective signals. Yet, these studies suggest that multiple messages cannot be transmitted simultaneously, because one pathway tends to dominate the response. Here we demonstrate that cells can truly multiplex signals: we show that they can transmit at least two signals simultaneously through a common pathway, and yet respond specifically to each of them.

Cells employ a number of coding strategies for transducing signals, such as encoding stimuli in the temporal dynamics, like the duration [91] or frequency [172], of an intracellular signal. In principle, any coding strategy could be used to multiplex signals. Here, we consider what is arguably the simplest and most generic coding strategy cells could choose, namely one in which the signals are encoded in the concentrations of the signaling proteins. We will call this strategy AM multiplexing.

7.2 Model

We will consider the biochemical network shown in Fig. 7.1a. It consists of N input species $\mathrm{S}_1, \ldots, \mathrm{S}_N$ with copy numbers $S_1, \ldots, S_N$, a common signal transduction pathway $\mathcal{V}$ and N output species $\mathrm{X}_1, \ldots, \mathrm{X}_N$. The copy number of each input species S_i can be in one of K states, $s_i = 0, \ldots, K-1$, which are labeled in order of increasing copy number, $S_i^{(0)} < S_i^{(1)} < \cdots < S_i^{(K-1)}$. The input pattern is denoted by the vector $\mathbf{s} = (s_1, \ldots, s_N)$. Similarly, the copy number of each output species X_i can be in one of L states $X_i^{(x_i)}$, with

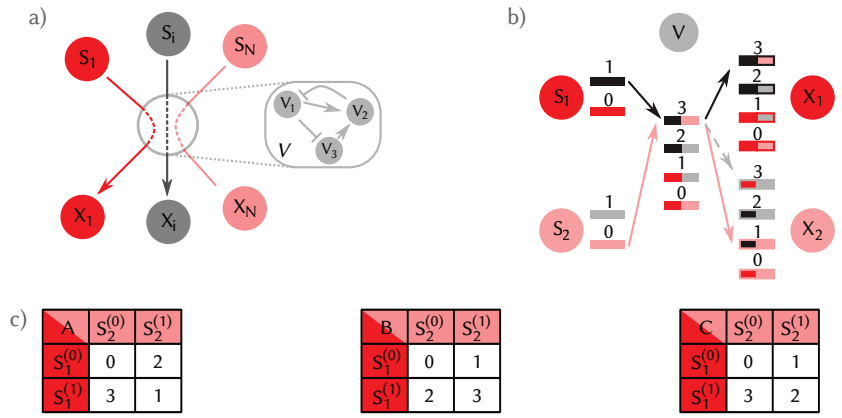


Figure 7.1: **a)** Biochemical multiplexing: N different signals are encoded in the state of a common pathway $\mathcal{V}$, and then decoded such that each output species X_i responds to the corresponding input S_i. **b)** Multiplexing is a mapping problem. The states of two inputs S_1 and S_2 are mapped onto the concentration of V, which is then mapped onto states of the output species X_1 and X_2; we require that the two lowest (highest) levels of X_i correspond to the lowest (highest) level of S_i. The dashed arrow denotes a mapping that violates this requirement; levels of V and X_i are colored according to input pattern $\mathbf{s}=(s_1,s_2)$. **c)** The 3 distinct mappings of **s** to v; in panel **b)** mapping C is shown.

$x_i = 0, \ldots, L-1$, ordered by increasing copy number, and the output pattern is denoted by the vector $\mathbf{x} = (x_0, \ldots, x_N)$. A necessary condition for multiplexing is that the state space of $\mathcal{V}$ is large enough that it is possible to encode the total number of input patterns, K^N, in $\mathcal{V}$.

We imagine that the N input signals are independent, and that the signal transduction network $\mathcal{V}$ replaces N independent signaling pathways. We therefore require that X_i should provide reliable information about the state s_i, but not necessarily about $s_{j\neq i}$; the N different input signals **s** simply have to be transduced to **x**, not necessarily integrated. In general, however, the state x_i will be a function of the states of all the input species: $x_i = f(\mathbf{s})$. This reflects the fact that inevitably there is cross-talk between the different signals because they are transmitted via the same pathway. However, this cross-talk is not detrimental as long as it does not compromise the cell's ability to infer from x_i what s_i was.

Another key point is that while the precise mapping from **s** to **x** may not be critical for the amount of information transmitted *per se*, this is likely to be important for whether or not this information can be exploited. Let us imagine that the system contains three input species, say three sugars, each of which is either present or absent, $s_i = 0$ or 1; let us further assume that X_i is an enzyme needed to consume sugar S_i. With 8 input patterns X_i can, in the absence of noise, take 8 values, identified as states $x_i = 0, \ldots, 7$. Now, it seems natural to demand that when the sugar S_i is absent ($s_i = 0$), the copy number of enzyme X_i is low, while when S_i is present, the copy number of X_i is high; this means that the four lowest levels of X_i ($x_i = 0, 1, 2, 3$) should correspond to $s_i = 0$, while the four highest levels of X_i should correspond to $s_i = 1$. We therefore require that the mapping from **s** to **x** is such

that the output states $\{x_i\}$ corresponding to input $s_i = j$ are grouped into sets that are *contiguous* and either increase or decrease *monotonically* with j, for each signal i. This leads to a monotonic input-output relation between S_i and X_i for each i. We call this requirement the multiplexing requirement.

In the rest of the chapter, we make these ideas concrete for a network with two input species, S_1 and S_2, each of which has either a low ($s_i = 0$) or a high concentration ($s_i = 1$), and the shared pathway consists of only one species, V. Multiplexing requires that, in the absence of noise, the four input patterns **s** be mapped onto four distinct states of V, $V^{(v)}$ with $v = 0, \ldots, 3$, again labeled in order of increasing copy number. These four levels of V lead to four states for each of the two output species X_1 and X_2 (Fig. 7.1b). As explained above, we require that we can group these four states into two sets, called LOW and HIGH, such that the LOW set, containing $x_i = 0, 1$, corresponds to $s_i = 0$ and the HIGH set, containing $x_i = 2, 3$, corresponds to $s_i = 1$ (or *vice versa*, leading to an inverse input-output relation). We note that there exist different ways of mapping **s** to v, but not all of these mappings can necessarily be decoded into **x** in a manner that satisfies the multiplexing requirement. We therefore first address the question of which combinations of mapping from **s** to v and decoding from v to **x** fulfill the multiplexing requirement, and then we will discuss what encoding mechanisms actually allow for the required mapping from **s** to v.

Because of the symmetry in the problem, there are three distinct ways of mapping the four input patterns **s** to v (Fig. 7.1c). To determine whether there exists a scheme for decoding the signals from v to **x** that satisfies the multiplexing requirement. To determine whether there exists a scheme for decoding the signals from v to **x** that satisfies the multiplexing requirement, we examine for each mapping all possible network topologies between V, X_1 and X_2, except those that involve autoregulation or mutual regulation since these may lead to bistability. In particular, we allow for activation and repression of X_1 and X_2 by V, and for activation and repression of X_2 by X_1, leading to feedforward loops, a common motif in signal transduction pathways and gene networks [11]. In the deterministic mean-field limit the steady-state values of X_1 and X_2 are thus given by

$$X_1 = k_1 f(V; K_\alpha, n_\alpha)/\mu, \tag{7.1}$$

$$X_2 = k_2 f(V; K_\beta, n_\beta) \times f(X_1; K_\gamma, n_\gamma)/\mu, \tag{7.2}$$

where k is the maximum activation or production rate, μ is the degradation or deactivation rate, and each regulation function is either an activating or repressing Hill function, $f(V; K, n) = V^n/(V^n + K^n)$ or $f(V; K, n) = K^n/(V^n + K^n)$. The multiplication in Eq. 7.2 indicates that we assume that at X_2, X_1 and V are integrated according to AND logic [11]. To explore which architectures allow for multiplexing we performed extensive sampling of the space of parameters $k_1, k_2, K_\alpha, n_\alpha, K_\beta, n_\beta, K_\gamma, n_\gamma$ for each of the mappings in Fig. 7.1c.

Only for mapping C do we find decoding schemes that satisfy the multiplexing requirement for realistic parameter values. Interestingly, all valid decoding networks are incoherent feedforward loops [11]. Figure 7.2 illustrates the principle for one such motif. $X_1(V)$ is a simple activation curve with activation threshold K_α. In contrast, $X_2(V)$ starts low and

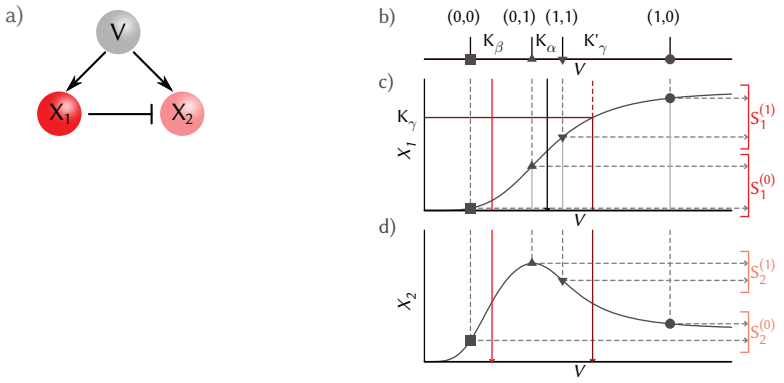


Figure 7.2: Decoding V using an incoherent feedforward loop. **a)** Network architecture. **b)** The values of V corresponding to the four input patterns **s** for mapping C (see Fig. 7.1c), with thresholds K_α,K_β,K_γ (see Eqs. 7.1-7.2). **c)** X_1(V). **d)** X_2(V). The non-monotonicity of X_2(V) swaps the states corresponding to (1,1) and (1,0) in the mapping from v to x_2.

rises around K_β, but then decreases again due to repression by X_1. This non-monotonicity, which is a result of the incoherent character of the feedforward loop, is critical since this makes it possible to swap the order of the states corresponding to $\mathbf{s} = (1,1)$ and $(1,0)$ in the mapping from v to x_2. For sharp regulation functions, this yields the intuitive requirement that $V^{(0)} < K_\beta < V^{(1)}$ and $X_1^{(2)} < K_\gamma < X_1^{(3)}$. We stress, however, that n_β and n_γ do not have to be large (they can be as small as unity), and then K_β and K_γ can even fall outside these ranges. In order for X_1 to be able to repress X_2 at $X_1^{(3)}$ but not $X_1^{(2)}$, X_1 should not be saturated at $V^{(2)}$; this means that $K_\alpha \gtrsim V^{(2)}$ and n_α should not become so large that $X_1(V)$ becomes two-valued. These mild constraints indicate that this is a robust decoding scheme that can be implemented for a wide range of parameter combinations.

We can now also understand why mappings A and B are difficult to decode: they would require an input-output relation between X_2 and V that rises more than once. This is difficult to achieve in a feedforward loop without mutual repression or activation.

7.3 Multiplexing

The above analysis shows that it is possible to decode multiple signals simultaneously, provided that the input **s** can be encoded in V according to mapping C. The next question is how this mapping, which corresponds to a particular input-output relation $V(S_1, S_2)$, can be generated. Multiplexing is most beneficial when the input signals are multiplexed at the beginning of signal transmission. We therefore consider whether two input signals can be multiplexed at the level of a single protein V, which could be an enzyme or a receptor at the very beginning of a signaling cascade. We consider a canonical motif where S_1 and S_2 bind competitively to V, which can exist in either an active conformational state A or an inactive

state I (Fig. 7.3a). In equilibrium the mean number of active V molecules will be

$$V(S_1, S_2) = \frac{V^{\max}(1 + q_1^A + q_2^A)}{1 + q_1^A + q_2^A + e^{\Delta E_0}(1 + q_1^I + q_2^I)}, \tag{7.3}$$

where $V^{\max}$ is the total number of V molecules, $\Delta E_0 = E_0^A - E_0^I$ is the free-energy difference between the active and inactive states of V in the absence of ligand binding, and $q_i^{I,A} = S_i/K_i^{I,A}$ with K_i^I and K_i^A the dissociation constants for the binding of S_i to inactive and active V, respectively. We find that a scheme where S_1 activates V more strongly than S_2, giving energy levels as in Fig. 7.3b, generates the required $V(S_1, S_2)$ (Fig. 7.3c); competition between the two ligands reduces the probability of S_1 binding the receptor when S_2 is present, ensuring that $V(S_1^{(0)}, S_2^{(1)}) < V(S_1^{(1)}, S_2^{(1)}) < V(S_1^{(1)}, S_2^{(0)})$. In the limits $S_i^{(0)} \ll K_i^{I,A} \ll S_i^{(1)}$ a sufficient condition for this ordering is $1 > K_2^A/K_2^I > K_1^A/K_1^I$. Clearly, therefore, the required encoding $V(S_1, S_2)$ could be implemented at the level of a single signaling protein V.

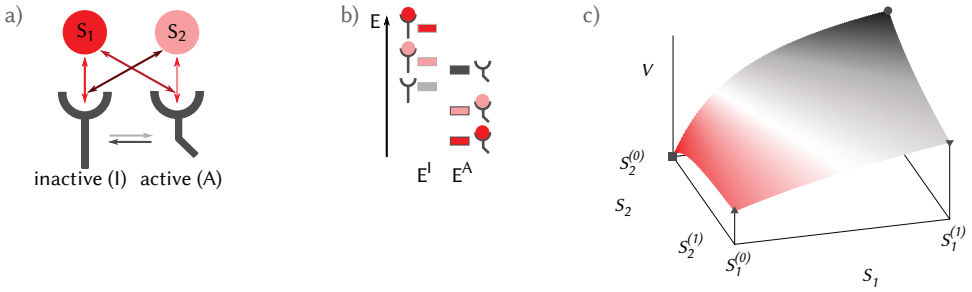


Figure 7.3: a) An encoding scheme where two ligands S_1 and S_2 competitively bind to a protein V, which can be in an active state A or an inactive state I. **b)** Energy levels of V in the active or inactive state with no ligand, ligand S_1 or ligand S_2 bound; $e^{-E_i^I} \equiv S_i/K_i^I$; $e^{-E_i^A} \equiv e^{-\Delta E_0} S_i/K_i^A$. **c)** Input-output relation $V(S_1,S_2)$ corresponding to the energy levels in **b)**, yielding mapping C (Fig. 7.1c); symbols ■,▲,●,▼ correspond to states in Fig. 7.2b,c,d.

The analysis above shows that in principle biochemical networks can multiplex signals in the mean-field, deterministic limit. However, there remains the question of whether signals can be multiplexed reliably in the presence of inevitable biochemical noise. To address this, we estimate a lower bound on the information about two binary signals S_1 and S_2 that are transmitted through the network studied above (Eqs. 7.1-7.3). We define the total information $I \equiv I(S_1, X_1) + I(S_2, X_2)$ as the sum of the mutual information for each of the individual signals[45]. Note that in the presence of noise X_i is not limited to 4 states but can in principle take any value. This definition of I makes it straightforward to directly compare the performance of this network with that of two independent pathways. If each of the two input states for each S_i is equally likely then the maximum value of $I(S_i, X_i)$ is 1 bit for each signal i; the maximum value of I is thus 2 bits.

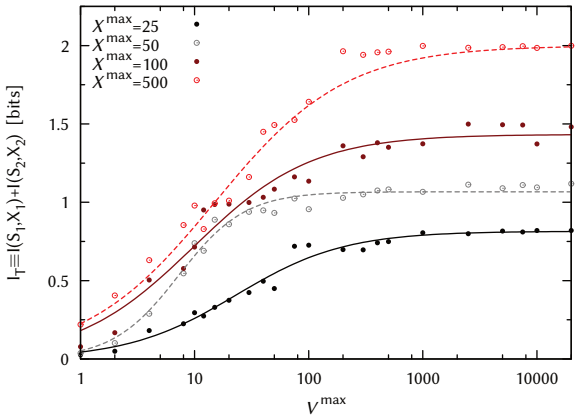


Figure 7.4: The transmitted information $I \equiv I(S_1,X_1)+I(S_2,X_2)$ as a function of V^{max} for four different values of X^{max}. I can reach the maximum value of 2 bits provided V^{max} and X^{max} are large enough. Typical values for the optimized parameters are: $K_\alpha \sim V^{max}$, $K_\beta \sim 0.3V^{max}$, $K_\gamma \sim 0.5X^{max}$, $\Delta E_0 \approx -2k_BT$, $S_1^{(0)}=S_2^{(0)}=0$, $K_1^I \gg S_1^{(1)}$, $K_2^I \sim S_2^{(1)}$, $K_1^A \ll S_1^{(1)}$, $K_2^A < S_2^{(1)}$.

To maximize the lower bound on I we optimize the network parameters using a simulated-annealing algorithm. We fix the deactivation rate of X_1 and X_2 to be $\mu = 1s^{-1}$ and set $n = n_\alpha = n_\beta = n_\gamma = 2$; the results are insensitive to the precise value of n for $1 \leq n \leq 4$. We set the maximum value of each X_i to be $X^{\max}$; this specifies the activation rates k_1 and k_2. We optimize the remaining decoding parameters K_α, K_β and K_γ over the range $[0, V^{\max}]$ or $[0, X^{\max}]$ as appropriate, and the encoding parameters q_i^j and ΔE over the range $[10^{-3}, 10^3]$ and $[-10k_BT, 10k_BT]$, respectively. Ligand binding is assumed to be fast compared to receptor activity switching; the rate of switching of V from the active to the inactive state is fixed at $1s^{-1}$, and the switching rate in the reverse direction then follows from detailed balance given Eq. 7.3. For each parameter set we compute the mutual information using the linear-noise approximation [45, 68]. Its accuracy was verified by performing Gillespie simulations of the optimized networks [75].

Using this procedure we compute I as a function of $V^{\max}$ and $X^{\max}$, which determine the intrinsic noise in the encoding and decoding processes. Figure 7.4 shows that below a threshold copy number $V_c^{\max} \approx 10$ the total information is low regardless of $X^{\max}$ because four distinct states of V cannot be generated. Once $V^{\max}$ becomes sufficiently large that the four encoded signals are well separated, I saturates at a value determined by the level of noise in the production and decay of X_i. For large $X^{\max}$ the information I reaches 2 bits, thus matching the performance of two independent channels. Importantly, I reaches 2 bits for $V^{\max} \approx X^{\max} \approx 500$, which is well within the range of typical protein copy numbers inside living cells. This shows that biochemical networks can multiplex two signals reliably at biologically relevant noise levels.

In summary, our results suggest that cells can transmit at least two binary signals through one and the same pathway, and yet respond specifically and reliably to each of them. The

proposed mechanism for biochemical multiplexing is based on swapping the order of states during the encoding and decoding steps. It is clear that the principle is generic, and could be implemented in any biochemical network that uses an incoherent feedforward loop [11]. Indeed, the architecture of a number of well-studied systems resembles that of the system studied here. It is well known that G-protein coupled receptors (GPCRs) can be stimulated by many different ligands, yet give rise to different cell fates, a phenomenon referred to as "multiplicity" [173]. A GPCR activates two G proteins, Gα and G$\beta\gamma$. Of particular interest is the system where Gα_q activates RhoGef63 [174], leading to cytoskeletal remodeling, while G$\beta\gamma$ activates PLCβ [175], ultimately controlling cell proliferation. Interestingly, PLCβ binds not only G$\beta\gamma$, but also Gα_q, thereby inhibiting the activation of RhoGef by Gα [174]. Hence, we have a scheme where one protein V (the GPCR) activates two proteins X_1 (RhoGef63) and X_2 (PLCβ), whereby X_2 effectively inhibits X_1. This is highly similar to our proposed multiplexing scheme, and it is indeed tempting to speculate that multiplicity is achieved via multiplexing. Secondly, Ras is a major hub in cell signaling: it is activated by many cellular stimuli, including growth factors, differentiation factors and cell survival factors. Activation of Ras leads to the activation of the Raf-Mek-Erk MAPK pathway, ultimately leading to differentiation or proliferation, and also via PI3K to the activation of Akt and hence the mTOR pathway, controlling many processes such as metabolism. Importantly, Akt also *de*activates Raf [176], generating a Ras (V)-Raf (X_1)-Akt (X_2) incoherent feedforward loop. Finally, CREB, a major transcription factor in neuronal cells, is activated by many different stimuli, and, in turn, regulates many different target genes, some of which include incoherent feedforward loops [177]. It will be interesting to test whether these systems employ biochemical multiplexing when they are stimulated by two signals simultaneously.

7.4 Acknowledgements

I thank Tom Shimizu, Frank Bruggeman and Nils Blüthgen for a critical reading of the chapter.

Chapter 8

Multiplexing oscillatory biochemical signals

In recent years it has been realized more and more that biochemical signals are not necessarily constant in time, and more importantly that the temporal dynamics of a signal can be the information carrier. In this chapter we show that living cells can multiplex both constant and oscillating signals, *i.e.*, transmit multiple signals through the same signaling pathway simultaneously, and yet respond to them very specifically. We study the information transmission for each signal and find that, under biologically relevant conditions, more than 8 bits of information can be transmitted simultaneously. Moreover, the use of an oscillatory signal does not reduce the fidelity of information transmission for the constant signal at all, showing indeed that multiplexing a constant signal and an oscillatory signal is an interesting strategy for a cell to transmit information.

Based on manuscript W.H. de Ronde, A. Kan and P.R. ten Wolde. Multiplexing oscillatory biochemical signals *In preparation.* Anton Kan provided the study on the biological examples.

8.1 Introduction

Cells depend on stimuli from their environment to initiate behaviors, including growth, division, differentiation, and death. One strategy to transmit different stimuli is to use distinct signaling pathways for the respective signals. An increasing number of experiments, however, show that components are often shared between different pathways. Even more strikingly, cells can transmit multiple signals through a single pathway and yet respond to them specifically. For example, in rat PC-12 cells both the epidermal (EGF) and neuronal growth factor (NGF) signal through the same Mitogen-Activated Protein Kinase (MAPK) pathway, yet these two signals give rise to different cell fates, differentiation and proliferation [91]. These observations suggest that cells are able to transmit multiple messages through the same signal transduction network, just as many telephone calls can be transmitted via a single wire. Indeed, the intriguing question that arises is whether biochemical networks, like electronic circuits, can multiplex signals, meaning that multiple input signals are combined (encoded) simultaneously in the dynamics of a common signaling pathway, which is then decoded such that cells respond specifically to each signal.

