Stochasticity in signal transduction pathways

Vidal Rodriguez, J.

Publication date
2009

Chapter 4

Response Time in Two-Component Signalling Systems

It is widely accepted that an important survival mechanism developed primarily in prokaryotic cells is the sensing of environmental stimuli in order to adapt and thrive under changing conditions (Mascher et al., 2006). In bacteria two-component signalling systems (TCS) are the most common sensing mechanism for external stimuli Nixon et al. (1986) ; Hoch (2000) ; Koretke et al. (2000) ; Wolanin et al. (2002) ; Bekker et al. (2006). Two-component signalling systems are a specific case related to those studied in chapter 3. Sensors are embedded in the plasma membrane, as with PTS and chemotaxis. However, unlike with the chemotactic cluster, sensors are usually found isolated and scattered throughout the plasma membrane, and consequently there is no co-operative kinetics among them. Moreover, the number of most sensory molecules in TCSs are thought to be small (10 molecules of histidine-kinase EnvZ in *E. coli* (Stock et al., 1989 ; Mizuno et al., 1982). Transcription factors concentration levels are similar to those used in the gene expression model. In contrast to the PTS, cytosolic molecule numbers are usually low, ranging from tens to a few hundred. Only TCSs transcription factors need to receive a phosphoryl group from a sensor in order to bind the promoter.

As we introduced in section 1.5, our prototypical sensory system design consists of only two chemical species: the first being a sensor, usually a histidine-kinase (HK); the second being a response regulator (RR), the mode of transport of the signal to the receiver system. In most cases, the receiver of a signal may be one or more genes, for which an RR is also referred to as a transcription

The contents of this chapter are based on the results published in

- Dobrzyński et al. Prokaryotic signaling has been optimzed for quick but robust response. Submitted (2009)

factor (TF).

Delivering a signal entails the completion of two transport, diffusion-based, processes: first, an inactive TF receives a phosphoryl group from an active HK upon a reactive collision; second, the now-active TF needs to encounter and react with the promoter region. These search processes rely on the diffusive motion of TF throughout the cytosol. TCSs also have features such as localised reactions on the membrane, and a single static target molecule in the cytosol. This low number of molecules suggests that these systems may display stochastic features similar to those found in gene expression (i.e. non-exponential distribution). The signalling process is similar to gene expression regulation, only that now transcription factors need to be activated first by molecules embedded on the plasma membrane.

Despite the ubiquitous presence of two-component signalling systems in bacteria and archaea (and few eukaryotes), few studies consider their fundamental design properties at the mesoscopic level. To this end, we focus on a minimalistic model with emphasis on four structural properties: spatial organisation of sensors; numbers of histidine-kinases and transcription factors; and the diffusive motion of transcription factors. We argue that the features of these principles are not completely free to vary; instead, they are most likely constrained for efficiency. Efficiency is to be understood in terms of how quickly a bacterium may respond to environmental stimuli by using the lowest number of molecules. In this chapter, we quantify this efficiency by the *response time* of the signalling system which is defined as the time elapsed between the activation of sensors and the delivery of the signal to a receiver system. We shall emphasise that this particular definition does not account for the processes that take place in the receiver system as to execute the information receive. For instance, if the receiver system is gene expression, the response time is the time elapsed between the first activation of the sensors till the first transcription factor binds the promoter. The transcription and translation processes, thus, are not included.

We aim at understanding the limits of the response time that arise from the interplay of only fundamental common design principles, not adjusted to any system in particular. Consequently, we refrain from using a fully kinetic modelling approach, as this, would require detailed knowledge at the mesoscopic level for each particular system. Instead, reactions are driven exclusively by the diffusive encounter rate, in other words, reactions are diffusion limited. However, the model may be further extended to accept reaction-limited rates. The fundamental signalling processes have been investigated in similar forms in Berg and Purcell (1977) ; Redner (2001) ; il Lim and Yin (2005) ; Lu et al. (2006) ; van Zon et al. (2006) ; Holcman and Schuss (2004) ; Taflia and Holcman (2007) ; Holcman and Schuss (2008). Those more theoretical oriented studies, however, differ from the presented here in that they did not tackle the two-component signalling in the detail we present here.

The remainder of this chapter is organised as follows. In the next section, we formalise the two-process, two-component signalling system. In section 4.2 we analyse in detail and separately the influences that spatial distribution and number of histidine-kinases, or sensors, have on the phosphorylation process, as

well as the timing of the diffusive transport from the membrane to a target at the centre of the cell. The response time is analysed under two different types of constraints in section 4.3. The first type of constraint assumes that the total number of molecules is fixed. The second, that their numbers are correlated. Using these constraints sheds light onto the efficiency of the design principles, so as to minimise the response time without using an excess of molecules.

4.1 The Two-Component Signalling System Model

Our model of two-component signal transduction (TCS) is written in Fig. 4.1, and follows the schema described in section 1.5.1 and in Hoch (2000).

The two-reaction model is:

$$TF + HK^* \rightarrow HK + TF^* \tag{4.1}$$

$$TF^* + P_{gene} \rightarrow TF^* \cdot P_{gene}, \tag{4.2}$$

where the first reaction is the phosphoryl-group transfer (*) from an active histidine-kinase (HK^*) to an inactive transcription factor (TF). In the second reaction, the now active TF^* may bind to its target. We may assume it is a promoter site located at the centre of the cell. Just as with the gene expression models used in section 3.1, it remains immobile, since the movement of the nucleoid is assumed be much slower. Transcription factor molecules are the only mobile molecules in the system.

The goal of this model is to understand the elementary processes that shape the response time (τ_{resp}). To this end, we minimise the number of parameters in the model to describe both the spatial and the diffusion-limited kinetic properties of the system. In first instance, we use a spherical cell of radius R and volume V, which for an *E. coli* it is often taken to be $R = 1\,\mu$m. (It seems that the cell's shape does not qualitatively affect the results, be it cigar-shaped (Blom and Peletier, 2002) or, hypothetically, cubical (Nagar and Pradhan, 2003).) We assume that all HKs are in their active state throughout the whole simulation; therefore, their output domain is ready to transfer a phosphoryl group to the input domain of a transcription factor. This initial condition implies that the signal sensed is present in a sufficiently high concentration and for a sufficiently extended period of time to maintain all sensors activated. Likewise, it is assumed that a phosphorylated transcription factor does not undergo autodephosphorylation.

Furthermore, we idealise a histidine-kinases as spherical caps with a radius r_{hk} on the surface of the cell which are distributed randomly. (HKs diffuse laterally, about two orders of magnitude slower than RRs according to Deich et al. (2004).) Note that the membrane surface of a spherical cap is $2\pi R(R + \sqrt{R^2 - r_{hk}^2})$, which for very small $r_{hk} \ll R$ is approximated by $2\pi r_{hk}^2$, the surface of a two-dimensional disk.

Transcription factors are represented by spherical particles with a radius r_{tf} undergoing normal diffusion throughout the cytosolic space. For the cell's membrane we use a reflective boundary condition — that is, when a particle

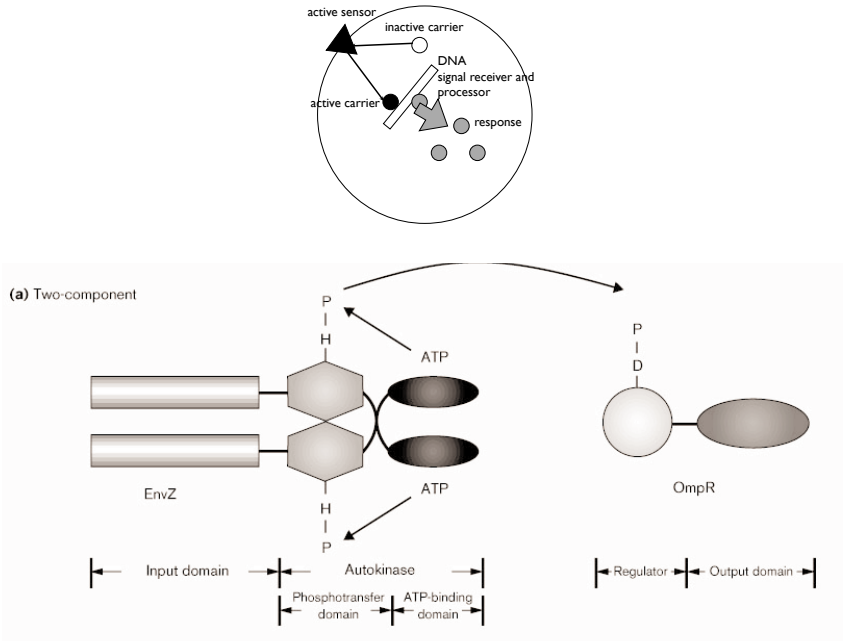

Figure 4.1: *Schema of a generic two-component signal transduction pathway where the receiver system is a gene expression system. The signal is detected by a transmembrane sensor which transfers a phosphoryl group extracted from ATP to an inactive response regulator. The active response regulator diffuses freely through the cytoplasm until it binds to the receiver system, often a promoter site on the DNA. (bottom) Details of the different composing units in histidine-kinases (left) and response regulators (right).*