One of the key problems in multiplexing is the unwanted crosstalk between the different signals: from the perspective of one signal, the presence of other signals constitutes noise. In recent years several mechanisms for ensuring signaling specificity have been proposed. One is spatial insulation, where the shared components are incorporated into distinct macromolecular complexes on scaffold proteins [168, 169]. Other proposals are based on the temporal dynamics of the system, such as cross-pathway (or mutual) inhibition [39, 170, 178, 179, 180] or kinetic insulation [171]. However, these studies only considered scenarios in which the system is stimulated with one signal, or if stimulated with multiple signals, can only transduce one signal at the time.

Rensing and Ruoff studied what happens when two or three MAPK pathways that share components are stimulated simultaneously [181] and found that one pathway tends to dominate the response, suggesting that multiple messages cannot be transmitted simultaneously. In contrast to this study, in a previous study we have shown that cells indeed can employ a multiplexing strategy. The concentration levels of two independent signals can be encoded in a shared pathway and decoded to yield two independent responses, such that each response is specific to its corresponding input signal but insensitive for its read-out to the other signal. We called this strategy AM multiplexing [182].

Interestingly, biological signals do not necessarily have constant level. Indeed, oscillations in intracellular signals (NF-κB [157, 183], Ca^{2+} [172] or nuclear Erk localization [156]) have been observed experimentally. For some extracellular signals it is also known that under physiological conditions the concentration is time-dependent, like Gonadotropin Release Hormone (GnRH) [184] or Insulin [185]. For many other extra-cellular signals the time-dependent behavior to our knowledge is not precisely known (e.g. EGF).

It is well-known that specific signals activating a common pathway, can give rise to different dynamical behaviors of the intermediate component. One example is the already introduced EGF and NGF dependent ERK-response [91, 186]. Another interesting example is the p53 response in damaged cells. Upon double-stranded DNA breaks, p53 shows an oscillatory response, while upon single-stranded DNA breaks p53 shows a constant response

[187]. These studies raise the idea that cells can encode constant and temporal signals simultaneously into a shared pathway.

Importantly, *in vivo* studies have shown that the response of biological systems can depend on the precise dynamics of the input. Dynamic stimuli have many different characteristics (e.g. frequency, amplitude, number of peaks), each of which could be an information carrying property. GnRH is a pulsatile extracellular signal. It has been shown that the response of the system is sensitive to the GnRH signal frequency [188, 189]. Another well-known frequency sensitive system is the response of NF-κB and NFAT to Ca^{2+} [190, 191], where gene-expression depends on the Ca^{2+} frequency. The use of information in the periodicity of the signal is also well-known for processes related to the cell-cycle [192, 193] or morphogenesis [194]. These examples suggests that cells can not only encode multiple signals into the temporal dynamics of a shared signaling pathway, but also decode this dynamics.

Indeed, the combination of signal — dependent temporal dynamics of intermediate components and subsequent temporal — dynamics sensitive read-out has been observed experimentally. The temporal profiles for the central node kinase IKK or transcription factor NF-κB are different for different inflammatory stimuli and the eventual read-out is dependent on the specific temporal dynamics [183, 195, 196, 197]. More examples are provided in the recent review article by Behar and Hoffmann [153].

The question whether two signals EGF and NGF in the MAPK system can be read out reliably has also been studied in more detail. The constant signal ERK, in the case of NGF stimulation, can be reliably read-out by a coherent feed-forward loop, as has been observed downstream of the NGF-response [106, 198]. The coherent feed-forward ensures that the transient ERK signal in response to EGF stimulation does not lead to a response, thereby preventing crosstalk between NGF and EGF signaling. However, whether a feed-forward loop also prevents cross-talk for a continuous oscillatory signal, or a series of pulses is unclear. For example, if presented with a pulse-train, the coherent feed-forward motif is only insensitive to pulses of a very specific shape. In a more detailed study of a different, yet similar, system, the downstream read-out of ERK after stimulation by EGF of HRG has been analyzed [199]. This study revealed that a combination of coherent feed-forward loops and negative feedback act together to block cross-talk from transient signals (or intermediary components). But for this system as well, it is unclear whether the combination of multiple loops prevents cross-talk for a continuous series of pulses or oscillations.

On the other hand, it has been suggested that a frequency modulated signal could be read-out without cross-talk from the constant signal, by a system that acts as a pulse counter [200]. It is, however, difficult to imagine how a simple biochemical system can act as a true pulse counter. Many pulse-counting mechanism are in fact time-integrators of pulses [153], with a response at a specific concentration threshold and therefore do not truly count (or respond to) pulses.

Since information can be encoded both in the frequency and in the amplitude of a specific signal, we wondered whether it is possible to combine frequency-encoded and amplitude-encoded signals to multiplex information. In our setup, a shared pathway is stimulated by two signals, one of which is time-dependent. The information in the time-dependent signal

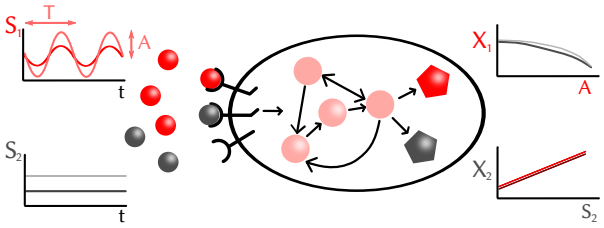


Figure 8.1: Two extracellular signals (S_1,S_2) stimulate a shared pathway, here by activating the same receptor. S_1 is time-dependent, S_2 is constant. The concentration of the output X_1 only depends on the characteristics of S_1, but not on the concentration of S_2. In contrast, the output X_2 depends on the concentration level of S_2, but not on the amplitude A or frequency T of S_1.

is either encoded in the frequency of the signal, or in its amplitude. The information in the constant level is encoded in its mean concentration level. These two signals are combined and simultaneously transmitted through a common pathway. The common signal is then decoded and split in two output responses. One output is only sensitive to the amplitude (or frequency) of the corresponding time-dependent input, such that the level of the output depends only on the amplitude (or frequency) of the input. The other output depends only on the concentration of the constant signal of the input (see Fig. 8.1).

8.2 The model

In this section we specify in detail the encoding and decoding steps. A cartoon of the model is shown in Figs. 8.1 and 8.2.

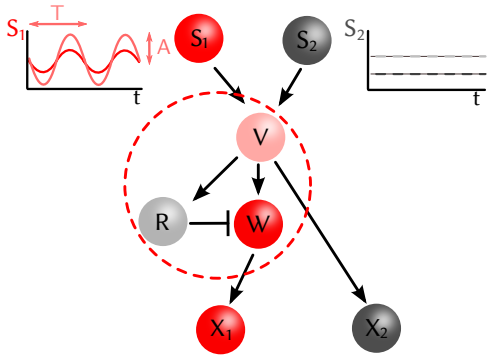


Figure 8.2: Schematic drawing of the multiplexing system. Encircled is the adaptive motif.

8.2.1 Encoding

In the encoding step of the motif, the two signals S_1, S_2 are combined into the shared pathway. The signals are modeled as a simple sinusoidal function

$$S(t) = \mu\left(1 + A\sin\left(2\pi\frac{t}{T}\right)\right). \tag{8.1}$$

μ is the signal mean, A is the signal amplitude (bounded between $[0:1]$) and T is the period of the signal oscillation. We assume that the signals are deterministic and discuss the effects of noise later. S_1 is an oscillatory signal, with kinetic parameters A_1, T_1 and constant μ_1. S_2 is constant, $A_2 = 0$, and the concentration level μ_2 carriers the information in the signal. In recent years it has been shown that biochemical systems can tune separately the amplitude and frequency of a signal [201, 202].

The most simple shared pathway is a single component (V), which could be a receptor on the cell or nuclear membrane, but could also be an intracellular enzyme or a gene-regulatory protein. We image that each signal is a kinase for this intermediate component V, which can switch between an active (e.g. phosphorylated) state (V^P) and an inactive (e.g. unphosphorylated) state, such that

$$\frac{dV^P}{dt} = \frac{k_V\left[\sum_i S_i(t)\right]\left(V_T - V^P\right)}{K_V + \left(V_T - V^P\right)} - m_V\frac{V^P}{M_V + V^P}, \tag{8.2}$$

where we sum over the total number of (time-varying) signals $S_i(t)$. The dephosphorylation is mediated by a phosphatase, that has a constant copy number. In Eq. 8.2 we assume Michaelis-Menten dynamics for V (see App. 8.A.1 for more details).

If the signal information is encoded into its dynamics, it is important to propagate this dynamics as accurately as possible, up to the point where the information is eventually decoded to a read-out. Otherwise, information will get lost due to the inaccurate tracking of the signal. Therefore we require that the component V accurately tracks the dynamics of the input signals. It is well-known that a linear transfer function between S and V does not lead to a deformation of the dynamic behavior, but only to a rescaling of the absolute levels (see App. 8.A.2). A linear transfer function can be realized if the kinase acts in the saturated regime, while the phosphatase is not saturated ($K_V \ll \left(V_T - V^P(t)\right), M_V \gg V^P(t)$), leading to

$$\frac{dV^P}{dt} = k_V\left(\sum_i S_i(t)\right) - m'_V V^P. \tag{8.3}$$

with $m'_V = m_V/M_V$.

8.2.2 Decoding V^P to X_1, X_2

The second part of the multiplexer is the decoding of the information in V^P into a functional output. The signals that are encoded in V^P have to be decoded into two output signals,

X_1 and X_2. Ultimately, an optimal system allows the cell to infer from an instantaneous measurement of the output signal the information in that signal. Therefore, the outputs of the multiplexing motif are the concentration levels of X_1 and X_2. X_1 is the response of S_1, while X_2 is the response of S_2.

In our simple model there is only one time-dependent signal S_1; S_2 is a constant signal. The response X_2 should be sensitive to the concentration of S_2, but be blind to any characteristics of S_1. Since V^P has a linear transfer function of the signals (Eq. 8.3), the average level of V^P $\langle V^P \rangle$ is independent of either A_1 or T_1. $\langle V^P \rangle$ does depend on the mean concentration level of the two signals, and since S_1 has a constant mean, changes in $\langle V^P \rangle$ reflect a change in the mean of S_2 μ_2. As a result, a simple linear time integration motif can be used as the final read-out for S_2. We therefore model X_2 as

$$\frac{dX_2}{dt} = k_{X_2} V^P(t) - m_{X_2} X_2(t). \tag{8.4}$$

Since Eq. 8.4 is linear, $\langle X_2 \rangle$ is a function of $\langle V^P \rangle$ only. Moreover, if the response time of X_2, τ_{X_2} $\left(= m_{X_2}^{-1}\right)$, is much longer than the oscillation period, the effect of the oscillations on the instantaneous concentration X_2 is integrated out. This is important to reduce the variability in $\langle X_2 \rangle$ due to dynamics in the system [203].

For X_1 a simple time-integration scheme does not work. The information that has to mapped on the output concentration X_1 is either the amplitude or the frequency of S_1. This information in S_1 is then propagated to V^P. The output X_1 should therefore depend on the frequency or the amplitude of V^P, but not on the mean response $\langle V^P \rangle$, since the mean represents the information in S_2. One possible way to construct an output that is insensitive to the mean, is an output that has a frequency-dependent response. This could be achieved via a motif whose power spectrum has a band-pass structure. Then, the response X_1 is insensitive to oscillations with very small or very high frequencies, but has a large gain for oscillations with frequencies which are within the band-pass regime. If the frequency of the incoming signal is within the band-pass range, the oscillations are strongly amplified at the response X_1. The amplitude of the oscillations of the response thus depends on the frequency T_1^{-1} and amplitude A_1 of the incoming signal, which are precisely the information carrying properties of S_1. A common biochemical motif with a frequency band-pass filter is an adaptive motif [67].

As discussed before, the amplitude of the oscillation in a protein level is not an quantity that cells can directly read-out. However, the information that is contained in the amplitude can me mapped onto the concentration of the output X_1, which can be read-out instantly. In the final step of our model therefore the output of the adaptive motif (W) is time-integrated to obtain X_1, and the concentration X_1 instantly provides all the information on S_1.

The key feature of an adaptive system is that the steady-state response is independent of the signal input, meaning that

$$\langle W \rangle = f\left(\{\text{all parameters}\} \notin \left\langle V^P \right\rangle\right). \tag{8.5}$$

We stress here two important properties of an adaptive system. First, the fact that in an

adaptive system the steady-state output level is independent of the constant input level, does not mean that the output level is zero. Second, although the output of an adaptive network is insensitive to constant inputs, the output of an adaptive system is not insensitive to dynamical inputs, even if these have a time-independent mean. This second observation is well-known and is the basis for the chemotactic behavior of *E. coli*, where the system does respond to a change in the input concentration, and the strength of the response depends on the magnitude of the change in input concentration. An adaptive system does not respond to very fast oscillations, due to the limited response time of any biochemical network. Moreover, an adaptive system does not respond to very slowly varying signals, since due to the adaptation, the system already adapts to the slowly changing signal before a change in the output can be observed.

Two common ways to construct an adaptive motif are known [101], the negative feedback motif and the incoherent feed-forward motif. In this multiplexing system we use the incoherent feed-forward motif. In the incoherent feed-forward motif a signal (S) stimulates two downstream components (R,W). One of the downstream components (R) by itself is also a signal for the other downstream component (W). Importantly, the regulatory effect of the direct pathway (S $\rightarrow$ W) is opposite to the effect of the indirect pathway (S $\rightarrow$ R $\rightarrow$ W). As a result, if S activates W, this activation is counteracted by the regulation of W through R. We thus obtain

$$\frac{dR}{dt} = k_R V^P - m_R R, \tag{8.6}$$

$$\frac{dW}{dt} = k_W \frac{V^P \left(W_T - W^P\right)}{K_W + \left(W_T - W^P\right)} - m_W \frac{R W^P}{M_W + W^P}. \tag{8.7}$$

A schematic drawing of the full motif is given in Fig. 8.2 with the adaptive motif encircled.

Setting the time-derivative to zero in Eq. 8.7 and solving for the steady state $\langle W^P \rangle$ we obtain

$$0 = \frac{\left(k_W \left(W_T - \langle W^P \rangle\right)\right) \left(m_R \left(\langle W^P \rangle + M_W\right)\right)}{\left(K_W + \left(W_T - \langle W^P \rangle\right)\right) \left(m_W k_R \langle W^P \rangle\right)}. \tag{8.8}$$

Although the full solution for $\langle W^P \rangle$ is unwieldy to present, there is no dependence on S_1 (since there is no dependence on $\langle V^P \rangle$) (see Eq. 8.8). This motif is thus completely adaptive.

For correct separation of the signals, the response W should be insensitive to the average level of V^P, $\langle V^P \rangle$, since $\langle V^P \rangle$ should not carry information from S_1, but from S_2. If W does depend on $\langle V^P \rangle$, this necessarily leads to unwanted cross-talk between the two signals. While W^P indeed is insensitive to the mean of V^P for a constant input (see Eq. 8.8), this is not necessarily the case for a dynamic $V^P(t)$. Since Eq. 8.7 is non-linear, the response W is dependent on the precise functional form of V^P, and, more importantly will depend on $\langle V^P \rangle$. In App. 8.A.3 this is studied in greater detail.

The gain of the network, $g^2(\omega)$, provides more information on the frequency-dependence of W^P [107]. The gain shows the amplification of the input signal as a function of frequency.

The full expression of the gain is unwieldy to present here, but in simplified form we have

$$g_{WP}^2(\omega) \propto \frac{\alpha\omega^2}{\beta\left(\omega^2+\tau_R^{-2}\right)\left(\omega^2+\tau_W^{-2}\right)\left(\omega^2+\tau_V^{-2}\right)}, \tag{8.9}$$

where α and β are proportionality constants and τ_i are the response times of component i. For slowly varying signals ($\omega \to 0$), the amplitude of the response is negligible due to the ω^2-term in the numerator of Eq. 8.9, reflecting the adaptive nature of the network. Second, for $\omega \ll \min\left[\tau_V^{-1}, \tau_R^{-1}, \tau_W^{-1}\right]$, the power scales with ω^2. For very large ω the power scales with ω^{-4}. In the intermediate regime for ω, the scaling depends on the precise response times. The response times are the diagonal Jacobian elements for the linearized system (Eqs. 8.2,8.6,8.7),

$$\tau_V = \left[\frac{m_V M_V}{\left(M_V+\langle V^P\rangle\right)^2} + \frac{k_V K_V \mu}{\left(V_T-\langle V^P\rangle+K_V\right)^2}\right]^{-1}, \tag{8.10}$$

$$\tau_R = \mu_R^{-1}, \tag{8.11}$$

$$\tau_W = \left[\frac{m_W M_W \langle R\rangle}{\left(M_W+\langle W^P\rangle\right)^2} + \frac{k_W K_W \langle V^P\rangle}{\left(W_T-\langle W^P\rangle+K_W\right)^2}\right]^{-1}. \tag{8.12}$$

Eq. 8.11 is the response time for a protein with a simple birth-death reaction. The mathematical form of the response times, τ_V and τ_W, Eq. 8.10 and Eq. 8.12, resembles that of switching process with a forward and backward step; their values depend on the signal parameters. In the linear regime (Eq. 8.3), τ_V simplifies to $\tau_V \approx -\left(m_V/\left(M_V\right)\right)^{-1}$, which is just the linear decay rate of V^P. Next, if $\langle W^P\rangle$ is independent of μ_i, the response time τ_W is also independent of μ_i. These results are important, since they indicate that the response times are intrinsic properties of the dynamics of the motif, which do not depend on the signal(s).

The gain (Eq. 8.9) is shown in Fig. 8.3 for three different parameter sets. The bandpass structure, with corresponding resonance frequency (the peak in the gain) is observed. Further, with circles, the response times τ_V, τ_R and τ_W are shown which determine the position of the peaks in the gain; the peak occurs at a frequency in between the two largest response times. In Fig. 8.3 we observe the influence of increasing k_R, m_R. For very slow changes in R, corresponding to k_R, m_R being low, the network has a very large gain. Increasing the response time of R, decreases the amplitude at the resonance frequency considerably. Faster tracking of V^P by R makes the adaptation of the biochemical circuit very fast and as a result, W^P does not respond at all to changes in V^P.

Finally, we look at the last step in the motif, the conversion of the dynamic response of the adaptive motif W into X_1. The instantaneous concentration X_1 should inform the system upon the type of signal input S_1. Simple time-integration of W, similar to the response X_2 (Eq. 8.4), is not sufficient. While time-integration by itself is important to average over multiple oscillation cycles, it is not sufficient because time-integration is a linear function and hence a change in the amplitude of the signal does not change the response, assuming

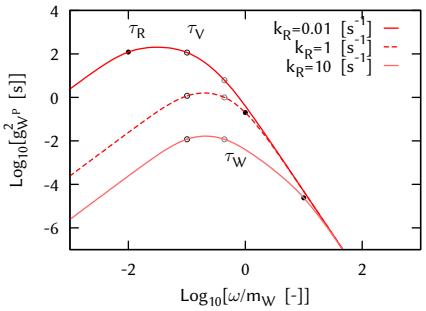


Figure 8.3: $g^2_{WP}(\omega)$ for different parameter sets. The circles indicate the response times τ_i. Parameters: k_V=1,m_V=600, k_W=m_W, K_V=$10^{-4}V_T$,M_V=$5V_T$, K_W=M_W=$W_T/2$, m_R=k_R, m_W sets the timescale.

that the oscillations are symmetric. Indeed to respond to different amplitudes, a non-linear transfer function is required (see Fig. 8.4).

$$\frac{dX_1}{dt} = k_{X_1}\frac{W^n}{W^n + K^n_{X_1}} - m_{X_1}X_1. \tag{8.13}$$

These Hill-type non-linear transfer functions are very common in biological systems, for example in gene regulation by transcription factors, or protein activation by multiple enzymes.

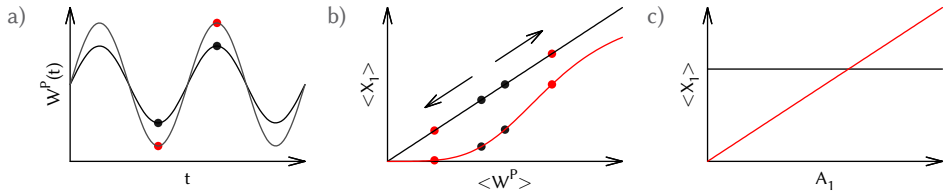


Figure 8.4: Schematic drawing of the decoding of the final step. **a)**: the oscillations in W are shown for two different input amplitudes. **b)**: The dose-response function of $\langle X_1\rangle$ versus $\langle W^P\rangle$. For a linear dose response function, a change in the amplitude of W^P is canceled, since the increase in $\langle X_1\rangle$ the rising part of W^P is precisely balanced by the decrease in the falling part. For a non-linear transfer function this is in general not the case and changes in amplitude lead to changes in the average output. **c)**: The dose-amplitude function of $\langle X_1\rangle$ versus A_1. For a linear dose-response function the dose-amplitude function is constant, while for a non-linear dose-response function, the output is a function of the amplitude.

8.3 Multiplexing

After specifying the model with its components we characterize its multiplexing capacity. We use the formalism of information theory (see the section: *Information theory*) to quantify

this capacity. We define two measures: $I_1(X_1, A_1)$, the mutual information between the concentration X_1 and the amplitude A_1 of signal S_1, and $I_2(X_2, \mu_2)$, the mutual information between the concentration X_2 and the concentration level μ_2 of S_2. The capacity of the model can then be defined by the total information $I_T = I_1(X_1, A_1) + I_2(X_2, \mu_2)$ that is transmitted through the system. The mutual information depends both on the kinetic parameters of the model and on the input distribution. In a previous study we have shown that using only constant signals, a simple biochemical system is capable of simultaneously transmitting two bits of information [182]. This indicates that this system can multiplex two signals with two input states perfectly. Here we wondered whether we can increase this information capacity. Therefore, we increase the number of input states (N_A and N_μ) for both signals S_1 and S_2 in steps of 1, where we assume a uniform distribution of the states. Next we optimize the total mutual information over the kinetic parameters. For the optimization we use an evolutionary algorithm (see App. 8.A.5).

Maximal information transmission between S_1 and X_1 occurs if a specific concentration X_1 maps uniquely to a specific amplitude A_1. In a deterministic system every possible combination of inputs (A_1, μ_2) leads to a single output concentration X_1. Ideally, X_1 should be a function of A_1 only, that is, it should be independent of μ_2. However, due to the cross-talk it is not; this means that a given concentration X_1 can correspond to multiple values of A_1 (each corresponding to different μ_2 (see Fig. 8.6). Information transmission is enhanced when this overlap is minimized.