wants to jump to a lattice site farther than the cell's radius R, the jump is not executed and the particle remains unchanged. Phosphorylation of a transcription factor occurs upon collision with any HK. In the simulations this happens when a TF tries to cross the boundary of the cell and the distance to a HK is shorter than 5 nm. This is a special case of the boundary condition, as the state of the TF changes to phosphorylated. Binding between TF* and the promoter site, also idealised as a small sphere, occurs when the distance between both particles is $< 5\,\mathrm{nm}$, which is equivalent to the radius of the promoter site r_{promoter}. Note that this is the typical diameter of a protein in *E. coli* (van Zon et al., 2006). In order to obtain accurate results of the first-passage time we use a fine lattice with a jump length of $\lambda = 1\,\mathrm{nm}$. The diffusion coefficient for TFs is assumed to be $D = 5\,\mu\mathrm{m}^2 s^{-1}$, and we use Eqn. 2.1 to obtain the time for each diffusion step.

Clusters containing, say, k histidine-kinase sensors are also modelled as a spherical cap, instead of a cluster of adjacent k spherical caps, with a total

surface area equal to the sum of k individual sensors. Thus the radius of a cluster scales with $k^{-1/2}$ times the area of one single histidine-kinase. The membrane's fraction covered by sensors is designated as Φ. The total cell's membrane fraction covered by sensors or a clusters of sensors is designated as Φ.

4.2 Analysis of the Individual Processes

In this section we first analyse the effects that the number and distribution of histidine-kinases have on the phosphorylation time. This is referred to as the first process. The second process, that of promoter activation, is only influenced by the location of the target, not the spatial distribution of sensors.

4.2.1 Scattered Sensors Reduce Response Time

The first process in the signal transduction is the phosphorylation of inactive transcription factors by histidine-kinases, also referred to as sensors. Schuss et al. (2007) showed that the average time for one particle (one TF) to hit for the first time one small area on the surface of a sphere, $\Phi \ll 1$, is to a first order approximation

$$\tau_{\text{phos},1} = \frac{V}{4D\epsilon} = \frac{V}{4D\Phi^{1/2}}, \tag{4.3}$$

where ϵ is the radius of small opening, and the subscript indicates that only one diffusing particle is involved. We use our simulations to extend that range and analyse the limits for which that result remains valid and how it is affected by the spatial organisation of sensors.

Initially, using one cluster, we vary the surface coverage Φ from 0.01 to 1. In Fig. 4.2 we show the time it takes for a transcription factor to reach a sensor, starting at a random position, relative to the time taken when the whole membrane is a sensor. The solid line shows that the scaling in Eqn. 4.3 remains a good approximation for a relatively large sensor (up to $\Phi = 0.1$). We should note that this fraction coverage is rather large when compared to the size of a single sensor. Assuming an effective diameter of 5 nm for a sensor on the membrane of an *E. coli*, the fraction the sensor covers is very small (approximately $\Phi = 10^{-6}$). Thus the approximation given in Eqn. 4.3 holds for cluster containing up to approximately 10^4 individual sensors. This is a large number, considering that the number of aspartate-Tar-CheA complexes in the chemotactic cluster is approximately 1200 (as used in the model in section 3.3 and in Lipkow et al. (2005) with the Smoldyn tool).

Dividing up the single cluster into many parts (i.e. 10, 100 and 400) broadens the range of the scaling in the direction of a larger Φ (dotted, dashed and dash-dot lines respectively in Fig. 4.2). More interesting is that at equal coverage splitting the cluster into smaller clusters and scattering them randomly reduces the phosphorylation time. Additionally, in Dobrzynski et al. (2009) we find that the phosphorylation time scales with the inverse of the number of sensors. These, however, shall be small and not close to each other in order to be independent

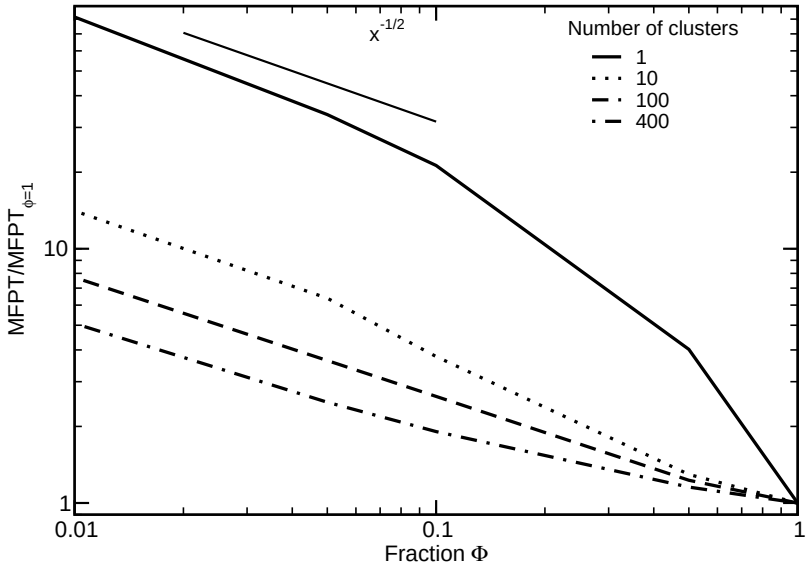

Figure 4.2: *The plot shows the mean first-passage time for one random walker starting at a random position to reach any of the clusters occupying a fraction of the whole membrane* Φ*. The results are normalised to the MFPT for the fully covered membrane* $\Phi = 1$*. Note that the scaling is proportional to* $\Phi^{-1/2}$ *for fractions smaller than* 0.1 *and for any number of clusters. This is in agreement with Schuss et al. (2007) result.*

and not affecting the first-passge time (Holcman and Schuss, 2008). Thus the phosphorylation time as a function of the number of HKs reads:

$$\tau_{\text{phos}}(N_{\text{HK}}) = \frac{\tau_{\text{phos},1}}{N_{\text{HK}}}. \tag{4.4}$$

This redistribution of smaller clusters is equivalent to that used by il Lim and Yin (2005), although with a subtle difference. In their approach the binding particles were external and entering the system, as opposed to our internal particles (TFs). Thus, it seems that the same principle applies, qualitatively, regardless of the location of the binding particles. We may attribute this effect to an increase in the radial symmetry given by the scattering of sensors on the membrane, reducing the average distance travelled by a transcription factor before finding a sensor.

Simulations offer us the possibility to calculate the noise of the first-passage time. If the TF were always to start at the centre of the cell and, assuming $\Phi = 1$, then its noise is $\eta^2 = 2/5$, which is possible to obtain analytically since the first passage to any point on the surface of a sphere is known (Nagar and Pradhan,

2003)*. For full coverage of the membrane and with a random initial position, as this is supposed to be the state of TFs in the cytosol, the noise becomes 1.8. A value larger than one indicates that the distribution has a power-law head, which originates due to the proximity of some particle to the membrane. This phenomenon is comparable to the distributions obtained of rebinding for two particles showed in section 2.4.1. As the covered fraction Φ decreases, the noise converges towards 1, which occurs for a value of $\Phi = 0.01$. This is an indication that the process becomes Poissonian, or, equivalently, that the distribution of the time to reach a receptor is exponential, thus, the system has no memory of the initial state, or equivalently a noise value of 1.

4.2.2 Few Transcriptions Suffice to Reduce Time Significantly

Another obvious strategy to reduce the phosphorylation time is to add more transcriptions factors. In Fig. 4.3 we show the time for the first phosphorylation to occur for a given sensor covering fraction $\Phi = 0.01$ for a number, N_{TF}, of transcription factors in the range 1–1000. We observe that the phosphorylation time is inversely proportional to the number of transcription factors (TFs). So, considering the phosphorylation time for one TF, as described in Eqn. 4.3, we may write:

$$\tau_{\mathrm{phos}}(N_{\mathrm{TF}}) = \frac{\tau_{\mathrm{phos},1}}{N_{\mathrm{TF}}}. \tag{4.5}$$

This proportionality is unaffected by the number of clusters and holds for a broad range of TFs that reach values considered high for two-component systems (see non-solid lines in Fig. 4.3). From this inverse proportionality, it is interesting to note that only the first few TFs have a significant contribution to the time reduction.