Biochemical systems are not deterministic but stochastic. Therefore we include the effects of noise. For oscillatory non-linear systems, analytical expressions for the noise can not be obtained. Therefore we follow a different route to approximate the noise strength. We calculate for the multiplexing system the noise strength using the linear-noise approximation [68] assuming constant input signals with strength μ_1, μ_2. The oscillations increase the variations in the response (which is a different source of variability than variations from the stochastic production and degradation). To prevent this from having a strong influence on our results, the response times of $\mathsf{X}_1, \mathsf{X}_2$ are longer than the oscillation period. The effect of the oscillations on the response is therefore time-integrated. Effectively, we set $m_{X_1} = m_{X_2} = \left(NT_p\right)^{-1}$ s^{-1}, such that the output averages over N oscillations with period T_p. We take $N = 10$.

We show the results in two different plots in Fig. 8.5, a deterministic network (top row) and a network with noise (bottom row). The total transmitted information does not reflect whether all information is transmitted reliably. For example, increasing the number of input states N_A can increase the mutual information $I_1(A_1, X_1)$ since the total input information has increased, but a specific output concentration X_1 can be less informative about a specific input amplitude. So maximal information transmission is not necessarily equal to maximal reliable information transmission. In Fig. 8.5 we therefore show precisely this quantity, the reliable information transmission for a deterministic network. Instead of the total mutual

information, the relative mutual information is

$$I_R\left(\left(A_1, X_1\right), \left(\mu_2, X_2\right)\right) = \frac{I_1\left(X_1, A_1\right)}{H\left(A_1\right)} + \frac{I_2\left(X_2, \mu_2\right)}{H\left(\mu_2\right)} \tag{8.14}$$

$$= \frac{I_1\left(X_1, A_1\right)}{\log_2\left[N_A\right]} + \frac{I_2\left(X_2, \mu_2\right)}{\log_2\left[N_\mu\right]}. \tag{8.15}$$

Note that $I_R((A_1, X_1), (\mu_2, X_2))$ has a maximum value of two, meaning that each channel transmits with 100% fidelity. In Fig. 8.5d,e,f the same results are shown, but now with additional noise.

In Fig. 8.5a the relative information for the channel $\mathsf{S}_1 \rightarrow \mathsf{X}_1$ is shown, in Fig. 8.5b for the channel $\mathsf{S}_2 \rightarrow \mathsf{X}_2$ and in Fig. 8.5c the sum of the two channels, all for deterministic networks. We observe that in a deterministic network the relative information is approximately 100%, independent of the number of input states N_A, N_μ. This is a remarkable result. We would expect cross-talk between the two channels to reduce information transmission. However, the results show that this cross-talk is very small. Only in the upper right corners, where the number of input states N_A and N_μ become large, do the simulation results show, a slight decrease in I_R. However, for the maximum number of input states, I_R is again 100%. Therefore the, scattered, decrease in I_R for values of N_A, N_μ does not reflect that maximal information transmission is not possible, but that the optimization has not reached convergence. The reason for this is that with a larger number of input states N_A, N_μ it becomes increasingly difficult to partition the phase space (set by the maximum attainable values of $\langle X_1 \rangle$ and $\langle X_2 \rangle$) in such a way that no cross-talk is present. Since $I_R = 100\%$ at the largest input states $N_A = N_\mu = \log_2(20)$, apparently, the parameter space is large enough that all input states can be divided over the phase space such that no overlap in the read-out $\langle X_1 \rangle$ is present. Phrased differently, given the observation of a specific $\langle X_1 \rangle$ the network can uniquely infer what the input signal A_1 has been, although every A_1 in itself can, due to the influence from the different states of S_2 give rise to multiple values of $\langle X_1 \rangle$ (see Fig. 8.6a). Either a decrease of the phase space, and/or an increase in the number of input states is required to observe the reduction in I_R due to the cross-talk between the two channels in a deterministic system. In terms of total transmitted information, the network is capable of transmitting at least $\log_2(16) + \log_2(16) = 8$ bits with absolute fidelity in the deterministic limit.

In the bottom row in Fig. 8.5 we again show the relative information I_R, but here for a system with noise. The noise is approximated following the linear-noise approximation [68] for the network with constant inputs, and we thus neglect the effect on the noise strength either due to the non-linear nature of the network, or due to the increase in variability of the read-outs caused by the oscillatory input. These two effects are not expected to qualitatively change the following observations. The multiplexing network is asymmetric in its information transmission capacity. In our network, the average of V^P, $\langle V^P \rangle$, is independent of the signaling characteristics of S_1 (which are the amplitude A_1 and frequency T_1, see App. 8.A.2); Eq. 8.4 then shows that the output of the right channel, $\langle X_2 \rangle$, is independent of the left channel. Figure 8.5e indeed shows that the relative information transmitted via the $\mathsf{S}_2 \rightarrow \mathsf{X}_2$ channel, $I_2(X_2, \mu_2)/\log_2(N_\mu)$ hardly depends on N_A; the right channel is thus

not much affected by the cross-talk from the left channel. More importantly, the channel is also less susceptible to the influence of noise. The origin of this is that X_2 is, by construction, a linear function of S_2, allowing an ideal separation of the input states N_μ in phase space [56, 57]. The left channel, $\mathsf{S}_1 \rightarrow \mathsf{X}_1$, is more susceptible to noise (Fig. 8.5d). Indeed, an increase in the number of states N_A directly reduces the relative information considerably. The accessible phase space for this channel is apparently smaller than for the $\mathsf{S}_2 \rightarrow \mathsf{X}_2$ channel and therefor the influence of a small noise source directly leads to a reduction in I_R. The effect of the cross-talk from S_2 is clear in the regime $\log_2(N_A) \geq 3$, where an increase in N_μ leads to a further reduction of I_R. We do note that even in the presence of noise, maximal relative information is obtained for $N_A = N_\mu = 4 (= 2\,\text{bits})$ (Fig. 8.5f) showing that without loss of any information 4 input states in each channel simultaneously can be transmitted.

8.4 Experimental observations

Here we connect our work to two biological systems. The first system is the p53 DNA damage response system. The p53 protein is a cellular signal for DNA-damage. Multiple forms of DNA damage exist and they lead to different temporal profiles of the p53 concentration. Double stranded breaks cause oscillations in the p53 concentration, while single stranded damage leads to a sustained p53 response [187, 204, 205]. Compared to our simple multiplexing motif, the encoding scheme in this system is more involved. In our system two external signals activate the shared component V. In the p53 system, p53 itself is V, but interestingly, negative (indirect) autoregulation of p53 is required to obtain sustained oscillations. To be more precise, without interaction with p53, the external signal due to a double stranded break is constant.

Although the encoding structure is different, the main result is that the system is able to encode two different signals into different temporal profiles simultaneously; depending on the type of damage either a constant and/or an oscillatory profile of p53 is present. These two signals can therefore be transmitted simultaneously due to their difference in the temporal profiles as we have shown in section Sec. 8.3. For the p53 system the input signals are binary, e.g. either there is DNA damage or not, although some experiments suggest that the amount of damage also could be transmitted [160]. The maximum information that can be transmitted following our simplified model is much larger than that required for two binary signals. A mathematical model, based upon experimental observations, shows that the encoding step creates a temporal profile for p53 that could be decoded by our suggested decoding module (not shown, to be published).

Another system of interest is the MAPK (or RAF-MEK-ERK) signaling cascade. The final output of this cascade is the protein ERK, which shuttles between the cytoplasm and the nucleus. ERK is regulated by many different incoming signals of which EGF, NGF and HRG are well-known [206]. The temporal profile of ERK depends on the specific input that is present. NGF and HRG lead to a sustained ERK level [186], while EGF leads to a transient or even oscillatory profile of the ERK level [156, 186, 207]. Compared to our model ERK would be the shared component V. Experiments show that oscillations in the ERK concentration

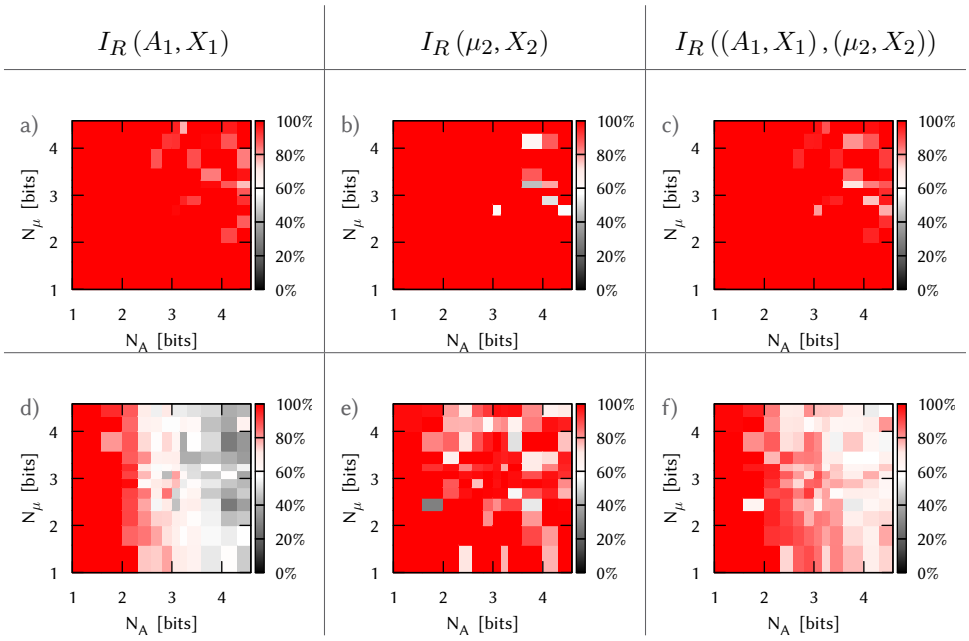


Figure 8.5: The transmitted relative information I_R (Eq. 8.14) as function of the number of input states N_A,N_μ, where 2 bits correspond to 2^2=4 input states. In panels **a,b,d,e** 100% corresponds to I_R=1, while in **c,f** 100% corresponds to I_R=2. Panels **a,b,c** for a deterministic system, panels **d,e,f** for a stochastic system. **a)** the relative mutual information $I_R(A_1,X_1)$ for the $S_1 \to X_1$ channel in a deterministic network is shown. The total mutual information is obtained by multiplying I_R with $\log_2(N_A)$, the horizontal axis. For (almost) all input states perfect information transmission is obtained. **b)** the relative mutual information $I_R(\mu_2,X_2)$ for the $S_2 \to X_2$ channel in a deterministic network is shown. The total mutual information is obtained by multiplying I_R with $\log_2(N_\mu)$, the vertical axis. For all input states perfect information transmission is obtained. **c)** the relative information of the total network $I_R((A_1,X_1),(\mu_2,X_2)) = I_R(A_1,X_1) + I_R(\mu_2,X_2)$. The noise strength σ_{X_i} are obtained following the linear-noise approximation for the network with constant input states. The noise strength sets the spread of each component around its deterministic mean, e.g. $X_{1,\text{noise}} = \langle X_1 \rangle \pm 2\sigma_{X_1}$. **d)** the relative mutual information $I_R(A_1,X_1)$; both the decrease in $I_R(A_1,X_1)$ as a function of N_A due to the presence of biochemical noise, and the decrease in $I_R(A_1,X_1)$ as a function of N_μ for $N_A > 3$ is observed due to the presence of the cross-talk is observed. **e)** the relative mutual information $I_R(\mu_2,X_2)$; the effect of noise is relatively small on information transmission. **f)** the relative information of the total network $I_R((A_1,X_1),(\mu_2,X_2))$. All results are obtained through numerical optimization (see App. 8.A.5). The scatter in the data reflects that the optimization procedure has not converged. The results should therefore be interpreted as lower bounds on information transmission.

can arise due to intrinsic dynamics of the system. However, they could also arise due to oscillations in the signal EGF, especially since, to our knowledge, it is unclear what the temporal behavior of EGF is under physiological conditions. Intriguingly, the fact that it can oscillate with constant inputs, suggests that it has band-pass characteristics, making the

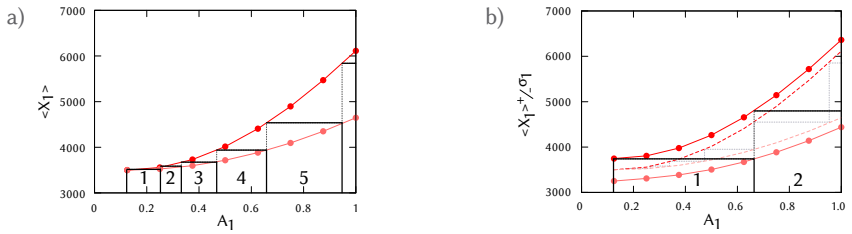


Figure 8.6: Schematic representation of the influence of noise and cross-talk on the information transmission in pathway $S_1 \rightarrow X_1$. **a)** Even in the absence of noise for every A_1 multiple values of $\langle X_1 \rangle$ are obtained, each corresponding to a specific value of μ_2. The dark red corresponds to the maximum value of $\langle X_1 \rangle$ for each A_1, while the light red line is the minimum value. The black line in between the red lines visualizes the range for which a specific $\langle X_1 \rangle$ uniquely maps to a single input amplitude A_1. In the deterministic limit, for this dose-amplitude response, 5 input states N_A can be transmitted with 100% fidelity, as indicated by the boxes. The number of states is only limited by the cross-talk from the $S_2 \rightarrow X_2$ channel. **b)** The influence of noise reduces the number of possible input states. Shown are the results for the dose-response curve as in panel **a**, but now with biochemical noise present. Dotted lines give the deterministic results, while solid lines are for a network with noise (color as in panel **a**). Since for each A_1 a larger range of $\langle X_1 \rangle$ values is obtained, less states A_1 can be uniquely encoded in the phase space. This is reflected in the increase in the width of the boxes; indeed, here only two input states can be transmitted with absolute reliability.
Please note that these results are obtained for simulations which are optimized for information transmission in the presence of noise. If the network were optimized in the deterministic limit, a dose-amplitude curve would be obtained which would allow for the reliably transmission of a larger number of input states. Please note that the dose-response curve of X_2 is a simple linear function and that the channel $S_2 \rightarrow X_2$ is not influenced by S_1. Results are therefore not shown.

system an ideal candidate for multiplexing.

For both experimental systems, we have only described the encoding step. In both cases, two signals are encoded in a shared component V, where one signal leads to a constant response, while the other signal creates oscillations. Both p53 and ERK are transcription factors for many downstream genes [208, 209]. For the decoding of the constant signal, a simple birth-death process driven by V is required. Many genes are regulated in this way. The decoding of the oscillatory signal requires an adaptive motif. Although adaptive motifs are common in biological processes, it is unclear whether downstream of either p53 or ERK an adaptive motif is present, which would complete our suggested multiplexing motif.

8.5 Discussion

We have discussed a system consisting of elementary motifs, which can simultaneously transmit two signals reliably. One of these signals is constant, and its corresponding information is encoded in its concentration level, while the other signal is dynamic, and its information is encoded in the dynamical properties but not in its average concentration level. The decoding of the constant signals for the constant signal is performed by a time-integration motif, while the decoding of the oscillatory signal requires a frequency sensitive

motif, for example an adaptive motif.

The system is capable of transmitting over 7 bits of information. In the presence of noise the information transmission decreases, but still with considerable noise levels, more than 4 bits of information can be transmitted. To transmit signals without errors it is preferable to send most information using the continuous channel and a smaller number of states through the frequency encoded channel. The reason for this is twofold. First, the amplitude encoded signal is less noisy, since the number of intermediate steps is smaller *and* second the frequency encoded signal is corrupted by cross-talk from the amplitude encoded signal, leading to overlaps in the state space of A_1 as a function of μ_2 (see Fig. 8.6). Nonetheless, the two channels can transmit in the presence of noise three or more states. Taken together, this is a considerable increase in the information transmission, compared to a system where both signals are constant [182], which can transmit two binary signals with absolute fidelity. This suggests that oscillatory signals can enhance the information transmission capacity of biochemical systems.

The main problem in multiplexing biochemical signals is cross-talk between the two signals. In this system the signals are encoded based upon their dynamical profile — S_1 is oscillatory and S_2 is constant. The decoding module for the frequency modulated signal, an adaptive motif, is non-linear. Therefore, this motif is not only sensitive to the temporal properties like the amplitude, but is also sensitive to the mean or average of its input. This inevitably leads to cross-talk between S_1 and S_2 of the output X_1 and this is a limitation for information transmission.

In this system we have assumed that the amplitude of the oscillatory signal is the information carrier of one signal. Equivalently, the same analysis can be performed for a signal at constant amplitude but at different frequencies. Qualitatively, the results will be similar. From the analysis of the gain one observes that varying the frequency of the input signal, changes the amplitude at the output. The amplitude of the output thus characterizes the signal frequency. However, an intrinsic redundancy is present in using the frequency as the information carrier, which can be understood from the symmetry of the gain. The response of the system is equal for frequencies that are positioned symmetrically with respect to the resonance frequency. As a result, for any given output, there are always two possible input frequencies, and without additional information, the cell can not resolve which of the two frequencies is present. Of course, one way to avoid this, would be to use only a part of the gain, in which the gain increases monotonically with frequency.

Besides the possibility of multiplexing, the use of oscillatory signals has other advantages. Oscillatory signals minimize the prolonged exposure to high levels of calcium, which can be toxic for cells [161]. In systems with cooperativity [162] an oscillating signal effectively reduces the signal threshold for response activation. Pulsed signals also provide a way of controlling the relative expression of different genes [159]. Encoding of stimuli into oscillatory signals can reduce the impact of noise in the input signal and during signal propagation [163]. Frequency encoded signals can be decoded more reliably than constant signals [203].

In this study we have assumed that the input signals are deterministic. Results are obtained following deterministic simulations, where noise is added following a solution of

the linear-noise approximation assuming non-oscillatory inputs. The effect of noise is a reduction of the information transmission.However, the effect of noise can always be counteracted by increasing the copy number. At the cost of producing and maintaining more proteins, similar results can therefore be obtained [182]. Next, the effect of oscillations on the variability of the output is small since the response times of X_1 and X_2 are much longer than the oscillation period. Slower responding outputs would time-average the oscillations cycle even more, reducing the variability in the response.

Interestingly, the output X_2 is a simple linear birth-death process driven by the input V^P. Comparing the dynamics of R and X_2, we note that these are similar, although they differ in their kinetic parameters. The response time of X_2 is much slower than the oscillation period to reduce the variability in the output. The response time of R has to be small, since that reduces the adaptation time. If the kinetic parameters of X_2 and R were equivalent, R could function as the read-out of S_2, which would reduce the complexity of the multiplexing motif.

The occurrence of temporal signals has driven many studies. Here we show that information can be encoded in the amplitude or frequency of oscillatory signals, which can then be decoded using a non-linear integration motif. Moreover, we show that oscillating signals are ideally suited for multiplexing; we give two biological systems that may have implemented this multiplexing strategy. The idea to use the temporal kinetics as the information carrier in a signal has been studied in a slightly different context, where the dose information is encoded in the duration of the intermediate component V, which in turn is time-integrated by the response X_1 [210]. However, such a dose-response relation could also be obtained by a much simpler network. Here, we show that encoding signals into the temporal dynamics of a signaling pathway allows for multiplexing, making it possible to simultaneously transmit multiple input signals through a common network with high fidelity.

8.A Supplementary Information

We use the following two general definitions for the mean and the maximum of a specific component, if the period of the input signal is T_p

$$\langle Z \rangle = \frac{1}{T_p} \int_t^{t+T_p} Z\left(t'\right) dt', \tag{A8.1}$$

$$Z_{\max} = \sup_{T_p} Z\left(t\right). \tag{A8.2}$$

8.A.1 Encoding

MM-approximation

In the derivation of Eq. 8.2, we have assumed, as is commonly done, Michaelis-Menten (MM) kinetics. However, the MM-approximation may not hold for a dynamical system [211, 212]. Since the MM-approach is a coarse graining of the full mass-action kinetics of the system, we evaluate numerically whether the MM-approximation is valid. The full set of reactions is

$$\begin{aligned} \mathsf{S} + \mathsf{V} &\underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} \mathsf{SV}, \\ \mathsf{SV} &\xrightarrow{k_2} \mathsf{S} + \mathsf{V^P}, \\ \mathsf{V^P} + \mathsf{E} &\underset{r_{-1}}{\overset{r_1}{\rightleftharpoons}} \mathsf{V^PE}, \\ \mathsf{V^PE} &\xrightarrow{r_2} \mathsf{V} + \mathsf{E}, \end{aligned} \tag{A8.3}$$

where in the MM-approximation one assumes

$$\frac{dSV}{dt} = 0, \; \frac{dV^PE}{dt} = 0. \tag{A8.4}$$

We compare the results for the full system in which the phosphorylation reaction is linear in S and E, with the Michaelis-Menten approximation in the linear regime (Eq. 8.3), in which the phosphorylation reaction is linear in the enzyme concentration (Fig. 8.7a). This linearity implies $K_V \left(= \left(k_{-1} + k_2\right)/k_1\right) \ll V_T - V^P$, such that the phosphorylation reaction of V to $\mathsf{V^P}$ simplifies to $k_V \sum_i S_i\left(t\right)$, which is indeed linear in S_i (and zero-order in V). If $V \gg S$, this condition will also be fulfilled for all moments in time when the system is dynamical [211]. In this case all S directly binds to V and the complex SV is very stable.

A very stable complex, however, can influence the dynamics of the signal oscillations. Assume that the oscillations in the signal S are driven due to external factors (like hormone pulsing), or that the oscillations depend on the (saturated) degradation of the signal S [213]. The oscillating signal directly forms the stable complex and, since in these scenarios the complex is not regulated, the absolute signal level $S_T = S + SV$ increases with increasing

numbers of oscillations, until all V is saturated. This will influence the oscillatory dynamics of the signal, reducing its periodicity and strongly increasing the mean level of signal present(Fig. 8.7b). The interaction between the signal and V, due to the influence on the oscillations, corrupts the information that is encoded in the oscillations. We note here that to overcome this problem, instead of assuming $S \ll V$, a scenario is possible where $V \ll S$. However, in this regime it is unclear if the dynamical behavior of the MM-approximation accurately represents the dynamics of the full mass-action equations. Moreover, if $S \gg V$, a small concentration S directly saturates all the V molecules. As a result, the concentration of V is insensitive to any oscillation in the concentration S, when the minimum concentration S is larger than the concentration V.