It is important to note that the noise (η^2) for a single particle is, due to the very small size of targets, nearly 1.

4.2.3 Time to Deliver the Signal to a Promoter

The second process in the signalling model is the transport of the signal from the sensor to the DNA binding site by the now-activated transcription factor. Unlike in the first process, described above, in this one the initial separation of the active transcription factor and the target is fixed and amounts to R. The average time for a single TF starting on the membrane to find a small spherical target in the centre of the cell is in first order approximation (Redner, 2001):

$$\tau_{\mathrm{act},1} = \frac{V}{4\pi D r_{\mathrm{promoter}}}, \tag{4.6}$$

where r is the radius of the small target. Note that this time only differs by a factor of $1/\pi$ compared to that of finding one small target on the membrane with initial random position of a TF, as shown in Eqn. 4.3.

*We should note that this result is also included in Redner (2001), albeit wrongly reported, it is corrected in the errata pages.

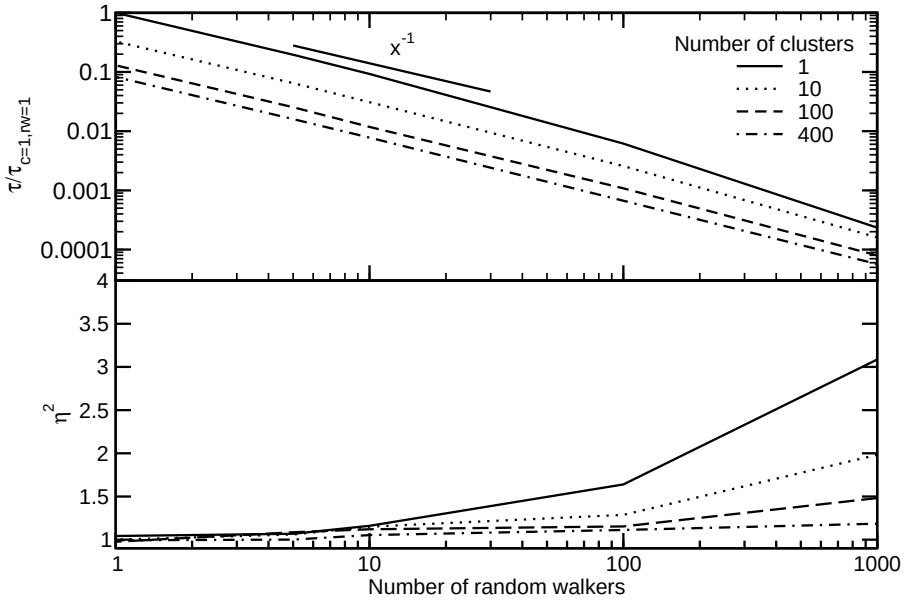

Figure 4.3: *The top plot shows that the mean-first passage time to a cluster of* $\Phi = 0.01$ *is inversely proportional the number of particles diffusing. Thus, the time is only significantly reduced by the first few particles. Noise* (η^2)*, however, increases with the number of particles due to increased probability of a particle being found close to a receptor, giving rise to a power-law distribution for short times.*

It is interesting to note that for more than one transcription factor the inverse scaling also applies, provided all of them have the same initial condition, that is, starting simultaneously. Consequently we may write again a simple scaling function for the activation time for a number N_{TF} of transcription factors:

$$\tau_{\mathrm{act}}(N_{\mathrm{TF}}) = \frac{\tau_{\mathrm{act},1}}{N_{\mathrm{TF}}}. \tag{4.7}$$

However, as we shall see later, this initial condition is not completely accurate, since the phosphorylation time for each TF occurs depends on the stochastic diffusion process. Therefore the initial conditions are not homogenous.