To overcome this problem, we add a small extension to the model. This extension is biologically inspired, since many external signals are sensed by receptors Q, which in turn activate (or phosphorylate) intracellular proteins. The crucial ingredient is that the signal-bound receptor dissociates on a much faster timescale than the oscillations. Due to this very fast receptor dissociation, the signal-bound state is very small.

$$\begin{aligned}
\mathsf{S} + \mathsf{Q} &\underset{kr_{-1}}{\overset{kr_1}{\rightleftharpoons}} \mathsf{Q}^{\mathsf{A}}, \\
\mathsf{Q}^{\mathsf{A}} + V &\underset{kv_{-1}}{\overset{kv_1}{\rightleftharpoons}} \mathsf{VQ}^{\mathsf{A}}, \\
\mathsf{VQ}^{\mathsf{A}} &\overset{kv_2}{\rightarrow} \mathsf{V}^{\mathsf{P}} + \mathsf{Q}^{\mathsf{A}}, \\
\mathsf{V}^{\mathsf{P}} + \mathsf{E} &\underset{r_{-1}}{\overset{r_1}{\rightleftharpoons}} \mathsf{V}^{\mathsf{P}}\mathsf{E}, \\
\mathsf{V}^{\mathsf{P}}\mathsf{E} &\overset{r_2}{\rightarrow} \mathsf{V} + \mathsf{E},
\end{aligned} \tag{A8.5}$$

With this small extension, the dose-response curve is similar to the dose-response curve for the Michaelis-Menten approximation with small K_V (in the linear regime), also for oscillatory signals (Fig. 8.7c).

8.A.2 Linear Approximation

The MM-approximation, assuming a linear regime for $\mathsf{S}_1, \mathsf{S}_2$ in the phosphorylation of V (Eq. 8.3) does not change the dynamics of the signals $\mathsf{S}_1, \mathsf{S}_2$, which we demonstrated with simulations, as is described in App. 8.A.1. To show that the linear regime does not change the mean of V^{P}, $\langle V^P \rangle$ as function of the signal characteristics, we perform in this section numerical simulations and for completeness we compare the linear regime with two other regimes.

In regime A there are many more V molecules than kinase (S_i) and phosphatase molecules, meaning that the kinase and phosphatase enzymes are saturated ($V_T \to \infty$). As a result, the dynamics can be simplified, since saturation of kinase and phosphatase molecules

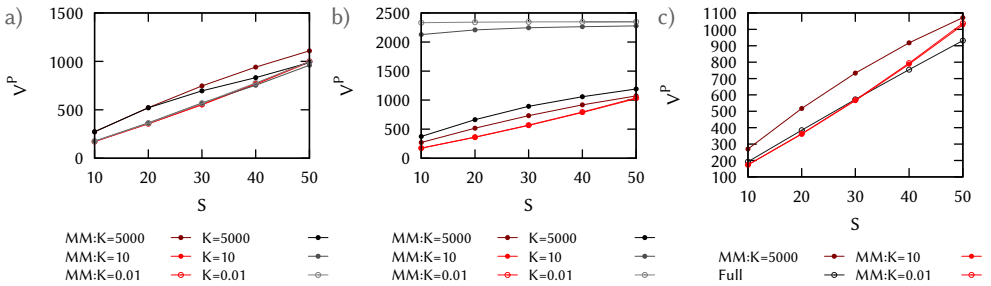


Figure 8.7: Comparison of the Michaelis Menten approximation (red) with the full mass action equation (Eq. A8.3, black) using Gillespie simulations. **a)** Dose response curve for constant signal S. The Michaelis Menten approximation approximates the non-coarse grained system very well (only for K=5000 the curves do not precisely overlap). **b)** Dose-response for sinusoidal input S with T=100 s. Due to the strong complex formation of VS, the signal S is not external to the system, but all the signaling molecules bind until all V is saturated. The effective concentration S is thus much larger than the mean of the oscillations. Parameters: r_1 s_{-1},E_T=150,V_T=2500,M_V=5000, and for K_W=5000,K_W=10,K_W=0.01 respectively $\{k_1,k_{-1},k_2,r_2\}$=$\{1,4997.5,2.5,1\}$ s^{-1},$\{1,9,1,2\}$ s^{-1},$\{100,0,1,2\}$ s^{-1}. **c)** The dose-response curve for (Eq. A8.5) compared to the Michaelis-Menten equations kr_1=0.05 s^{-1},kr_{-1}=4000 s^{-1},kv_1=1 s^{-1},kv_{-1}=4000,kv_2=25 s^{-1},r_2=2 s^{-1},r_1=1 s^{-1},r_{-1}=4998 s^{-1}(M_V=5000),E_T=50,R_T=1000,V_T=5000.

implies that $M_V, K_V \ll V_T$. This leads to

$$\frac{dV^P}{dt} \approx k_V \left(\sum_i S_i\,(t) \right) - m_V, \tag{A8.6}$$

This is the well-known regime of zero-order dynamics for V. In this regime $\langle V^P \rangle$ can be approximated by a binary function $\langle V^P \rangle = 0$ or $\langle V^P \rangle = V_T$, with the transition depending on the kinase(s) $\left(\sum_i S_i \right)_{\text{crit}}$ (see Fig. 8.8, open gray symbols). V thus acts as a switch. However, a switch-like functional dependence of V on S does not lead to perfect tracking of the signal S, and therefore not to reliable propagation of the oscillations.

The second regime, regime B, is the opposite of the previous. In this regime, V is limiting (e.g. $M_V, K_V \gg V_T$, see Eq. 8.2). The resulting dynamics is described by

$$\frac{dV^P}{dt} = k'_V \left(\sum_i S_i\,(t) \right) \left(V_T - V^P \right) - m'_V V^P, \tag{A8.7}$$

where $k'_V = k_V / K_V$ and $m'_V = m_V / M_V$. In this regime, the phosphorylation reaction is non-linear, but degradation is a linear process. A typical dose-response curve is shown in Fig. 8.8 (closed gray symbols).

In regime C (Fig. 8.8, closed black symbols), the two preceding regimes are combined. There is saturation of the kinases in the production, but saturation of V^P in the dephospho-

rylation, leading to Eq. 8.3

$$\frac{dV^P}{dt} = k_V \left(\sum_i S_i\left(t\right) \right) - m'_V V^P. \tag{A8.8}$$

In Fig. 8.8a the dose-response curve between S and V^P is shown, where we study the mean $\langle V^P \rangle$ as a function of a constant signal S with increasing mean μ. Regime C (closed red symbols) has an approximate linear relation between S and $\langle V^P \rangle$. Regime B (closed gray symbols) increases hyperbolically to saturation, while regime A (open gray symbols) shows the switch-like response.

A sinusoidal oscillation can only be propagated perfectly as a sinusoidal if the dose-response function is linear. For non-linear dose-response functions, oscillations with small amplitude are propagated correctly, since for small perturbations every function has linear characteristics. However, larger amplitude oscillations are deformed by the non-linear transfer function. As a result, the mean $\langle V^P \rangle$ changes as a function of A and/or T, the oscillation parameters. A stronger non-linear dose-response function decreases the amplitude-range of oscillations that can be propagated without this deformation. Figure 8.8b shows $\langle V^P \rangle$ for signals with different properties A and T. Since μ (the signal mean) is constant, $\langle V^P \rangle$ should be constant as well, if the transfer function would precisely track the signal dynamics. We observe that both in regime B and C $\langle V^P \rangle$ does not depend on the oscillation parameters A, T, while in regime A a strong dependence on these parameters exist. In Fig. 8.9a-c we show corresponding time traces. Again the strong non-linear response for regime A is observed, while regime B and C exhibit oscillations that are very similar to the signal sinusoidal oscillations. Please also note the reduction in amplitude in regime B, compared to regime C. This can be explained by the hyperbolic shape of the dose-response curve for regime B, which dampens changes in S (Fig. 8.8a).

8.A.3 Decoding

The adaptive motif

We take a closer look at the adaptive motif, especially focusing on the behavior of the average response $\langle W^P \rangle$ as a function of $V^P\left(t\right)$ in different parameter regimes. As discussed in the main text, an adaptive system is insensitive to a constant input concentration. In other words, the output of the adaptive system does not depend on the input, as long as the input is constant. For time-dependent inputs, however, this is in general not true.

In the multiplexing framework, V^P is characterized by the mean μ_1, frequency T_1 and amplitude A_1 of S_1 and the mean μ_2 of S_2. The response W^P should only depend on T_1 and A_1, since these carry the information in S_1, but be insensitive to the mean level of the signals S_1,S_2, and thus V^P. We restate the mathematical description of the adaptive motif

$$\frac{dR}{dt} = k_R V^P - m_R R, \tag{A8.9}$$

$$\frac{dW}{dt} = k_W \frac{V^P \left(W_T - W^P\right)}{K_W + \left(W_T - W^P\right)} - m_W \frac{R W^P}{M_W + W^P}. \tag{A8.10}$$

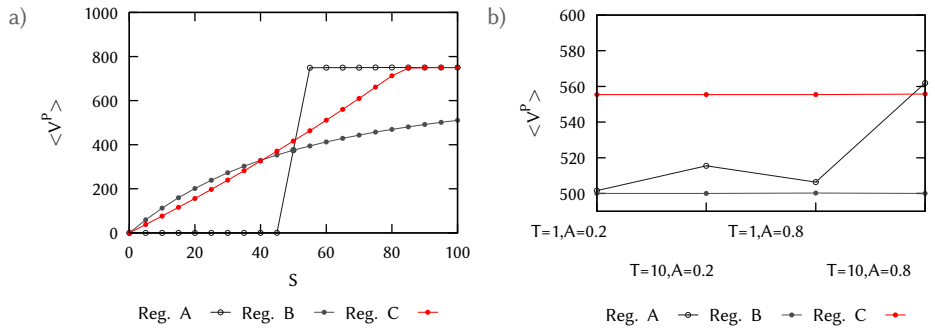


Figure 8.8: **a)** The dose-response curve is shown for the regime in which both the production and degradation are zero-order in V^P (regime A, Eq. A8.6), in which production is and degradation are linear in V^P (regime B, Eq. 8.3) and zero-order production but linear degradation of V^P (regime C,Eq. A8.7). The curve that is linear over the widest S-range is that for linear degradation of V^P, but zero-order production (regime C). Parameters: m_V=50,K_V=$10^{-4}V_T$,M_V=$10^{-4}V_T$, m_V=500,K_V=$10^{-4}V_T$,M_V=$5V_T$, and m_V=50,K_V=$5V_T$,M_V=$5V_T$. k_V sets the timescale. **b)** The time average over a single oscillation period, $\langle V^P \rangle_T$, is shown for four different simulations where the signal characteristics are as indicated and μ_S=50.

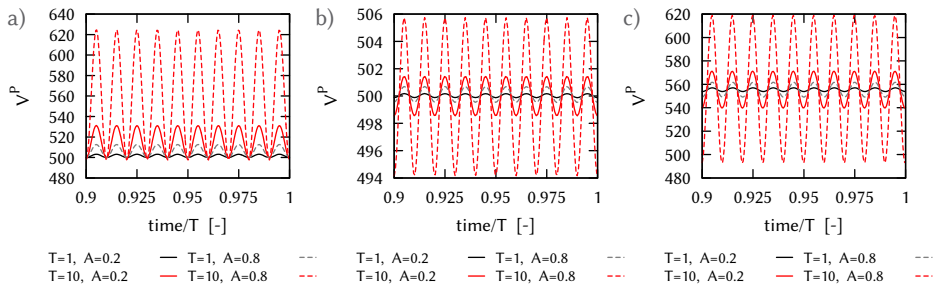


Figure 8.9: Time traces for regime A (panel **a**) zero-order for V in both production and degradation (Eq. A8.6), regime B (panel **b**) linear in V in both production and degradation (Eq. A8.7) and regime C (panel **c**) degradation linear in V and production zero-order in V (Eq. 8.3), all for a single driving signal with different amplitude A levels and frequencies T^{-1}, while μ=50. The non-linear response for the zero-order regime is clearly visible in **a**. The difference between panel **b** and **c** is the amplitude of the response. The process with a saturated production (B, panel **b**) has a much smaller amplitude than the linear production (C, panel **c**).

Similar to the discussion in App. 8.A.2 we distinguish four different regimes for the adaptive motif. Regime A is characterized by the full non-linear description Eq. A8.10. In regime B, where $K_W, M_W \gg W_T$, indicating that protein W is fully saturated, Eq. A8.10

simplifies to

$$\frac{dW^P}{dt} = \frac{k_W}{K_W} V^P \left(W_T - W^P\right) - \frac{m_W}{M_W} R W^P, \tag{A8.11}$$

$$\frac{\left(W_T - \langle W^P \rangle\right)}{\langle W^P \rangle} = \frac{K_W m_W}{k_W M_W} \frac{k_R}{m_R}. \tag{A8.12}$$

Eq. A8.11 describes bimolecular reactions between $\mathsf{V^P}$ and W, and R and $\mathsf{W^P}$ respectively; note that the the production rate of $\mathsf{W^P}$ is bounded by the total amount of unphosphorylated W. More importantly, we restate that adaptation requires that the deactivation reaction of $\mathsf{W^P}$ is non-linear, scaling as $R \times W^P$. This is the origin of the non-linearity in the network, turning multiplexing into a non-trivial problem.

In regime C, in a sense the opposite of regime B, both $\mathsf{V^P}$ and R operate in a saturated state ($K_W \ll W_T - X_1^P, M_W \ll X_1^P$), and this leads to

$$\frac{dW^P}{dt} = k_W V^P - m_W R, \tag{A8.13}$$

where the dynamics of $\mathsf{W^P}$ is independent of $\mathsf{W^P}$ itself, but does depend strongly on changes in $\mathsf{V^P}$ and R. This corresponds to the zero-order regime, with an ultra-sensitive response of $\mathsf{W^P}$ to changes in the ratio $\mathsf{V^P}$/R. This regime is only valid in a small concentration range of $\mathsf{V^P}$, due to the copy number constraint. If $W^P \sim W_T$ or $W^P \sim 0$, the approximation fails.

We stress again that we can not construct a linear regime for $\mathsf{W^P}$, equivalent to Eq. 8.3 for $\mathsf{V^P}$. This is because, in contrast to Eq. 8.2, in Eq. A8.10 the dephosphorylation reaction is non-linear. The dephosphorylation of $\mathsf{V^P}$ is dictated by a constant number of enzymes, while the dephosphorylation of $\mathsf{W^P}$ depends on $R\,(t)$, which oscillates. In regime D, we assume a functional form which is most similar to Eq. 8.3, meaning that the activating enzyme ($\mathsf{V^P}$) is saturated, while the deactivating enzyme (here R) is not,

$$\frac{dW^P}{dt} = k_W V^P - \frac{m_W}{M_W} R W^P, \tag{A8.14}$$

$$\left\langle W^P \right\rangle = \frac{k_W M_W m_R}{m_W k_R}. \tag{A8.15}$$

For these four regimes we show the dependence of $\langle W^P \rangle$ on k_W at fixed μ_1, with $\mu_2 = 0$ (Fig. 8.10a), the gains (Fig. 8.10b) and typical time traces (Fig. 8.11).

The dose-rate curve of $\langle W^P \rangle$ vs k_W (Fig. 8.10a) shows for a constant signal S_1 the change in $\langle W^P \rangle$ with a change in k_W. From Eq. A8.10, we observe that a change in k_W is equivalent to a change in $\langle V^P \rangle$. Since $\mu_2 = 0$, the dose-rate curve therefore is equivalent to a dose-response curve of $\langle W^P \rangle$ vs the men of S_1, μ_1. Note, that for an adaptive regime indeed the response is independent of the signal, but not independent of the kinetic rates of the adaptive motif itself, therefore a dose-rate curve can be obtained. However, for constant k_W, the dose-rate curve indicates the actual magnitude of the transient response $\mathsf{W^P}$ to a change in the input. Regimes A and B have a similar dose-rate curve, which shows a

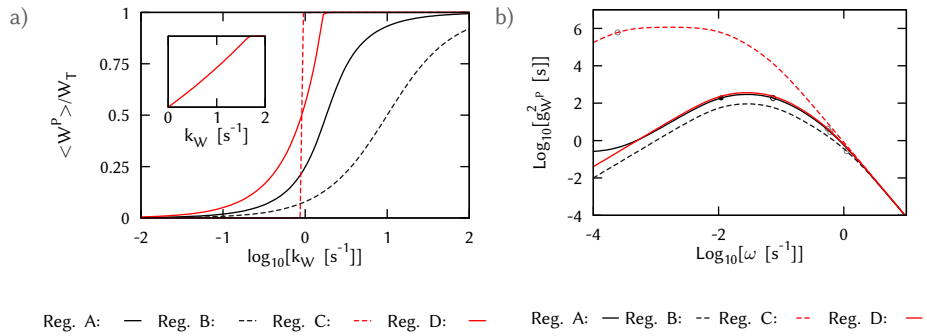


Figure 8.10: **a)** For Eq. A8.10 (black solid), Eq. A8.11 (black dashed), Eq. A8.13 (red dashed) and Eq. A8.14 (red solid) the dose-rate curve is shown for μ_1=50,μ_2=0. The very sharp dependence on k_W of the completely saturated regime C is shown. Note that in regime D $\langle W^P \rangle$ is linear in k_W over a wide range (see inset), making this an ideal candidate for multiplexing. Parameters: K_V=0.004V_T,m_V=10 s^{-1},M_V=30V_T,V_T=2500,m_R=0.01 s^{-1},k_R=0.011 s^{-1},W_T=1000; m_W=2(W_T/2.+M_W)/W_T s^{-1}. Reg. A: K_W=M_W=W_T/2, Reg. B: K_W=M_W=5W_T, Reg. C: K_W=M_W=$10^{-4}W_T$, Reg. D: K_W=$10^{-4}W_T$,M_W=5W_T. **b)** The power spectrum (equal coloring) for equal $\langle X^P \rangle$. Indicated with circles are τ_R (black solid), τ_V (black open) and τ_X (gray open). A change K_W,M_W only changes τ_X. Since $\tau_X^{-1} \gg \tau_R^{-1}, \tau_V^{-1}$, the resonance frequency is similar for regime A,B and D. Regime C has a complete different power spectrum. The larger amplitude reflects the switching behavior of the system. Next, the resonance frequency is present at smaller frequencies. The peak is less pronounced, reflecting that the switching behavior is not dependent on a specific signal frequency, but is present for many different signal inputs. Parameters: as in panel **a** except for k_W, which sets $\langle X_1^P \rangle$=W_T/2.

gradual non-linear increase of W^P as a function of k_W. Regimes C, where the kinase and phosphatase are saturated in the (de)phosphorylation reactions has a much stronger non-linear response function and shows a switch transition at a critical k_W. Regime D is linear in W^P as function of k_W. This indicates that this regime is a good candidate for multiplexing.

The gains (Fig. 8.10b) of regime A,B and D are similar, while for regime C the gain is much larger. The peak is at a smaller frequency compared to that of the other regimes and less pronounced.

The time traces (Fig. 8.11a-d) show some interesting features. First, close to the resonance frequency (ω_{res}) (Fig. 8.11a,c) the oscillations in W^P are large, while signal oscillations that are much faster are not propagated at all, but averaged out (Fig. 8.11b,d). Next, the amplitude of the response W^P has a strong dependence on the signal amplitude (A_1). The systems with non-saturated dynamics for the kinase and phosphatase (Eqs. A8.10, A8.11) show clear oscillatory behavior, although the shape of the oscillations is strongly altered with respect to the shape of the incoming sinusoidal function due to the non-linear transfer function (dose-response function). A closer comparison with Fig. 8.10a shows that the dose-response curve is steeper in regime A than in regime B. As a result, a small change in V^P leads to larger change in W^P in regime A. The oscillations in regime A therefore have a larger amplitude than in regime B. The (binary) switch characteristic of the dose-rate curve of regime C (Eq. A8.13) is clearly seen. Next note that, except for regime D, the oscillations

are (strongly) asymmetric. The decrease in the minimum with respect to the steady-state average $\langle W^P \rangle$ is stronger than the increase in the maximum.

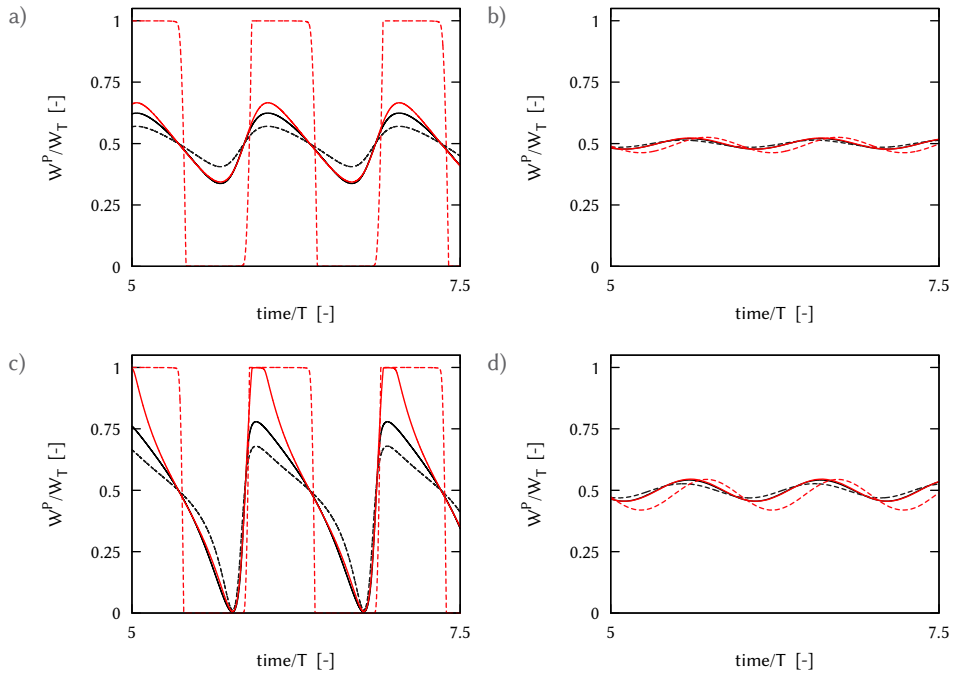


Figure 8.11: Respective time traces for different frequencies T^{-1} and amplitudes A of S_1, shown for the four different regimes A-D, colors and parameters as in Fig. 8.10b (black solid, regime A, black dashed, regime B, red dashed, regime C, red solid, regime D). Panels **a**, **c** T=10 s, panels **b**, **d** T=1000 s. Panels **a**, **b** A=0.5, **c**, **d** A=1.

The behavior of $\langle W^P \rangle$ as function of A_1 and μ_2

As discussed in the main text the potential cross-talk between S_1 and S_2 is an important problem. Since the transfer function from V^P to W^P is non-linear, a change in $\langle V^P \rangle$ will influence the response W^P. Indeed, in Fig. 8.11a,c the non-linearity of the response given a sine-input is clearly present. A very important consequence of this nonlinearity is that the mean response $\langle W^P \rangle$ is a function of $\langle V^P \rangle$. Although W is adaptive for steady-state inputs, the mean is dependent on dynamic inputs.

In Fig. 8.12 the mean response $\langle W^P \rangle$ and the maximum expression $W^P_{\max}$ are shown for specific characteristics of S_1 (A_1, T_1 and constant μ_1), but changing μ_2 for the different parameter regimes A, B, C, D (as defined earlier). The mean response $\langle W^P \rangle$ close to the resonance frequency ($T_{\text{res}} \approx 1000\,s$) increases with increasing μ_2, but constant A_1 (Fig. 8.12b, d). As important, also the maximum response is not constant, but decreases as a function of μ_2. Both effects are larger for larger A_1. This indicates, that for a specific A_1, T_1 a range of maximum and average W^P concentrations is observed as function of μ_2, and therefore a

unique mapping between W^P and A_1 is not present. Indeed, for every A_1 there are multiple outputs W^P, each corresponding to a different value of μ_2. This cross-talk between μ_2 and A_1 is precisely the cross-talk that a multiplexing system should minimize to obtain reliable signal transmission performance.

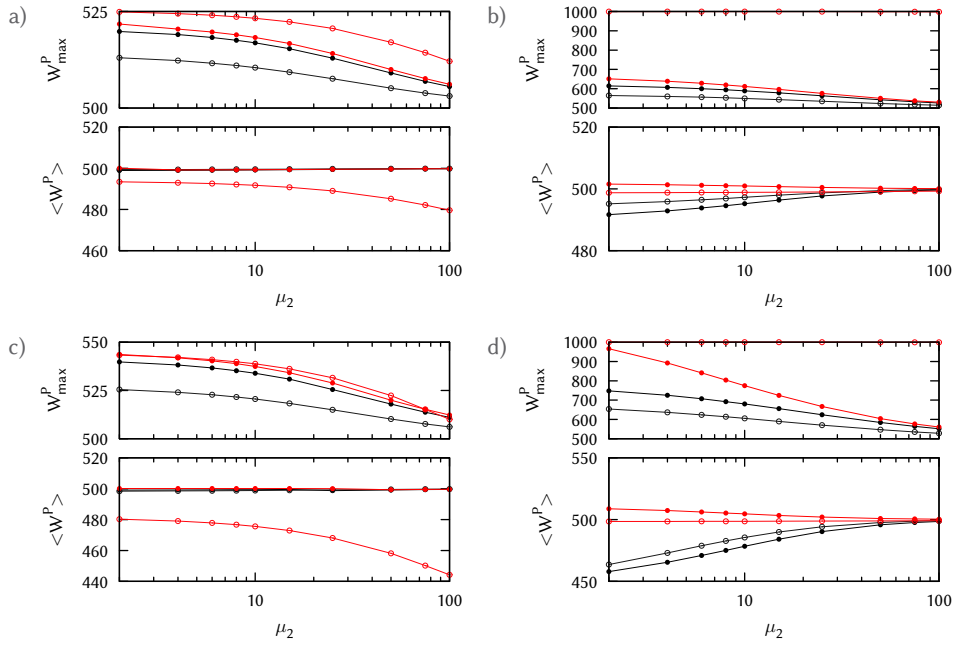


Figure 8.12: Dose-response curves for $\langle W \rangle$ (bottom) and W^{max} (top) as function of μ_2, for constant μ_1=25. Color for the four different regimes A-D as in Fig. 8.10b (black solid, regime A, black dashed, regime B, red dashed, regime C, red solid, regime D). Panels **a**, **c** T=10 s, **b**, **d** T=1000 s. Panels **a**, **b** A=0.5, **c**, **d** A=1.

As important is the fact that the non-linear transfer function between V^P and W not only creates a dependence of $\langle W^P \rangle$ on $\langle V^P \rangle$, but also a dependence of $\langle W^P \rangle$ on A_1. In other words, even for a signal with constant mean, the mean of the output is not constant, but depends on the amplitude. Indeed, for constant mean $\langle V^P \rangle$, $\langle W^P \rangle$ changes as a function of the amplitude (Fig. 8.13a, b). The direction of the change depends on the specific regime, in regime A $\langle W^P \rangle$ decreases with increasing amplitude, while in regime D $\langle W^P \rangle$ increases with increasing amplitude. In Fig. 8.13a, b the mean response $\langle W^P \rangle$ and the maximum expression W^P_{max} are shown as a function of amplitude A_1 for different mean μ_2 and T_1 for regime A,D.

In Fig. 8.13c,d the corresponding overlap plots are shown for regime D. These capture the described sensitivity to both A_1 and μ_2. In the preceding paragraph we have shown that $\langle W^P \rangle = f(A_1, \mu_2)$ and $W^P_{\text{max}} = g(A_1, \mu_2)$: both are functions of A_1 and μ_2. In the plots the range of concentrations W^P is shown as a function of the amplitude A_1 defined over all μ_2. Each horizontal line describes $\langle W^P \rangle$ as a function of μ_2 for a given A_1, where

μ_2 is varied either in the range $[2:25]$ (red dashed) or in the range $[2:50]$ (red solid). In red lines the result for $T = 1000\,\mathrm{s}$ is shown, while the black line correspond to $T = 10\,\mathrm{s}$ and a μ_2-range of $[2:50]$. From Fig. 8.13c,d we note a couple of interesting observations. First, we observe the increase in the minimum of $\langle W^P \rangle$ for increasing A_1, which reflects the increase in the mean with increasing amplitude. Second, the presence of cross-talk between A_1 and μ_2 in the overlap plots is visualized by an overlap between the concentration ranges of μ_2. Within the overlap regime, the specific concentration W^P corresponds to different values of A_1 and therefore a unique mapping is not present. The overlap depends strongly on the resonance frequency and on the number of inputs N_μ of S_2. Since the states are uniformly spaced within a fixed interval ($A : [0:1]$, $\mu_2 : [2 : \mu_{2\,\max}]$), the number of states determines the input distribution. A reduction in the number of input states, increases the separation between each state in phase space and thus reduces the overlap and thus the cross-talk between the left and right pathway. Of course, in the limit $N_\mu = 1$, cross-talk is absent.

In order to prevent cross-talk between the two signals at the amplitude read-out, the mean response $\langle W^P \rangle$, should be independent of μ_2, while W^P should discriminate between different amplitudes. We quantify this trade-off as a function of the parameters K_W, M_W using the mutual information. We define

$$I_1 = I\left(W^P_{\max}, A_1\right) \text{ and } I_2 = I\left(\left\langle W^P \right\rangle, A_1\right). \tag{A8.16}$$

For maximal discrimination of the amplitude, I_1 should be maximized, while no response of the mean $\langle W^P \rangle$ with respect to the amplitude indicates that I_2 should be minimized. In Fig. 8.14 we show therefore $I_T = I_1 - I_2$, where large I_T reflects good amplitude discrimination while having a relatively constant mean response. From Fig. 8.14 we observe that the mutual information is maximized for $K_W \ll W_T$ and $M_W \gg W_T$, which corresponds to regime D. This coincides with the behavior we have observed from the overlap and dose-amplitude figures. Indeed, these figures show as well that large amplitude discrimination $W^P_{\max}$, while preserving constant $\langle W^P \rangle$, is optimal in regime D.

8.A.4 The conversion of W^P to X_1

In App. 8.A.3 we have discussed that due to the non-linearity of the adaptive motif, $\langle W^P \rangle$ is not independent of the temporal profile of $V^P(t)$, since $\langle W^P \rangle$ depends both on the maximum and minimum expressions and on the average expression. The fact that the mean $\langle W^P \rangle$ is not constant as a function of the amplitude A_1, has an influence on the suggested conversion from W^P to X_1.

As discussed in the main text, if $\langle W^P \rangle$ would be constant as function A_1, the transfer function between W^P and X_1 should be non-linear to optimally transfer the information encoded in the amplitude. However, since $\langle W^P \rangle$ depends on A_1 this general analysis is not necessarily correct. In Fig. 8.15 the dose-amplitude relation is shown for a system where the mean $\langle W^P \rangle$ changes with varying A_1 (compare with Fig. 8.4). Since a linear response function is only sensitive to the mean input, and not to the minimum and maximum response (for a temporal profile that is symmetric with respect to the mean), the output $\langle X_1 \rangle$

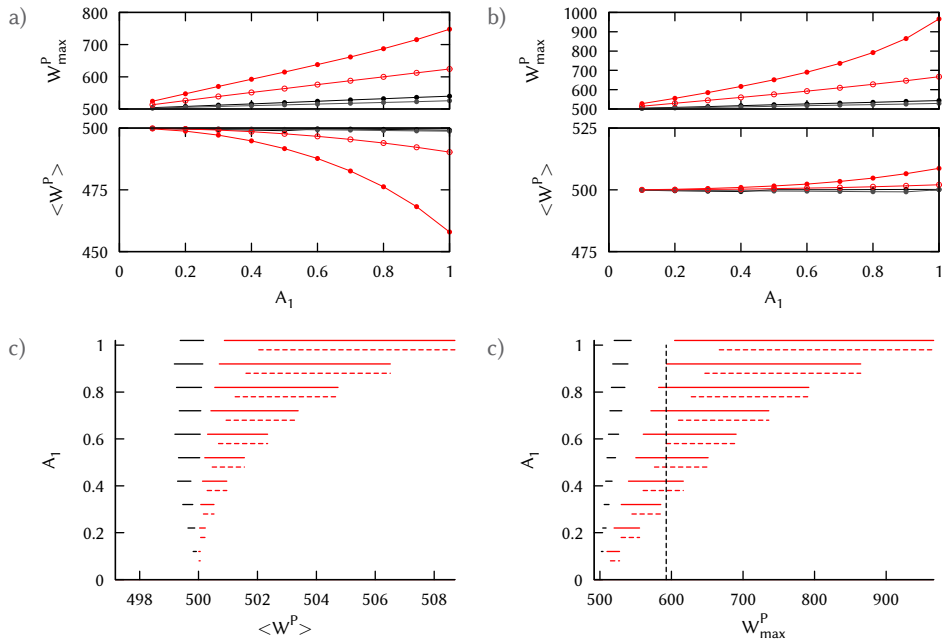


Figure 8.13: Panels **a**, **b** Dose-amplitude curves for $\langle W \rangle$ (bottom) and W^{max} (top) as function of A_1, for constant μ_1=25. Color code: black T=10 s, red T=1000 s. Closed circles: μ_2=2 au, Open circles μ_2=25 au. In regime A (panel **a**) for T=10 s (black) there is almost no change for increasing A. For T=1000 s (red), the mean $\langle W^P \rangle$ decreases, while the maximum response W^P_{max} increases. In regime D (panel **b**) similar behavior with respect to T is observed, although the mean $\langle W^P \rangle$ here increases, but the increase is relatively small and the maximum response has a stronger increase with respect to A_1. This is because regime D has a linear response over a larger range of inputs $\langle Y^P \rangle$ (see Fig. 8.10a). Panels **c**, **d**: Overlap plots for regime D. We show the range of the response $\langle W^P \rangle$ for a given amplitude A_1 over all input μ_2. Black color for T=10 s, red color for T=1000 s. Thick red μ_2=[2−50] au, dashed red μ_2=[2−25] au. The overlap between the different lines visualizes the redundancy (or non-uniqueness) of a specific response. The vertical dashed black line shows that a specific output W^P_{max} corresponds to multiple inputs A_1. For the dashed red line, with a smaller number of S_2 inputs, the dashed black line crosses a smaller number of lines as the range in μ_2 decreases. Therefore the two red lines show the reduction of overlap (or cross-talk) for increasing $\mu_{2,max}$ (solid $\mu_{2,max}$=50 au, dashed $\mu_{2,max}$=25 au). Note the scale on the horizontal axis in panel **c**, the change in $\langle W^P \rangle$ is negligible.

then is a function of A_1. A unique mapping between $\langle W^P \rangle$ and A_1 thus follows for a linear relation between W^P and X_1. Note that this is in contradiction with the analysis in the main text, where the linear relation between W^P and X_1 has no influence given a constant mean $\langle W^P \rangle$ and symmetric temporal profile. A non-linear transfer function between W^P and X_1 depends on the maximum W^P_{max}, the minimum W^P_{min} and the change in $\langle W^P \rangle$. If the mean $\left\langle W^P \right\rangle$ is a function of the amplitude A_1, due to the non-linearity of the transfer function, the mean $\langle X_1 \rangle$ depends, non-linearly, on the mean $\langle W^P \rangle$. This additional

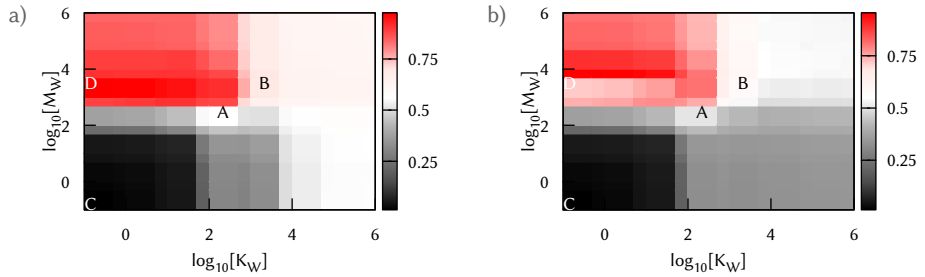


Figure 8.14: Contourplot for the mutual information (I_T=I_1−I_2, Eq. A8.16) as function of the parameters K_W,M_W, color indicates I_T. The input distributions are uniformly distributed with μ_2^j=50j/N_μ,A_1^i=(i−1)/(N_A−1). Parameters as in Fig. 8.10b and T=1000 s. Panel **a** N_A=4,N_μ=2 **b** N_A=4,N_μ=8.

dependency of $\langle X_1 \rangle$, next to the dependency on $W^P_{\max}, W^P_{\min}$, could corrupt the amplitude discrimination as suggested in the main text. We observe in Fig. 8.11b,d that the temporal profile of $W^P(t)$ is not symmetric with respect to time or the mean $\langle W^P \rangle$. Indeed, the minimum amplitude can be "larger" than the maximum amplitude, which provides an explanation why the time averaged mean $\langle W^P \rangle$ decreases. This also suggests that $\langle X_1 \rangle$ as a function of A_1 can decrease with increasing amplitude A_1, since the increase in amplitude creates a more asymmetric $W^P(t)$, which is biased to smaller values.

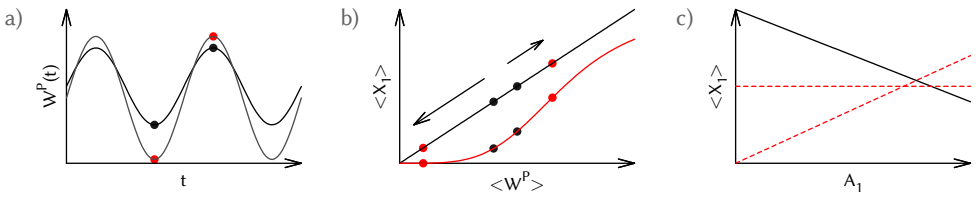


Figure 8.15: Schematic drawing of the decoding of the final step for a change in $\langle W^P \rangle$ for different amplitudes. **a)** the oscillations in W are shown for two different input amplitudes. An increase in the amplitude leads to a decrease in $\langle W^P \rangle$. **b)** The dose-response function of X_1 versus W. For a linear dose response function, a change in the amplitude of W does lead to a change decrease in the output X_1, since $\langle W^P \rangle$ decreases. For a non-linear transfer function the influence of the decrease of $\langle W^P \rangle$ depends on the precise form of the dose-response function, since the decrease in $\langle W^P \rangle$ has an opposite effect of the increase in the amplitude. **c)** The dose-amplitude function of X_1 versus A_1. For a linear dose-response function the dose-amplitude function is decreases, while for a non-linear dose-response function, the dose-amplitude function depends on the precise shape of the dose-response function.

The story is more intriguing, since $\langle W^P \rangle$ depends not only on A_1, but also on μ_2. Maximal information transmission occurs if the transfer function can discriminate different output levels maximally with respect to the corresponding inputs. Since it is difficult to analytically derive expressions, we perform numerical simulations where we study the total mutual information transmitted between two pathways as a function of the parameters K_W, M_W

(Eq. A8.10).

$$\frac{dX_1}{dt} = k_{X_1} W^P(t) - m_{X_1} X_1(t), \tag{A8.17}$$

$$\frac{dX_1}{dt} = k_{X_1} \frac{\left(W^P\right) n_{X_1}(t)}{\left(W^P\right)^{n_{X_1}}(t) + K_{X_1}^{n_{X_1}}(t)} - m_{X_1} X_1(t). \tag{A8.18}$$

In Fig. 8.16a,b the results for Eq. A8.17 are shown, while Fig. 8.16c,d the results for Eq. A8.18 are shown. We observe that the non-linear transfer function has a larger total information transmission, than a linear transfer function between W^P and X_1. Compared to Fig. 8.14, we note that optimal results are obtained for a larger range of parameters. First, regime D (Eq. A8.14) still provides correct results. For $K_W, M_W \gg W_T$ the oscillations are strongly damped and this reduces the transmitted information. For $K_W, M_W \ll 1$, W has a switch like response and this again strongly reduces the information transmission.

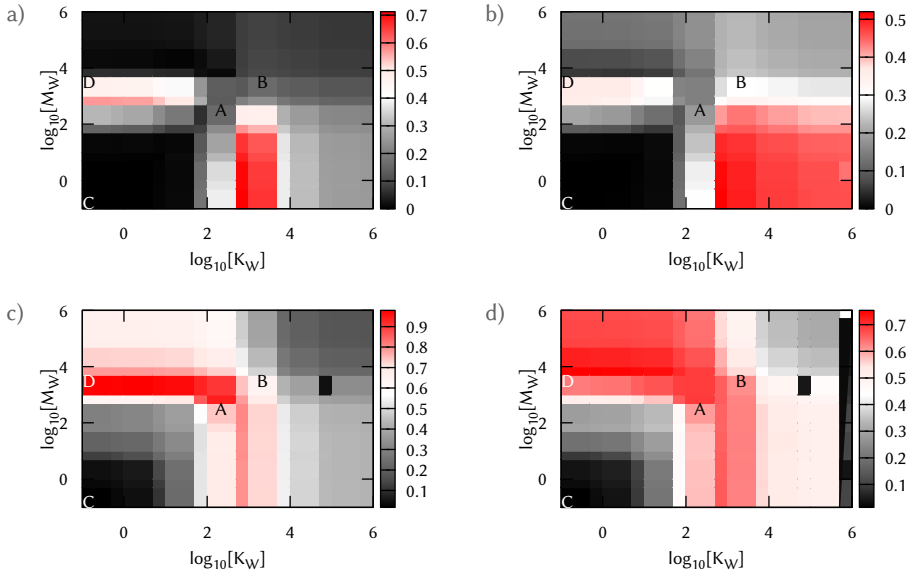


Figure 8.16: Contourplot for the mutual information ($I_1(X_1,A_1)$) as function of the parameters K_W,M_W, color indicates $I_1(X_1,A_1)$. The input distributions are uniformly distributed with $\mu_2^j = 50j/N_\mu$,$A_1^i = (i-1)/(N_A-1)$. Parameters as in Fig. 8.10b and T=1000 s. Panel **a**, **c)** N_A=4,N_μ=2 **b**, **d)** N_A=4,N_μ=8. Panels **a**, **b**: X_1 as given by Eq. A8.17 $k_{X_1} = 5m_{X_1} = 1/10\,\mathrm{T\,s}^{-1}$, panels **c**, **d**: X_1 as given by Eq. A8.18, K_{X_1}=5000,n_{X_1}=5, k_{X_1}=3500,$m_{X_1} = 1/10\,\mathrm{T\,s}^{-1}$.

8.A.5 Numerical optimization

The numerical optimization used to produce Fig. 8.5 is based upon the Wright-Fisher model for population evolution. In the simulations N systems are present, each consisting of a

single multiplexing network. In the optimization procedure, we optimize over the kinetic parameters of the network, for given uniform input distributions for the two signals.

$$p\left(A_1 = a\right) = \frac{1}{N_A}, \text{ with amplitude values } a_i = \frac{i}{N_A}, i \in [1 : N_A], \tag{A8.19}$$

$$p\left(\mu_2 = \mu\right) = \frac{1}{N_\mu} \text{ with concentration values } \mu_j = \frac{j}{N_\mu}\mu_{\max}, j \in \left[1 : N_\mu\right]. \tag{A8.20}$$

In the initialization step, each network is assigned random parameters where each parameter is within the ranges specified in Eq. A8.22 and Eq. A8.23. At every step for all input states $(A_{1,i}, \mu_{2,j})$ the output concentrations $\langle X_1\rangle, \langle X_2\rangle$ and corresponding variances $\sigma_{X_1}, \sigma_{X_2}$ are determined. With these the mutual information $I_1\left(A_1, X_1\right)$ and $I_2\left(\mu_2, X_2\right)$ is obtained. Using the total mutual information $I_T = I_1 + I_2$ for each network, in the selection step N new systems are selected, where the likelihood of reproduction of each system is proportional to its fitness I_T. A larger mutual information increases the likelihood of being used as a template for any of the new systems. Each new network is then "mutated" by multiplying a randomly selected parameter by the factor (1 + δ), where δ is drawn uniform randomly from the range [–Δ: Δ]; we take $\Delta = 0.3$. Then the cycle repeats.

We take the kinetic parameters of the encoding module to be fixed, since the transmission of oscillations is optimized within the linear regime (see Eq. 8.3). We further take the mean of the oscillatory signal to be constant and if necessary $X_{1,T}$

$$\begin{aligned} &k_V = 1\,\mathrm{s}^{-1},\ m_V = 5625\,\mathrm{s}^{-1},\ M_V = 0.01V_T,\ K_V = 30V_T,\ V_T = 1000,\\ &\mu_1 = 25,\ X_{1,T} = 3000. \end{aligned} \tag{A8.21}$$

Some parameters are constrained based upon values of other parameters

$$\begin{aligned} &k_W \text{ to set } \left\langle W^P \right\rangle = W_T/2 \text{ for a constant signal,}\\ &m_{X_2} = m_{X_1} = \left(10T^{-1}\right),\\ &k_{X_2} = 5m_{X_2},\\ &k_{X_1} = 3500m_{X_1} f\left(X_{1,T}, 2, K_{X_1}, n_{X_1}\right). \end{aligned} \tag{A8.22}$$

This leads with the following set of parameters to optimize over for given N_A, N_μ, where we assume for the response X_1 a functional description following Eq. A8.18. In square brackets we give the minimum and maximum range of each parameter.

$$\begin{aligned} &K_W\ \left[1\,\mathrm{au} : 75000\,\mathrm{au}\right], M_W\ \left[1\,\mathrm{au} : 75000\,\mathrm{au}\right], m_W\ \left[10^{-3}\,\mathrm{s}^{-1} : 10^{3}\,\mathrm{s}^{-1}\right],\\ &T\ \left[10^{1}\,\mathrm{s} : 10^{4}\,\mathrm{s}\right], \mu_{2,\max}\ \left[1\,\mathrm{au} : 100\,\mathrm{au}\right], n_{X_1}\ \left[1 : 5\right], K_{X_1}\ \left[1\,\mathrm{au} : 75000\,\mathrm{au}\right]. \end{aligned} \tag{A8.23}$$

References

[1] Schwann T (**1839**) Microscopic investigations on the accordance in the structure and growth of plants and animals. Berlin: Sydenham Society (English Translation, 1847).

[2] Adler J (**1975**) Chemotaxis in bacteria. *Annual Review of Biochemistry* **44**:341–56. doi:10.1146/annurev.bi.44.070175.002013

[3] Monod J (**1966**) From Enzymatic Adaptation to Allosteric Transitions. *Science* **154**:475–483. doi:10.1126/science.154.3748.475

[4] Levi JD (**1956**) Mating Reaction in Yeast. *Nature* **177**:753–754. doi:10.1038/177753b0

[5] Cover TM, Thomas JA (**1991**) Elements of Information Theory. Wiley Series in Telecommunications. New York, USA: John Wiley & Sons, Inc., second edition, 776 pp. doi:10.1002/0471200611.

[6] Rieke F, Baylor D (**1998**) Single-photon detection by rod cells of the retina. *Reviews of Modern Physics* **70**:1027–1036. doi:10.1103/RevModPhys.70.1027

[7] Ueda M, Shibata T (**2007**) Stochastic signal processing and transduction in chemotactic response of eukaryotic cells. *Biophysical journal* **93**:11–20. doi:10.1529/biophysj.106.100263

[8] Gregor T, Tank DW, Wieschaus EF, Bialek W (**2007**) Probing the limits to positional information. *Cell* **130**:153–64. doi:10.1016/j.cell.2007.05.025

[9] Guptasarma P (**1995**) Does replication-induced transcription regulate synthesis of the myriad low copy number proteins of Escherichia coli? *BioEssays : news and reviews in molecular, cellular and developmental biology* **17**:987–97. doi:10.1002/bies.950171112

[10] Milo R, Shen-Orr SS, Itzkovitz S, Kashtan N, Chklovskii D, et al. (**2002**) Network motifs: simple building blocks of complex networks. *Science* **298**:824–7. doi:10.1126/science.298.5594.824

[11] Shen-Orr SS, Milo R, Mangan S, Alon U (**2002**) Network motifs in the transcriptional regulation network of Escherichia coli. *Nature genetics* **31**:64–8. doi:10.1038/ng881

[12] Yeger-Lotem E, Sattath S, Kashtan N, Itzkovitz S, Milo R, et al. (**2004**) Network motifs in integrated cellular networks of transcription-regulation and protein-protein interaction. *Proceedings of the National Academy of Sciences of the United States of America* **101**:5934–9. doi:10.1073/pnas.0306752101

[13] Kashtan N, Alon U (**2005**) Spontaneous evolution of modularity and network motifs. *Proceedings of the National Academy of Sciences of the United States of America* **102**:13773–8.
doi:10.1073/pnas.0503610102

[14] Alon U (**2007**) Network motifs: theory and experimental approaches. *Nature reviews Genetics* **8**:450–61.
doi:10.1038/nrg2102

[15] Tyson JJ, Novák B (**2010**) Functional motifs in biochemical reaction networks. *Annual Review of Physical Chemistry* **61**:219–40.
doi:10.1146/annurev.physchem.012809.103457

[16] Burda Z, Krzywicki A, Martin OC, Zagorski M (**2011**) Motifs emerge from function in model gene regulatory networks. *Proceedings of the National Academy of Sciences of the United States of America* :10.
doi:10.1073/pnas.1109435108

[17] Mangan S, Alon U (**2003**) Structure and function of the feed-forward loop network motif. *Proceedings of the National Academy of Sciences of the United States of America* **100**:11980–5.
doi:10.1073/pnas.2133841100

[18] Barkai N, Leibler S (**1997**) Robustness in simple biochemical networks. *Nature* **387**:913–917.

[19] Ferrell Jr JE (**1998**) The Biochemical Basis of an All-or-None Cell Fate Switch in Xenopus Oocytes. *Science* **280**:895–898.
doi:10.1126/science.280.5365.895

[20] Prill RJ, Iglesias PA, Levchenko A (**2005**) Dynamic properties of network motifs contribute to biological network organization. *PLoS biology* **3**:e343.
doi:10.1371/journal.pbio.0030343

[21] Brandman O, Meyer T (**2008**) Feedback loops shape cellular signals in space and time. *Science* **322**:390–5.
doi:10.1126/science.1160617

[22] Tyson JJ, Chen KC, Novák B (**2003**) Sniffers, buzzers, toggles and blinkers: dynamics of regulatory and signaling pathways in the cell. *Current Opinion in Cell Biology* **15**:221–231.
doi:10.1016/S0955-0674(03)00017-6

[23] Nakajima M, Imai K, Ito H, Nishiwaki T, Murayama Y, et al. (**2005**) Reconstitution of circadian oscillation of cyanobacterial KaiC phosphorylation in vitro. *Science* **308**:414–5.
doi:10.1126/science.1108451

[24] Zwicker D, Lubensky DK, ten Wolde PR (**2010**) Robust circadian clocks from coupled protein-modification and transcription-translation cycles. *Proceedings of the National Academy of Sciences of the United States of America* **107**:22540–5.
doi:10.1073/pnas.1007613107

[25] Elowitȝ MB, Levine AJ, Siggia ED, Swain PS (**2002**) Stochastic gene expression in a single cell. *Science* **297**:1183–6.
doi:10.1126/science.1070919

[26] Ozbudak EM, Thattai M, Kurtser I, Grossman AD, van Oudenaarden A (**2002**) Regulation of noise in the expression of a single gene. *Nature genetics* **31**:69–73. doi:10.1038/ng869

[27] Newman JRS, Ghaemmaghami S, Ihmels J, Breslow DK, Noble M, et al. (**2006**) Single-cell proteomic analysis of S. cerevisiae reveals the architecture of biological noise. *Nature* **441**:840–6. doi:10.1038/nature04785

[28] Raj A, van Oudenaarden A (**2009**) Single-molecule approaches to stochastic gene expression. *Annual Review of Biophysics* **38**:255–70. doi:10.1146/annurev.biophys.37.032807.125928

[29] Taniguchi Y, Choi PJ, Li GW, Chen H, Babu MM, et al. (**2010**) Quantifying E. coli proteome and transcriptome with single-molecule sensitivity in single cells. *Science* **329**:533–8. doi:10.1126/science.1188308

[30] Locasale JW (**2008**) Signal duration and the time scale dependence of signal integration in biochemical pathways. *BMC systems biology* **2**:108. doi:10.1186/1752-0509-2-108

[31] Scott M, Hwa T, Ingalls BP (**2007**) Deterministic characterization of stochastic genetic circuits. *Proceedings of the National Academy of Sciences of the United States of America* **104**:7402–7. doi:10.1073/pnas.0610468104

[32] Bruggeman FJ, Blüthgen N, Westerhoff HV (**2009**) Noise management by molecular networks. *PLoS computational biology* **5**:e1000506. doi:10.1371/journal.pcbi.1000506

[33] Heinrich R, Neel BG, Rapoport TA (**2002**) Mathematical models of protein kinase signal transduction. *Molecular cell* **9**:957–70.

[34] Thattai M, van Oudenaarden A (**2001**) Intrinsic noise in gene regulatory networks. *Proceedings of the National Academy of Sciences of the United States of America* **98**:8614–9. doi:10.1073/pnas.151588598

[35] Tan C, Reza F, You L (**2007**) Noise-limited frequency signal transmission in gene circuits. *Biophysical journal* **93**:3753–61. doi:10.1529/biophysj.107.110403

[36] Chaves M, Sontag ED, Dinerstein RJ (**2004**) Optimal Length and Signal Amplification in Weakly Activated Signal Transduction Cascades. *The Journal of Physical Chemistry B* **108**:15311–15320. doi:10.1021/jp048935f

[37] Ghosh B, Karmakar R, Bose I (**2005**) Noise characteristics of feed forward loops. *Physical biology* **2**:36–45. doi:10.1088/1478-3967/2/1/005

[38] Cağatay T, Turcotte M, Elowitȝ MB, García-Ojalvo J, Süel GM (**2009**) Architecture-dependent noise discriminates functionally analogous differentiation circuits. *Cell* **139**:512–22. doi:10.1016/j.cell.2009.07.046

[39] Hu B, Levine H, Rappel WJ (**2011**) Design principles and specificity in biological networks with cross activation. *Physical biology* **8**:026001.
doi:10.1088/1478-3975/8/2/026001

[40] Lipshtat A, Purushothaman SP, Iyengar R, Ma'ayan A (**2008**) Functions of bifans in context of multiple regulatory motifs in signaling networks. *Biophysical journal* **94**:2566–79.
doi:10.1529/biophysj.107.116673

[41] Shahrezaei V, Ollivier JF, Swain PS (**2008**) Colored extrinsic fluctuations and stochastic gene expression. *Molecular systems biology* **4**:196.
doi:10.1038/msb.2008.31

[42] Cohen-Saidon C, Cohen AA, Sigal A, Liron Y, Alon U (**2009**) Dynamics and variability of ERK2 response to EGF in individual living cells. *Molecular cell* **36**:885–93.
doi:10.1016/j.molcel.2009.11.025

[43] Pedraza JM, van Oudenaarden A (**2005**) Noise propagation in gene networks. *Science* **307**:1965–9.
doi:10.1126/science.1109090

[44] Shibata T, Fujimoto K (**2005**) Noisy signal amplification in ultrasensitive signal transduction. *Proceedings of the National Academy of Sciences of the United States of America* **102**:331–6.
doi:10.1073/pnas.0403350102

[45] Shannon CE (**1948**) The mathematical theory of communication. *The Bell System Technical Journal* **27**:379–423,623–656.

[46] Tkačik G (**2010**) From statistical mechanics to information theory: understanding biophysical information-processing systems. *Arxiv preprint* .

[47] MacKay DJC (**2003**) Information theory, inference and learning algorithms, volume 22. Cambridge: Cambridge University Press, 640 pp.

[48] Borst A, Theunissen FE (**1999**) Information theory and neural coding. *Nature neuroscience* **2**:947–57.
doi:10.1038/14731

[49] Ziv E, Nemenman I, Wiggins CH (**2007**) Optimal signal processing in small stochastic biochemical networks. *PloS one* **2**:e1077.
doi:10.1371/journal.pone.0001077

[50] Walczak AM, Mugler A, Wiggins CH (**2009**) A stochastic spectral analysis of transcriptional regulatory cascades. *Proceedings of the National Academy of Sciences of the United States of America* **106**:6529–34.
doi:10.1073/pnas.0811999106

[51] Mehta P, Goyal S, Long T, Bassler BL, Wingreen NS (**2009**) Information processing and signal integration in bacterial quorum sensing. *Molecular systems biology* **5**:325.
doi:10.1038/msb.2009.79

[52] Tkačik G, Callan C, Bialek W (**2008**) Information capacity of genetic regulatory elements. *Physical Review E* **78**:1–17.
doi:10.1103/PhysRevE.78.011910

[53] Mugler A, Walczak A, Wiggins C (**2010**) Information-Optimal Transcriptional Response to Oscillatory Driving. *Physical Review Letters* **105**:1–4.
doi:10.1103/PhysRevLett.105.058101

[54] Walczak AM, Tkačik G, Bialek W (**2010**) Optimizing information flow in small genetic networks. II. Feed-forward interactions. *Physical Review E* **81**:1–16.
doi:10.1103/PhysRevE.81.041905

[55] Cheong R, Rhee A, Wang CJ, Nemenman I, Levchenko A (**2011**) Information Transduction Capacity of Noisy Biochemical Signaling Networks. *Science* .
doi:10.1126/science.1204553

[56] Tkačik G, Walczak AM, Bialek W (**2009**) Optimizing information flow in small genetic networks. *Physical Review E* **80**:1–18.
doi:10.1103/PhysRevE.80.031920

[57] Tkačik G, Walczak AM (**2011**) Information transmission in genetic regulatory networks: a review. *Journal of Physics: Condensed matter* **23**:153102.
doi:10.1088/0953-8984/23/15/153102

[58] Lestas I, Vinnicombe G, Paulsson J (**2010**) Fundamental limits on the suppression of molecular fluctuations. *Nature* **467**:174–178.
doi:10.1038/nature09333

[59] Raser JM, O'Shea EK (**2005**) Noise in gene expression: origins, consequences, and control. *Science* **309**:2010–3.
doi:10.1126/science.1105891

[60] Samoilov M, Plyasunov S, Arkin AP (**2005**) Stochastic amplification and signaling in enzymatic futile cycles through noise-induced bistability with oscillations. *Proceedings of the National Academy of Sciences of the United States of America* **102**:2310–5.
doi:10.1073/pnas.0406841102

[61] Munakata T, Kamiyabu M (**2006**) Stochastic resonance in the FitzHugh-Nagumo model from a dynamic mutual information point of view. *The European Physical Journal B* **53**:239–243.
doi:10.1140/epjb/e2006-00363-x

[62] Tănase-Nicola S, ten Wolde PR (**2008**) Regulatory Control and the Costs and Benefits of Biochemical Noise. *PLoS computational biology* **4**.
doi:10.1371/journal.pcbi.1000125

[63] Locke JCW, Young JW, Fontes ME, Jiménez MJH, Elowitz MB (**2011**) Stochastic Pulse Regulation in Bacterial Stress Response. *Science* **366**.
doi:10.1126/science.1208144

[64] Mitra PP, Stark JB (**2001**) Nonlinear limits to the information capacity of optical fibre communications. *Nature* **411**:1027–30.
doi:10.1038/35082518

[65] Tostevin F, ten Wolde PR (**2009**) Mutual Information between Input and Output Trajectories of Biochemical Networks. *Physical Review Letters* **102**:1–4. doi:10.1103/PhysRevLett.102.218101

[66] Fano R (**1961**) Transmission of Information: A Statistical Theory of Communications. Cambridge, MA: MIT Press.

[67] Tostevin F, ten Wolde P (**2010**) Mutual information in time-varying biochemical systems. *Physical Review E* **81**:1–15. doi:10.1103/PhysRevE.81.061917

[68] van Kampen NG (**2007**) Stochastic Processes in Physics and Chemistry. North-Holland: Elsevier, third edition.

[69] Gillespie DT (**2000**) The chemical Langevin equation. *The Journal of Chemical Physics* **113**:297. doi:10.1063/1.481811

[70] Paulsson J (**2005**) Models of stochastic gene expression. *Physics of Life Reviews* **2**:157–175. doi:10.1016/j.plrev.2005.03.003

[71] Berg HC (**2003**) The rotary motor of bacterial flagella. *Annual review of biochemistry* **72**:19–54. doi:10.1146/annurev.biochem.72.121801.161737

[72] Nevozhay D, Adams RM, Murphy KF, Josic K, Balázsi G (**2009**) Negative autoregulation linearizes the dose-response and suppresses the heterogeneity of gene expression. *Proceedings of the National Academy of Sciences of the United States of America* **106**:5123–8. doi:10.1073/pnas.0809901106

[73] Lorenz E (**1963**) Deterministic Nonperiodic Flow. *Journal of the Atmospheric Sciences* **20**:130–141.

[74] Gillespie DT (**1976**) A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *Journal of Computational Physics* **22**:403–434. doi:10.1016/0021-9991(76)90041-3

[75] Gillespie DT (**1977**) Exact stochastic simulation of coupled chemical reactions. *The journal of physical chemistry* **93555**:2340–2361.

[76] Wallace EWJ (**2010**) A simplified derivation of Van Kampen's system size expansion. *Arxiv preprint* :5.

[77] Paulsson J (**2004**) Summing up the noise in gene networks. *Nature* **427**:415–8. doi:10.1038/nature02257

[78] Risken H (**1984**) The Fokker-Planck Equation. Methods of Solution and Applications. Berlin — Heidelberg — New York — Tokyo: Springer-Verlag, first edition, 454 pp.

[79] Elf J, Ehrenberg Mn (**2003**) Fast evaluation of fluctuations in biochemical networks with the linear noise approximation. *Genome research* **13**:2475–84. doi:10.1101/gr.1196503

[80] Wright SC (**1931**) Evolution in Mendelian Populations. *Genetics* **16**:97–159.

[81] Fisher R (**1930**) The genetical theory of natural selection. Oxford: Clarendon Press.

[82] Hermsen R (**2008**) Transcription Regulation and Genome Organization. Ph.D. thesis, Vrije Universiteit Amsterdam.

[83] Berg HC, Purcell EM (**1977**) Physics of chemoreception. *Biophysical journal* **20**:193–219. doi:10.1016/S0006-3495(77)85544-6

[84] Lim WA (**2010**) Designing customized cell signalling circuits. *Nature reviews Molecular cell biology* **11**:393–403. doi:10.1038/nrm2904

[85] Bennett MR, Pang WL, Ostroff NA, Baumgartner BL, Nayak S, et al. (**2008**) Metabolic gene regulation in a dynamically changing environment. *Nature* **454**:1119–22. doi:10.1038/nature07211

[86] Mettetal JT, Muzzey D, Gómez-Uribe CA, van Oudenaarden A (**2008**) The frequency dependence of osmo-adaptation in Saccharomyces cerevisiae. *Science* **319**:482–4. doi:10.1126/science.1151582

[87] Tu Y, Shimizu TS, Berg HC (**2008**) Modeling the chemotactic response of Escherichia coli to time-varying stimuli. *Proceedings of the National Academy of Sciences of the United States of America* **105**:14855–60. doi:10.1073/pnas.0807569105

[88] Swain PS, Elowitz MB, Siggia ED (**2002**) Intrinsic and extrinsic contributions to stochasticity in gene expression. *Proceedings of the National Academy of Sciences of the United States of America* **99**:12795–800. doi:10.1073/pnas.162041399

[89] Ma'ayan A, Jenkins SL, Neves SR, Hasseldine A, Grace E, et al. (**2005**) SI: Formation of regulatory patterns during signal propagation in a Mammalian cellular network. *Science* **309**:1078–83. doi:10.1126/science.1108876

[90] Segall JE, Block SM, Berg HC (**1986**) Temporal comparisons in bacterial chemotaxis. *Proceedings of the National Academy of Sciences of the United States of America* **83**:8987–91.

[91] Marshall CJ (**1995**) Specificity of receptor tyrosine kinase signaling: transient versus sustained extracellular signal-regulated kinase activation. *Cell* **80**:179–85.

[92] Yi TM, Huang Y, Simon MI, Doyle JC (**2000**) Robust perfect adaptation in bacterial chemotaxis through integral feedback control. *Proceedings of the National Academy of Sciences of the United States of America* **97**:4649–53.

[93] Samoilov M, Arkin A, Ross J (**2002**) Signal Processing by Simple Chemical Systems. *The Journal of Physical Chemistry A* **106**:10205–10221. doi:10.1021/jp025846z

[94] Ingalls BP (**2004**) A Frequency Domain Approach to Sensitivity Analysis of Biochemical Networks. *The Journal of Physical Chemistry B* **108**:1143–1152. doi:10.1021/jp036567u

[95] Simpson ML, Cox CD, Sayler GS (**2003**) Frequency domain analysis of noise in autoregulated gene circuits. *Proceedings of the National Academy of Sciences of the United States of America* **100**:4551–6.
doi:10.1073/pnas.0736140100

[96] Tănase-Nicola S, Warren PB, ten Wolde PR (**2006**) Signal Detection, Modularity, and the Correlation between Extrinsic and Intrinsic Noise in Biochemical Networks. *Physical Review Letters* **97**:21.
doi:10.1103/PhysRevLett.97.068102

[97] Warren PB, Tănase-Nicola S, ten Wolde PR (**2006**) Exact results for noise power spectra in linear biochemical reaction networks. *The Journal of chemical physics* **125**:144904.
doi:10.1063/1.2356472

[98] Alon U (**2007**) An Introduction to Systems Biology. Florida: Chapman & Hall/CRC, 301 pp.

[99] Hornung G, Barkai N (**2008**) Noise propagation and signaling sensitivity in biological networks: a role for positive feedback. *PLoS computational biology* **4**:e8.
doi:10.1371/journal.pcbi.0040008

[100] Macnab RM, Koshland DE (**1972**) The gradient-sensing mechanism in bacterial chemotaxis. *Proceedings of the National Academy of Sciences of the United States of America* **69**:2509–12.

[101] Ma W, Trusina A, El-Samad H, Lim WA, Tang C (**2009**) Defining network topologies that can achieve biochemical adaptation. *Cell* **138**:760–73.
doi:10.1016/j.cell.2009.06.013

[102] Levine E, Hwa T (**2007**) Stochastic fluctuations in metabolic pathways. *Proceedings of the National Academy of Sciences of the United States of America* **104**:9224–9.
doi:10.1073/pnas.0610987104

[103] Lipan O, Wong WH (**2005**) The use of oscillatory signals in the study of genetic networks. *Proceedings of the National Academy of Sciences of the United States of America* **102**:7063–8.
doi:10.1073/pnas.0403790102

[104] Austin DW, Allen MS, McCollum JM, Dar RD, Wilgus JR, et al. (**2006**) Gene network shaping of inherent noise spectra. *Nature* **439**:608–11.
doi:10.1038/nature04194

[105] Cournac A, Sepulchre Ja (**2009**) Simple molecular networks that respond optimally to time-periodic stimulation. *BMC systems biology* **3**:29.
doi:10.1186/1752-0509-3-29

[106] Murphy LO, MacKeigan JP, Blenis J (**2004**) A Network of Immediate Early Gene Products Propagates Subtle Differences in Mitogen-Activated Protein Kinase Signal Amplitude and Duration. *Molecular and Cellular Biology* **24**:144–153.
doi:10.1128/MCB.24.1.144-153.2004

[107] de Ronde W, Tostevin F, ten Wolde PR (**2010**) Effect of feedback on the fidelity of information transmission of time-varying signals. *Physical Review E* **82**.
doi:10.1103/PhysRevE.82.031914

[108] Goentoro L, Shoval O, Kirschner MW, Alon U (**2009**) The incoherent feedforward loop can provide fold-change detection in gene regulation. *Molecular cell* **36**:894–9. doi:10.1016/j.molcel.2009.11.018

[109] Dunlop MJ, Cox RS, Levine JH, Murray RM, Elowitz MB (**2008**) Regulatory activity revealed by dynamic correlations in gene expression noise. *Nature genetics* **40**:1493–8. doi:10.1038/ng.281

[110] Kalir S, Mangan S, Alon U (**2005**) A coherent feed-forward loop with a SUM input function prolongs flagella expression in Escherichia coli. *Molecular systems biology* **1**:2005.0006. doi:10.1038/msb4100010

[111] Kittisopikul M, Süel GM (**2010**) Biological role of noise encoded in a genetic network motif. *Proceedings of the National Academy of Sciences of the United States of America* **107**:13300–5. doi:10.1073/pnas.1003975107

[112] Hatakeyama M, Kimura S, Naka T, Kawasaki T, Yumoto N, et al. (**2003**) A computational model on the modulation of mitogen-activated protein kinase (MAPK) and Akt pathways in heregulin-induced ErbB signalling. *The Biochemical journal* **373**:451–63. doi:10.1042/BJ20021824

[113] Itzkovitz S, Levitt R, Kashtan N, Milo R, Itzkovitz M, et al. (**2005**) Coarse-graining and self-dissimilarity of complex networks. *Physical Review E* **71**:1–10. doi:10.1103/PhysRevE.71.016127

[114] Boeckh J, Kaissling KE, Schneider D (**1965**) Insect Olfactory Receptors. *Cold Spring Harbor Symp Quant Biol* **30**:1263–1280.

[115] Sourjik V, Berg HC (**2002**) Receptor sensitivity in bacterial chemotaxis. *Proceedings of the National Academy of Sciences of the United States of America* **99**:123–7. doi:10.1073/pnas.011589998

[116] Erdmann T, Howard M, ten Wolde PR (**2009**) Role of Spatial Averaging in the Precision of Gene Expression Patterns. *Physical Review Letters* **103**:2–5. doi:10.1103/PhysRevLett.103.258101

[117] Dubuis JO, Tkačik G, Wieschaus EF, Gregor T, Bialek W (**2011**) Positional information, in bits. *Arxiv preprint* :1–10.

[118] Bialek W, Setayeshgar S (**2005**) Physical limits to biochemical signaling. *Proceedings of the National Academy of Sciences of the United States of America* **102**:10040–5. doi:10.1073/pnas.0504321102

[119] Hu B, Kessler D, Rappel WJ, Levine H (**2011**) Effects of Input Noise on a Simple Biochemical Switch. *Physical Review Letters* **107**:1–5. doi:10.1103/PhysRevLett.107.148101

[120] Agmon N, Szabo A (**1990**) Theory of reversible diffusion-influenced reactions. *The Journal of Chemical Physics* **92**:5270.

[121] van Zon JS, ten Wolde PR (**2005**) Simulating Biochemical Networks at the Particle Level and in Time and Space: Green's Function Reaction Dynamics. *Physical Review Letters* **94**:1–4. doi:10.1103/PhysRevLett.94.128103

[122] van Zon JS, ten Wolde PR (**2005**) Green's-function reaction dynamics: a particle-based approach for simulating biochemical networks in time and space. *The Journal of chemical physics* **123**:234910. doi:10.1063/1.2137716

[123] Takahashi K, Tănase-Nicola S, ten Wolde PR (**2010**) Spatio-temporal correlations can drastically change the response of a MAPK pathway. *Proceedings of the National Academy of Sciences of the United States of America* **107**:2473–8. doi:10.1073/pnas.0906885107

[124] Rice SA (**1985**) Diffusion-limited reactions. Amsterdam: Elsevier.

[125] Popov AV, Agmon N (**2001**) Three-dimensional simulations of reversible bimolecular reactions: The simple target problem. *The Journal of Chemical Physics* **115**:8921. doi:10.1063/1.1412609

[126] Gopich IV, Szabo A (**2002**) Kinetics of reversible diffusion influenced reactions: The self-consistent relaxation time approximation. *The Journal of Chemical Physics* **117**:507. doi:10.1063/1.1482701

[127] Frenkel D, Smit B (**2001**) Understanding Molecular Simulation, Second Edition: From Algorithms to Applications (Computational Science). Academic Press, second edition.

[128] Chandler D (**1987**) Introduction to Modern Statistical Mechanics. Oxford University Press, 247 pp.

[129] Kaplan S, Bren A, Zaslaver A, Dekel E, Alon U (**2008**) Diverse two-dimensional input functions control bacterial sugar genes. *Molecular cell* **29**:786–92. doi:10.1016/j.molcel.2008.01.021

[130] Prehoda K, Lim WA (**2002**) How signaling proteins integrate multiple inputs: a comparison of N-WASP and Cdk2. *Current Opinion in Cell Biology* **14**:149–154. doi:10.1016/S0955-0674(02)00307-1

[131] Bray D (**1995**) Protein molecules as computational elements in living cells. *Nature* :307–312.

[132] Tamsir A, Tabor JJ, Voigt CA (**2011**) Robust multicellular computing using genetically encoded NOR gates and chemical 'wires'. *Nature* **469**:212–5. doi:10.1038/nature09565

[133] Regot S, Macia J, Conde N, Furukawa K, Kjellén J, et al. (**2011**) Distributed biological computation with multicellular engineered networks. *Nature* **469**:207–11. doi:10.1038/nature09679

[134] Dueber JE, Yeh BJ, Chak K, Lim WA (**2003**) Reprogramming control of an allosteric signaling switch through modular recombination. *Science* **301**:1904–8. doi:10.1126/science.1085945

[135] de Silva AP, McClenaghan ND (**2004**) Molecular-scale logic gates. *Chemistry (Weinheim an der Bergstrasse, Germany)* **10**:574–86.
doi:10.1002/chem.200305054

[136] Deonarine A, Clark S, Konermann L (**2003**) Implementation of a multifunctional logic gate based on folding/unfolding transitions of a protein. *Future Generation Computer Systems* **19**:87–97.
doi:10.1016/S0167-739X(02)00110-3

[137] Arkin AP, Ross J (**1994**) Computational functions in biochemical reaction networks. *Biophysical Journal* **67**:560–578.
doi:10.1016/S0006-3495(94)80516-8

[138] Hermsen R, Tans SJ, ten Wolde PR (**2006**) Transcriptional regulation by competing transcription factor modules. *PLoS computational biology* **2**:e164.
doi:10.1371/journal.pcbi.0020164

[139] Buchler NE, Gerland U, Hwa T (**2003**) On schemes of combinatorial transcription logic. *Proceedings of the National Academy of Sciences of the United States of America* **100**:5136–41.
doi:10.1073/pnas.0930314100

[140] Bintu L, Buchler NE, Garcia HG, Gerland U, Hwa T, et al. (**2005**) Transcriptional regulation by the numbers: models. *Current Opinion in genetics & development* **15**:116–24.
doi:10.1016/j.gde.2005.02.007

[141] Citri A, Yarden Y (**2006**) EGF-ERBB signalling: towards the systems level. *Nature reviews Molecular cell biology* 7:505–16.
doi:10.1038/nrm1962

[142] Barton GM, Medzhitov R (**2002**) Toll-like receptors and their ligands. *Current Topics in Microbiology and Immunology* **270**:81–92.

[143] Moghal N, Sternberg PW (**1999**) Multiple positive and negative regulators of signaling by the EGF-receptor. *Current Opinion in Cell Biology* **11**:190–6.

[144] Bray D (**1998**) Signaling complexes: biophysical constraints on intracellular communication. *Annual Review of Biophysics and Biomolecular Structure* **27**:59–75.
doi:10.1146/annurev.biophys.27.1.59

[145] Landau M, Ben-Tal N (**2008**) Dynamic equilibrium between multiple active and inactive conformations explains regulation and oncogenic mutations in ErbB receptors. *Biochimica et biophysica acta* **1785**:12–31.
doi:10.1016/j.bbcan.2007.08.001

[146] Minneman KP (**2007**) Heterodimerization and surface localization of G protein coupled receptors. *Biochemical pharmacology* **73**:1043–50.
doi:10.1016/j.bcp.2006.09.001

[147] Lauffenburger DA, Linderman JJ (**1996**) Receptors: Models for Binding, Trafficking, and Signaling. Oxford: Oxford University Press, second edition.

[148] Phillips R, Kondev J, Theriot J (**2008**) Physical Biology of the Cell. New York: Garland Science, first edition.

[149] Koshland DE, Némethy G, Filmer D (**1966**) Comparison of experimental binding data and theoretical models in proteins containing subunits. *Biochemistry* **5**:365–85.

[150] Monod J, Wyman J, Changeux JP (**1965**) On the nature of allosteric transitions: A plausible model. *Journal of Molecular Biology* **12**:88–118.
doi:10.1016/S0022-2836(65)80285-6

[151] Bray D, Duke T (**2004**) Conformational spread: the propagation of allosteric states in large multiprotein complexes. *Annual Review of Biophysics and Biomolecular Structure* **33**:53–73.
doi:10.1146/annurev.biophys.33.110502.132703

[152] Hansen CH, Sourjik V, Wingreen NS (**2010**) A dynamic-signaling-team model for chemotaxis receptors in Escherichia coli. *Proceedings of the National Academy of Sciences of the United States of America* .
doi:10.1073/pnas.1005017107

[153] Behar M, Hoffmann A (**2010**) Understanding the temporal codes of intra-cellular signals. *Current Opinion in genetics & development* **20**:684–93.
doi:10.1016/j.gde.2010.09.007

[154] Paszek P, Jackson DA, White MRH (**2010**) Oscillatory control of signalling molecules. *Current Opinion in genetics & development* **20**:670–676.
doi:10.1016/j.gde.2010.08.004

[155] Berridge MJ, Galione A (**1988**) Cytosolic calcium oscillators. *The FASEB journal* **2**:3074–3082.

[156] Shankaran H, Ippolito DL, Chrisler WB, Resat H, Bollinger N, et al. (**2009**) Rapid and sustained nuclear-cytoplasmic ERK oscillations induced by epidermal growth factor. *Molecular systems biology* **5**:332.
doi:10.1038/msb.2009.90

[157] Nelson DE, See V, Nelson G, White MRH, Se V, et al. (**2004**) Oscillation in transcription factor dynamics: a new way to control gene expression. *Biochemical Society Transactions* **32**:1090–1092.

[158] Jacquet M, Renault G, Lallet S, de Mey J, Goldbeter A (**2003**) Oscillatory nucleocytoplasmic shuttling of the general stress response transcriptional activators Msn2 and Msn4 in Saccharomyces cerevisiae. *The Journal of cell biology* **161**:497–505.
doi:10.1083/jcb.200303030

[159] Cai L, Dalal CK, Elowitz MB (**2008**) Frequency-modulated nuclear localization bursts coordinate gene regulation. *Nature* **455**:485–90.
doi:10.1038/nature07292

[160] Lahav G, Rosenfeld N, Sigal A, Geva-Zatorsky N, Levine AJ, et al. (**2004**) Dynamics of the p53-Mdm2 feedback loop in individual cells. *Nature genetics* **36**:147–50.
doi:10.1038/ng1293

[161] Trump BF, Berezesky IK (**1995**) Calcium-mediated cell injury and cell death. *The FASEB journal* **9**:219–228.

[162] Gall D, Baus E, Dupont G (**2000**) Activation of the liver glycogen phosphorylase by Ca(2+)oscillations: a theoretical study. *Journal of theoretical biology* **207**:445–54. doi:10.1006/jtbi.2000.2139

[163] Rapp PE, Mees AI, Sparrow CT (**1981**) Frequency encoded biochemical regulation is more accurate than amplitude dependent control. *Journal of theoretical biology* **90**:531–44.

[164] Suter DM, Molina N, Gatfield D, Schneider K, Schibler U, et al. (**2011**) Mammalian genes are transcribed with widely different bursting kinetics. *Science* **332**:472–4. doi:10.1126/science.1198817

[165] Chubb JR, Trcek T, Shenoy SM, Singer RH (**2006**) Transcriptional pulsing of a developmental gene. *Current biology* **16**:1018–25. doi:10.1016/j.cub.2006.03.092

[166] Schwanhäusser B, Busse D, Li N, Dittmar G, Schuchhardt J, et al. (**2011**) Global quantification of mammalian gene expression control. *Nature* **473**:337–42. doi:10.1038/nature10098

[167] Geva-Zatorsky N, Rosenfeld N, Itzkovitz S, Milo R, Sigal A, et al. (**2006**) Oscillations and variability in the p53 system. *Molecular systems biology* **2**:2006.0033. doi:10.1038/msb4100068

[168] Schwartz MA, Madhani HD (**2004**) Principles of MAP kinase signaling specificity in Saccharomyces cerevisiae. *Annual Review of Genetics* **38**:725–48. doi:10.1146/annurev.genet.39.073003.112634

[169] Patterson JC, Klimenko ES, Thorner J (**2010**) Single-cell analysis reveals that insulation maintains signaling specificity between two yeast MAPK pathways with common components. *Science signaling* **3**:1–11. doi:10.1126/scisignal.2001275

[170] Bardwell L, Zou X, Nie Q, Komarova NL (**2007**) Mathematical models of specificity in cell signaling. *Biophysical journal* **92**:3425–41. doi:10.1529/biophysj.106.090084

[171] Behar M, Dohlman HG, Elston TC (**2007**) Kinetic insulation as an effective mechanism for achieving pathway specificity in intracellular signaling networks. *Proceedings of the National Academy of Sciences of the United States of America* **104**:16146–51. doi:10.1073/pnas.0703894104

[172] Berridge MJ, Lipp P, Bootman MD (**2000**) The versatility and universality of calcium signalling. *Cell* **1**.

[173] Hermans E (**2003**) Biochemical and pharmacological control of the multiplicity of coupling at G-protein-coupled receptors. *Pharmacology & Therapeutics* **99**:25–44. doi:10.1016/S0163-7258(03)00051-2

[174] Lutz S, Freichel-Blomquist A, Yang Y, Rümenapp U, Jakobs KH, et al. (**2005**) The guanine nucleotide exchange factor p63RhoGEF, a specific link between Gq/11-coupled receptor signaling and RhoA. *The Journal of biological chemistry* **280**:11134–9. doi:10.1074/jbc.M411322200

[175] Rhee SG, Bae YS (**1997**) Regulation of phosphoinositide-specific phospholipase C isozymes. *The Journal of biological chemistry* **272**:15045–8.

[176] Zimmermann S, Moelling K (**1999**) Phosphorylation and regulation of Raf by Akt (protein kinase B). *Science* **286**:1741–4.

[177] MacGillavry HD, Stam FJ, Sassen MM, Kegel L, Hendriks WTJ, et al. (**2009**) NFIL3 and cAMP response element-binding protein form a transcriptional feedforward loop that controls neuronal regeneration-associated gene expression. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **29**:15542–50.
doi:10.1523/JNEUROSCI.3938-09.2009

[178] McClean MN, Mody A, Broach JR, Ramanathan S (**2007**) Cross-talk and decision making in MAP kinase pathways. *Nature genetics* **39**:409–14.
doi:10.1038/ng1957

[179] Hu B, Rappel WJ, Levine H (**2009**) Mechanisms and constraints on yeast MAPK signaling specificity. *Biophysical journal* **96**:4755–63.
doi:10.1016/j.bpj.2009.02.065

[180] Thalhauser CJ, Komarova NL (**2009**) Specificity and robustness of the mammalian MAPK-IEG network. *Biophysical journal* **96**:3471–82.
doi:10.1016/j.bpj.2008.10.076

[181] Rensing L, Ruoff P (**2009**) How can yeast cells decide between three activated MAP kinase pathways? A model approach. *Journal of theoretical biology* **257**:578–587.
doi:10.1016/j.jtbi.2009.01.015

[182] de Ronde W, Tostevin F, ten Wolde PR (**2011**) Multiplexing Biochemical Signals. *Physical Review Letters* **107**:1–4.
doi:10.1103/PhysRevLett.107.048101

[183] Hoffmann A, Levchenko A, Scott ML, Baltimore D (**2002**) The IkappaB-NF-kappaB signaling module: temporal control and selective gene activation. *Science* **298**:1241–5.
doi:10.1126/science.1071914

[184] Li Y, Goldbeter A (**1989**) Frequency specificity in intercellular communication. Influence of patterns of periodic signaling on target cell responsiveness. *Biophysical journal* **55**:125–45.
doi:10.1016/S0006-3495(89)82785-7

[185] Wu Z, Chui CK, Hong GS, Chang S (**2011**) Physiological analysis on oscillatory behavior of glucose-insulin regulation by model with delays. *Journal of theoretical biology* **280**:1–9.
doi:10.1016/j.jtbi.2011.03.032

[186] Sasagawa S, Ozaki Yi, Fujita K, Kuroda S (**2005**) Prediction and validation of the distinct dynamics of transient and sustained ERK activation. *Nature cell biology* **7**:365–73.
doi:10.1038/ncb1233

[187] Batchelor E, Loewer A, Mock CS, Lahav G (**2011**) Stimulus-dependent dynamics of p53 in single cells. *Molecular systems biology* **7**:488.
doi:10.1038/msb.2011.20

[188] Armstrong SP, Caunt CJ, Fowkes RC, Tsaneva-Atanasova K, McCartney RR (**2009**) Pulsatile and sustained gonadotropin-releasing hormone (GnRH) receptor signaling: does the Ca2+/NFAT signaling pathway decode GnRH pulse frequency? *The Journal of biological chemistry* **284**:35746–57.
doi:10.1074/jbc.M109.063917

[189] Lim S, Pnueli L, Tan JH, Naor Z, Rajagopal G, et al. (**2009**) Negative feedback governs gonadotrope frequency-decoding of gonadotropin releasing hormone pulse-frequency. *PloS one* **4**:e7244.
doi:10.1371/journal.pone.0007244

[190] Dolmetsch RE, Xu K, Lewis RS (**1998**) Calcium oscillations increase the efficiency and specificity of gene expression. *Nature* **392**:933–6.
doi:10.1038/31960

[191] Tomida T, Hirose K, Takizawa A, Shibasaki F, Iino M (**2003**) NFAT functions as a working memory of Ca2+ signals in decoding Ca2+ oscillation. *The EMBO journal* **22**:3825–32.
doi:10.1093/emboj/cdg381

[192] Maeda M, Lu S, Shaulsky G, Miyazaki Y, Kuwayama H, et al. (**2004**) Periodic signaling controlled by an oscillatory circuit that includes protein kinases ERK2 and PKA. *Science* **304**:875–8.
doi:10.1126/science.1094647

[193] Garmendia-Torres C, Goldbeter A, Jacquet M (**2007**) Nucleocytoplasmic oscillations of the yeast transcription factor Msn2: evidence for periodic PKA activation. *Current biology* **17**:1044–9.
doi:10.1016/j.cub.2007.05.032

[194] Hilioti Z, Sabbagh W, Paliwal S, Bergmann A, Goncalves MD, et al. (**2008**) Oscillatory phosphorylation of yeast Fus3 MAP kinase controls periodic gene expression and morphogenesis. *Current biology* **18**:1700–6.
doi:10.1016/j.cub.2008.09.027

[195] Ashall L, Horton CA, Nelson DE, Paszek P, Harper CV, et al. (**2009**) Pulsatile stimulation determines timing and specificity of NF-kappaB-dependent transcription. *Science* **324**:242–6.
doi:10.1126/science.1164860

[196] Werner SL, Barken D, Hoffmann A (**2005**) Stimulus specificity of gene expression programs determined by temporal control of IKK activity. *Science* **309**:1857–61.
doi:10.1126/science.1113319

[197] Sung MH, Salvatore L, De Lorenzi R, Indrawan A, Pasparakis M, et al. (**2009**) Sustained oscillations of NF-kappaB produce distinct genome scanning and gene expression profiles. *PloS one* **4**:e7163.
doi:10.1371/journal.pone.0007163

[198] Ballif BA, Blenis J (**2001**) Molecular mechanisms mediating mammalian mitogen-activated protein kinase (MAPK) kinase (MEK)-MAPK cell survival signals. *Cell growth & differentiation : the molecular biology journal of the American Association for Cancer Research* **12**:397–408.

[199] Nakakuki T, Birtwistle MR, Saeki Y, Yumoto N, Ide K, et al. (**2010**) Ligand-Specific c-Fos Expression Emerges from the Spatiotemporal Control of ErbB Network Dynamics. *Cell* **2**:884–896.
doi:10.1016/j.cell.2010.03.054

[200] Friedland AE, Lu TK, Wang X, Shi D, Church G, et al. (**2009**) Synthetic gene networks that count. *Science* **324**:1199–202.
doi:10.1126/science.1172005

[201] Tsai TYC, Choi YS, Ma W, Pomerening JR, Tang C, et al. (**2008**) Robust, tunable biological oscillations from interlinked positive and negative feedback loops. *Science* **321**:126–9.
doi:10.1126/science.1156951

[202] Stricker J, Cookson S, Bennett MR, Mather WH, Tsimring LS, et al. (**2008**) A fast, robust and tunable synthetic gene oscillator. *Nature* **456**:516–9.
doi:10.1038/nature07389

[203] Tostevin F, de Ronde W, ten Wolde PR (**2012**) Reliability of frequency and amplitude decoding in gene regulation. *In preparation* .

[204] Fei P, El-Deiry WS (**2003**) P53 and radiation responses. *Oncogene* **22**:5774–83.
doi:10.1038/sj.onc.1206677

[205] Sengupta S, Harris CC (**2005**) p53: traffic cop at the crossroads of DNA repair and recombination. *Nature reviews Molecular cell biology* **6**:44–55.
doi:10.1038/nrm1546

[206] Shaul YD, Seger R (**2007**) The MEK/ERK cascade: from signaling specificity to diverse functions. *Biochimica et biophysica acta* **1773**:1213–26.
doi:10.1016/j.bbamcr.2006.10.005

[207] Kholodenko BN (**2000**) Negative feedback and ultrasensitivity can bring about oscillations in the mitogen-activated protein kinase cascades. *European journal of biochemistry / FEBS* **267**:1583–8.

[208] von Kriegsheim A, Baiocchi D, Birtwistle MR, Sumpton D, Bienvenut W, et al. (**2009**) Cell fate decisions are specified by the dynamic ERK interactome. *Nature cell biology* **11**:1458–64.
doi:10.1038/ncb1994

[209] Wei CL, Wu Q, Vega VB, Chiu KP, Ng P, et al. (**2006**) A global map of p53 transcription-factor binding sites in the human genome. *Cell* **124**:207–19.
doi:10.1016/j.cell.2005.10.043

[210] Behar M, Hao N, Dohlman HG, Elston TC (**2008**) Dose-to-duration encoding and signaling beyond saturation in intracellular signaling networks. *PLoS computational biology* **4**:e1000197.
doi:10.1371/journal.pcbi.1000197

[211] Segel LA (**1988**) On the validity of the steady-state assumption of enzyme kinetics. *Bulletin of mathematical biology* **50**:579–593.

[212] Rami Tzafriri A, Edelman ER (**2007**) Quasi-steady-state kinetics at enzyme and substrate concentrations in excess of the Michaelis-Menten constant. *Journal of theoretical biology* **245**:737–48.
doi:10.1016/j.jtbi.2006.12.005

[213] Mengel B, Hunziker A, Pedersen L, Trusina A, Jensen MH, et al. (**2010**) Modeling oscillatory control in NF-κB, p53 and Wnt signaling. *Current Opinion in genetics & development* **20**:656–64.
doi:10.1016/j.gde.2010.08.008

Summary

All living systems have to respond to changes in their environment — this capacity could even be phrased as a defining property of a living system. In order to respond to a change in the environment, the first step is to register that a change has occurred. We humans have developed five senses, *seeing, hearing, feeling, smelling, tasting*, that allow us to actively register the environment and note changes within it. Not only for humans is this capacity important, but also for more simple forms of life, including unicellular organisms like bacteria. And indeed, also unicellular organisms have developed "tools" that allow them to measure changes in their environment.

However, monitoring the environment by itself is not sufficient. Cells also actively have to process the information from the environment and, possibly, change their current behavior. If for example *E. coli*, initially surrounded by an excess of glucose, observes that the glucose concentration decreases in favor of lactose, in order to survive, it has to change its behavior. The bacterium for example could change its internal metabolic circuit from glucose-powered to lactose-powered by expressing different enzymes, or decide to move to different places with possible higher concentrations of glucose. In both scenarios, the change in the environment leads to an active response of the bacterium.

The connection between the change in the environment, the signal, and the response of the cell is facilitated by networks of interacting proteins, which jointly form a signaling cascade. Cells have many different signaling cascades, since there are many different signals and signals can lead to a variety of responses. Arguably, the most important property of a signaling cascade is the transfer of information from the signal to the response as reliably as possible. To study the reliability quantitatively I use information theory, since this provides us with an excellent measure, the mutual information between signal and response. Using information theory it can be shown that the reliability of the information transfer depends upon the ratio of two quantities, the amplification of the signal, or the gain, and the inherent noise of any biochemical cascade. In this thesis we will study these quantities, the gain and the noise, in more detail, especially in *Chapters 2, 3, 4* and *6*.

Increasing the reliability of a cascade in general requires more energy, for example for the production of more proteins or additional phosphorylations. Since energy resources are limited for every cell, a large reliability therefore should be obtained while keeping the cost low. One possible solution for cost reduction is the simultaneous use of a single protein for multiple functions or in multiple cascades, a process called multiplexing. In *Chapters 5, 7* and *8* this idea is worked out in more detail.

Reliability of signaling cascades

In *Chapters 2* and *3* I look at the reliability of signaling cascades for time-varying signals. Recent studies have shown that some small network structures are overrepresented within signaling cascades, so-called network motifs, and for these motifs I study the reliability. The reliability is quantified using the mutual information between signal and response. For time-varying continuous signals, the mutual information is replaced by the mutual information rate, which i) is frequency dependent, and ii) is directly proportional to the gain-to-noise ratio of the cascade, which now is frequency dependent as well. By studying both the gain (amplification) and the noise simultaneously, we observe that positive feedback or positive autoregulation increases the gain-to-noise ratio at small frequencies, showing

that slowly varying signals can be transmitted more reliably if positive regulation is present in the cascade. Negative feedback enhances the reliability of signaling at high frequencies, but only if the feedback is within the cascade, upstream of the final response. Then negative feedback can even lead to a peak in the gain-to-noise ratio as a function of the signal frequency.

For cascades with feed-forward properties (either coherent, incoherent or diamond) I observe that the gain-to-noise ratio is enhanced if the signal is split in two separate intermediate proteins, which then later recombine at the response. However, to observe this, coincidence detection at the response is required, meaning that for activation of the response both intermediates should be present simultaneously. This form of activation can be obtained both through homo- or heterodimerization of the intermediate proteins. Next, in this chapter, we observe that for a feed-forward motif the gain-to-noise function can have high-pass or low-pass characteristics as a function of frequency which, surprisingly, does not depend on the type of feed-forward motif, either coherent or incoherent. By actively changing the different coupling strenghts between the two pathways in a feed-forward motif, the gain-to-noise ratio can be switched from a low-pass to a high-pass filter and vice-versa.

Diffusion noise

Many signaling cascades have as their natural starting point a receptor, although this is not always true. These receptors sense a specific ligand concentration which diffuses in the medium surrounding the receptor and stochastically binds and unbinds to the receptor. Over 35 years ago, an interesting study by Berg and Purcell has provided an answer to the question what the minimal limit in the uncertainty is of a concentration measurement if such a single receptor would measure the concentration over some time interval. This limit has been disputed in later papers, creating an open question. In *Chapter 4*, using a different analytical derivation and computer simulations, we confirm the findings of Berg and Purcell of 35 years ago, establishing, again, the fundamental limit on noise in a concentration measurement. The limit consists of two terms, one originating from the stochastic dynamics of the receptor and one from the diffusive behavior of the ligand particles. Importantly, the fundamental limit depends on the fractional occupancy of the receptor. This reflects the fact that a receptor that is bound, can not measure "new" ligand molecules, thereby increasing the uncertainty in the measurement of the outside concentration. Moreover, in this chapter we present a simple model which provides the correct correlation time, and therefore the correct measurement uncertainty, and the parameter range for which the simple model is accurate.

Signal integration

Focussing on the receptors in *Chapter 5* I study signal integration by receptors. Recent experiments have shown that receptors, or in general proteins, can act as logic gates, meaning that the response depends on the combined information of two signals. We wondered how versatile this mechanism is. Using a simple statistical mechanics model we show that indeed receptors are capable of performing any gate by tuning kinetic parameters. On evolutionary timescales, this would be a mechanism for cells to obtain receptors which function as any specific logic gate. As interesting, we show that on much shorter timescales, that of protein signaling, all logic gates can be formed from a limited set of receptor monomers that dimerize.

Oscillatory signals

It is increasingly recognized that cells often use oscillatory signals to transmit information. Intuitively, one would expect an oscillatory signal to increase the variability in the output compared to a constant signal. Since this increase in variability makes the signaling cascades less reliable, the use of oscillatory signals seems, from this point of view, disadvantageous. However, in *Chapter 6* we show that, counter-intuitively, oscillatory signals not necessarily lead to lower variability in the output of a gene regulatory network. This effect relies on the fact that a switch driven by an oscillatory signal, becomes more periodic, than when driven by a constant signal. The oscillatory input signal then leads to a more constant output level than the constant input signal.

Multiplexing

Multiplexing, the art of transmitting multiple different signals simultaneously through a shared pathway, is the topic of the final two chapters of this thesis. In *Chapter 7* the problem of multiplexing constant signals is studied. Here we show that two signals can indeed be encoded into a shared cascade and decoded into two responses, where each response is sensitive to a single signal only, with absolute reliability. Since the input signals are constant in time, we refer to this type of multiplexing as AM multiplexing.

In *Chapter 8* we extend the problem as posed in Chapter 7, by asking the question if oscillatory signals and constant signals jointly can be multiplexed. Here one response is sensitive to the properties of the oscillatory signal (like the oscillation period and the amplitude, but not the mean level), while the other response is sensitive to the concentration of the constant signal. We show that indeed networks can be constructed that multiplex these signals, and as expected, we find that, next to the intrinsic biochemical noise, also the cross-talk between the two signals acts as a bottleneck for information. However, the total amount of information through the network can be increased enormously with respect to the AM multiplexing strategy.

Samenvatting

Ieder levend wezen moet reageren op veranderingen in zijn omgeving — sterker nog, de mogelijkheid om dit te kunnen doen zou een definieerbare eigenschap van leven kunnen worden genoemd. Om überhaupt te kunnen reageren op veranderingen in de omgeving, moet allereerst de omgeving worden waargenomen, om zo een eventuele verandering op te merken. Mensen hebben voor deze taak zintuigen ontwikkeld, *zien, horen, voelen, ruiken en proeven*, die het mogelijk maken om actief, maar misschien wel vaker onbewust, de omgeving waar te nemen en veranderingen op te merken. Maar niet alleen mensen moeten hun omgeving waar nemen, ook eenvoudigere levensvormen, inclusief de ééncelligen, zoals bacteriën, moeten dit doen. Ook ééncellige organismen hebben, via evolutionaire processen, het gereedschap ontwikkeld dat nodig is om de de eigenschappen van de omgeving te meten.

Echter, alleen het actief waarnemen van de omgeving is onvoldoende. Naast de waarneming, moeten cellen de verkregen informatie over de omgeving ook verwerken, wat (vaak) tot een gedragsverandering leidt. Als bijvoorbeeld de bacterie *E. coli*, in beginsel omringd door de suiker glucose, opmerkt dat de glucose suikers worden vervangen door lactose suikers, dan moet de bacterie "actie" ondernemen om in leven te blijven. Een mogelijkheid is een aanpassing van het interne metabole circuit door de expressie van nieuwe, andere enzymen, zodat in plaats van glucose, nu lactose de energiebron wordt, en een andere mogelijkheid is om zich te verplaatsen naar regio's met een hogere concentratie aan glucose suikers. Belangrijk is dat in beide scenario's een verandering van de omgeving tot een actieve respons van de bacterie leidt.

In deze context is de veranderende omgeving het signaal, en de gedragsverandering, de respons. In cellen zijn signaal en respons verbonden door netwerken van verschillende eiwitten, die onderling reageren, bijvoorbeeld door phosphorylatie of dimerisatie. Een netwerk van zulke eiwitten wordt een signaalcascade genoemd, en signaalcascades zijn een belangrijk onderwerp van dit proefschrift. Cellen hebben veel verschillende signaalcascades, omdat er veel verschillende signalen zijn en omdat één signaal tot verschillende responses kan leiden. Ik durf te stellen dat de betrouwbaarheid van een signaalcascade het belangrijkste criterium is om goed te functioneren, maar andere criteria, zoals transmissiesnelheid of signaalintegratie, kunnen ook een rol spelen. De betrouwbaarheid van een cascade hangt van twee factoren af, enerzijds de versterking van het signaal, de zogenaamde gain, en anderzijds de intrinsieke ruis, die inherent aanwezig is in elke biochemische cascade. De betrouwbaarheid kan zeer goed gekwantificeerd worden met behulp van informatietheorie; immers een belangrijke resultaat binnen de informatietheorie is de gedeelde informatie tussen signaal en respons. Uit de informatietheorie blijkt inderdaad dat de betrouwbaarheid van een signaalcascade afhangt van de verhouding van de gain en de ruis. In dit proefschrift worden deze factoren, de gain, de ruis en gedeelde informatie, in meer detail bestudeerd, voornamelijk in de *Hoofdstukken 2, 3, 4* en *6*.

Het vergroten van de betrouwbaarheid is in principe een kostbare zaak. Energie is nodig om bijvoorbeeld extra eiwitten te maken voor een grotere versterking, of voor een grote aantal phosphorylatie stappen. Echter, energie is een schaars goed, en een hoge betrouwbaarheid moet daarom worden verkregen met minimaal energiegebruik. In andere woorden, een hoge betrouwbaarheid voor een lage kostprijs. Een mogelijkheid voor het reduceren van de kosten, is het hergebruiken van eenzelfde eiwit binnen verschillende cascades, een proces dat multiplexen wordt genoemd. In de *Hoofdstukken 5, 7* en

8 wordt hier in detail naar gekeken.

Betrouwbaarheid van signaalcascades

In de *Hoofdstukken 2 en 3* bestudeer ik de betrouwbaarheid van signaalcascades for signalen die tijdsafhankelijk zijn. Recente studies hebben aangetoond dat bepaalde kleine netwerkstrucuren oververtegenwoordigd zijn binnen grotere signaalcascades, zogenaamde netwerkmotieven, en voor deze kleine motieven bestudeer ik de betrouwbaarheid met behulp van de gedeelde informatie tussen het signaal en de respons. Omdat de signalen tijdsafhankelijk zijn, vervangen we de gedeelde informatie, door de zogenaamde gedeelde informatiesnelheid, die frequentie afhankelijk is. Verder is deze gerelateerd aan de verhouding van de gain en de ruis, maar nu zijn zowel de gain als de ruis ook frequentie afhankelijk. Door tegelijkertijd de gain en de ruis van een motief te bestuderen, blijkt dat voor een motief met positieve autoregulatie of positieve terugkoppeling, de informatieoverdracht van langzaam variërende signalen wordt vergroot. Echter, voor motieven met negatieve terugkoppeling wordt de informatie overdracht voor snel variërende signalen vergroot, maar alleen als de terugkoppeling zich "boven" (dichter bij het signaal) de uiteindelijke respons bevindt. In dat geval kan negatieve terugkoppeling zelfs leiden tot een maximum in de gain-ruis verhouding als functie van de signaalfrequentie.

In cascades met een feed-forward motif (coherent, incoherent of diamant) wordt de gain-ruis verhouding vergroot als het initiële signaal wordt gesplitst in twee gescheiden tussenliggende componenten of signaalpaden, die bij de uiteindelijke respons weer samenkomen. Echter, dit effect, de verhoogde gain-ruis verhouding, wordt alleen waar genomen bij zogenaamde gelijktijdige detectie, wat betekent dat voor activatie van de respons beide tussenliggende componenten/signaalpaden tegelijkertijd aanwezig moeten zijn. Bij zowel het homo- als heterodimerisatie proces van eiwitten is gelijktijdige detectie bijvoorbeeld aanwezig. Daarnaast laten we in dit hoofdstuk ook zien dat in een feed-forward motief de gain-ruis verhouding als een functie van de frequentie, zowel een karakteristieke hoog-band als een laag-band vorm kan hebben. Verrassend genoeg hangt dit niet af van het type feed-forward motief, want zowel het coherente als het incoherent motief kan beide karakteristieke vormen hebben. Door het actief regulieren van de koppeling tussen de respons en de signaalpaden is het mogelijk om de gain-ruis verhouding van een hoogband naar een laagband karakteristiek en vice versa te veranderen.

Ruis door diffusie

Het natuurlijke startpunt van vele signaalcascades is een receptor, al is dit niet altijd het geval. Het doel van een receptor is het registreren van de concentratie van een bepaalde ligand. Dit is niet zo eenvoudig, omdat het ligand diffundeert in het medium dat de receptor omgeeft en stochastisch bindt en ontbindt aan de receptor. Meer dan 35 jaar geleden hebben Berg en Purcell een antwoord geformuleerd op de vraag wat de minimale onzekerheid in de ligandconcentratie is wanneer deze door één receptor over een langere tijd wordt gemeten. Dit resultaat, de minimale onzekerheid, is echter in latere jaren betwist en is daarom een open vraag. In *Hoofdstuk 4* bevestigen wij de resultaten van Berg en Purcell, door middel van een nieuwe analytische theorie en computer simulaties, en daarmee herstellen wij de fundamentele Berg-Purcell limiet weer in ere. Deze limiet bestaat uit twee termen; de eerste term komt van de stochastische dynamica van het bind- en ontbindproces, de tweede term van het diffuse gedrag van de ligandmoleculen. Een belangrijk kenmerk van de limiet is de afhankelijkheid van de onzekerheid van de relatieve bezettingsgraad van de receptor. Deze afhankelijkheid reflecteert dat een receptor die relatief vaak gebonden is, relatief weinig "nieuwe" moleculen kan waarnemen en daardoor een hogere onzekerheid in de concentratie meting geeft. Daarnaast geven wij in dit hoofdstuk nog een eenvoudig model dat zeer goed de correlatietijd, en daarmee de onzekerheid in de

concentratie verklaart, en het parametergebied waarbinnen dit eenvoudige model geldig is.

Signaalintegratie

Nu we toch de receptor aan een kritische blik onderwerpen, bekijken we in *Hoofdstuk 5* naar signaalintegratie door receptoren. Recente experimenten hebben laten zien dat receptoren, of, meer in het algemeen, eiwitten, logische functies kunnen uitvoeren. Dit wil zeggen dat de uiteindelijke respons afhankelijk is van de geïntegreerde informatie van twee signalen. Wij waren benieuwd hoe flexibel dit mechanisme is. Met een simpel model uit de statistische mechanica kunnen we aantonen dat receptoren inderdaad als elke mogelijk logische poort kunnen werken, door het actief reguleren van kinetische parameters. Op evolutionaire tijdschalen is dit een mechanisme waarbij cellen receptoren kunnen creeëren die als elke logische poort zijn te gebruiken. Echter, en misschien nog interessanter, we laten ook zien dat op veel kortere tijdschalen cellen alle logische poorten kunnen verkrijgen door het combineren van een beperkte set receptoren in dimeren.

Oscillerende signalen

Steeds meer experimenten laten zien dat cellen gebruik maken van oscillerende signalen. In eerste instantie zou men verwachten dat een oscillerend signaal de variabiliteit van de respons vergroot in vergelijking met een niet-oscillerend signaal. De vergroting van de variabiliteit, en dus de ruis, maakt een signaalcascade minder betrouwbaar en daarom lijkt het toepassen van oscillerende signalen, vanuit het oogpunt van de betrouwbaarheid, onlogisch. Echter in *Hoofdstuk 6* laten we zien, tegen de intuïtie in, dat in een genregulatie netwerk oscillerende signalen niet noodzakelijkerwijs leiden tot een lagere variabiliteit in de respons. Dit effect is mogelijk omdat een genetische schakelaar, aangedreven door een oscillerend signaal, een meer periodiek tijdsgedrag vertoont, dan aangedreven door een constant signaal. Het gevolg hiervan is dat het oscillerende signaal een minder variabele respons genereert dan het constante signaal.

Multiplexen

Multiplexen, of de kunst van het tegelijkertijd doorsturen van meerdere signalen door een gedeeld signaalpad is het onderwerp van de laatste twee hoofdstukken van dit proefschrift. In *Hoofdstuk 7* wordt het probleem van het multiplexen van constante signalen bestudeerd. We laten zien dat twee signalen inderdaad kunnen worden geëncodeerd in een gedeeld signaalpad en vervolgens gedecodeerd in twee uitgangssignalen, zodanig dat elk uitgangssignaal alleen afhankelijk is van zijn eigen ingangssignaal, met een 100% betrouwbaarheid. Omdat de ingangssignalen constant zijn, noemen we deze vorm van multiplexen AM multiplexen.

In *Hoofdstuk 8* vergroten we de vraagstelling van hoofdstuk 7. In plaats van ons alleen te richten op constante signalen, richten we ons nu op de vraag of één constant en één oscillerende signaal kunnen worden gemultiplexed. Dit kan, als de respons van het oscillerende signaal alleen afhankelijk is van de kenmerken van het oscillerende signaal (zoals de periode of de oscillatie amplitude, maar niet de gemiddelde concentratie), terwijl de respons van het constante signaal alleen afhankelijk is van de concentratie. We laten zien dat het mogelijk is een netwerk te formeren dat aan deze voorwaarde voldoet. Zoals verwacht, is niet alleen de intrinsieke biochemische ruis een belangrijke bottle-neck voor informatietransmissie, maar ook de kruisbestuiving tussen de twee verschillende signaalpaden. Echter, met een oscillerend en een constant signaal kan de totale hoeveelheid informatie die door het netwerk wordt gestuurd, enorm vergroot worden ten opzichte van AM multiplexing.

Terugkijkend op

Terugkijkend op 4 jaar AMOLF besef ik me dat ik een hele intense en bijzondere tijd heb beleefd. AMOLF, door alle mensen, is een enorm inspirerende en verrijkende periode in mijn leven geweest, waar ik warme, vrolijke en hartelijke gevoelens bij heb. Inderdaad, zoals Godfried Bomans schreef

" Geluk wordt pas zichtbaar als het voorbij is."

Pieter Rein, 4 jaar onder jouw hoede, op zoek naar de wetenschapper in mij. Je enorme interesse in de wetenschap heeft me gestimuleerd en soms verbaasd en je motivatie om samen problemen op te lossen is ongekend. Daarvoor in de eerste plaats mijn grote dank. Met jouw hulp heb ik die wetenschapper gevonden, en ik geniet daar elke dag van. Niet alleen daarom weet ik zeker dat jij, voor mij, de perfecte promotor bent. Uitgedaagd en geholpen heb je me altijd, en daarnaast heb ik prachtige herinneringen aan extra-academische hoogtepunten. De wind kwam soms uit alle richtingen, maar dankzij jou was promoveren afkruisen naar de benedenwindse boei…

Vele duwtjes kreeg ik, en de laatste duw richting de promotie kreeg ik van Harm. Subtiel in het begin, en later ook ondersteunend in de dalen van mijn wetenschappelijke speurtocht. Ik ben er uiteindelijk in de wolken mee…

Filipe, your share in this thesis has been enormous. I can not thank you enough for this. Your sharp thinking makes you a fabulous scientist, and your British sense of humor irresistibly funny.

Andrew, also you I would like to thank. First in Los Alamos and Boston, later in Amsterdam I have benefited and learned greatly from your scientific skills. Your critical mind comes with your many humorous remarks.

Anton en Kaizu, the time we spend together at AMOLF was short, but I liked your enthusiasm for the problems we were studying and your contributions to the research in this thesis.

Guys, I have enjoyed being a small wheel in the clockwork of the Biochemical Networks Group, the curious interests, the passion for science, the urge to really understand, but also the drive to share a laugh and have fun together. Tomek, Nils, Laurens, Mark, Martijn, Joris, Chris, Tommy, David, Thorsten, Siebe, Rutger and Christian: thanks for this time. And of course: Amy en Sanne.
Natuurlijk wil ik alle mensen van de "biokant" bedanken voor een fantastische sfeer, zodat de dagelijkse routine never a dull moment was, en bijzondere herinneringen heb ik daarnaast aan Bela, Jose, Marina, Pierre, Liedewij, Paige, Tom, Martin, Christiaan, Marjon, Daan, Gianluca, Marjolein, Ana en Ernst Jan.

Samen met Sanne heb ik de afgelopen jaren vele, mooie reizen gemaakt; ver weg, maar juist ook dichtbij. Met geluk denk ik daaraan terug, en voor al die reizen ben ik jou erg dankbaar. Ik wens jou bijzondere reizen in de toekomst.

Lieve Joost-Marijn, Floor, Katrijn, papa en mama, alle steun die ik van jullie heb gekregen maakte mijn roller coaster soms wat draaglijker… weet dat ik ook van jullie houd.

Van terugkijken op naar vooruitkijken naar, de toekomst zonnig blauw als de voorkant samen met de passie van het rood van de binnenkant.
Ik sta te trappelen…

Wiet de Ronde
Februari 2012