4.3 Analysis of the Response Time

4.3.1 Modelling the Signalling Response Time

From the previous results we conclude that a combination of scattered receptors (or sensors) and a relatively low number of transcription factors may reduce

response times significantly. To understand the influence of those structural parameters, it is instructive to study a simplistic yet intuitive equation-based model for the response time τ_{resp}. Considering the response time is the sum of the phosphorylation and activation steps, and by using the scaling laws derived in equations 4.4, 4.5 and 4.7, we arrive at (Dobrzynski et al., 2009):

$$\tau_{\text{resp}} = (\tau_{\text{phos}}/N_{HK} + \tau_{\text{act}})/N_{TF}. \tag{4.8}$$

By a simple inspection of the properties of this equation, we may state that the phosphorylation time τ_{phos} is smaller than τ_{act}, since it is reduced by both the number of HK and TF. Consequently, since τ_{phos} and τ_{act} are roughly equivalent, the second term dominates the response time. This equation stresses that the number of transcription factors is likely to be more influential than the number of histidine-kinases to reduce the response time. However, as we can readily see in Fig. 4.4, this model fails to accurately reproduce the simulated model, especially for a high number of transcription factors. Unlike in the phosphorylation process, where the initial condition was homogeneous (at $t = 0$ all TFs and HKs were in the same state), in the target search process they are not. In Eqn. 4.8 it is assumed that all phosphorylated TFs start their search for the promoter site simultaneously, which is not the case. As we discussed earlier, phosphorylation occurs at different times. Nevertheless, the simplicity of this equation shall prove insightful at a qualitative level in understanding phenomena shown later in this chapter.

In Dobrzynski et al. (2009) we present a more detailed model that accounts for the stochasticity of the two processes involved in the signal transport. The time taken by a single inactive transcription factor to reach the promoter is the sum of two stochastic processes:

$$f(\tau_{\text{resp}}, t) = X(\tau_{\text{phos}}, t) + Y(\tau_{\text{act}}, t),$$

where X and Y are the two diffusion driven processes for phosphorylation and promoter activations respectively. The sum of two stochastic processes is the convolution of their probability density functions (*pdf*), and, as argued before, both processes are well approximated to possess a negative exponential distribution ($\eta^2 = 1$) (Springer, 1979 ; Dobrzynski et al., 2009). Treating each individual particle trajectory allows us to have homogeneous initial conditions. The response time is obtained by taking the first-passage of a TF out of N (Yuste and Lindenberg, 1996), which results in:

$$\tau_{\text{resp}} = \left(\frac{1}{\tau_{\text{phos}}} - \frac{1}{\tau_{\text{act}}}\right)^{-N} \sum_{k=0}^{N} \binom{N}{k} \frac{\left(\frac{1}{\tau_{\text{phos}}}\right)^{N-k} \left(\frac{1}{-\tau_{\text{act}}}\right)^{-k}}{\left(\frac{k}{\tau_{\text{phos}}} + \frac{N-k}{\tau_{\text{act}}}\right)}. \tag{4.9}$$

This exact model shows an excellent agreement with simulations (exact model as a solid line and simulations as squares in Fig. 4.4). The influence that each parameter of this exact model has in the response time is more complex to understand than the more direct and clearer expression of the simple model of Eqn. 4.8.

4.3.2 Similar Numbers of Histidine-Kinases and Transcription Factors for an Optimal Response Time

As we argued in the introduction of this thesis, a cell's processes may be limited by resource availability and, as a consequence, natural selected processes that make the most efficient use of those resources may prove advantageous for the organism life. In a two-component signalling system (TCS) this concerns the use of structural factors such as number of molecules and spatial organisation. A simple hypothetical example illustrates the role of efficient allocation of resources in two-component signalling systems. Let us assume that a TCS may invest energy in producing a total of, for example, 50 molecules. Then the system faces a trade-off between investing energy to synthesize histidine-kinase sensors or to synthesize response regulators. Indeed, with this constraint the numbers of molecules correlate inversely to one another. For now, we do not consider the energy invested in forming clusters, nor do we take into account that either there is one single cluster or all HKs are scattered randomly. In this scenario we may ask ourselves what would be an optimal ratio, if any, between the number of HKs and TFs molecules, HK/TF, as to reduce the response time τ_{resp}.

Fig. 4.4 (panel A) shows the response time (τ_{resp}) curves for scattered HKs (solid and dashed lines), and for all HKs grouped into one cluster (dotted line). While the total number of HK and RR is restricted to 50 molecules, their ratio (HK/RR) varies from one TF and 49 HKs (49/1), to the opposite configuration, 49 TFs and one HK (1/49).

The common feature to both types of histidine-kinases spatial distribution is the presence of a minimum point, which is also not an extreme point. For the scattered cluster configuration this optimal point is achieved for a ratio HK/TF of approximately 20/30. Increasing the total number of particles does not affect significantly the location of this optimal point (results shown in Dobrzynski et al. (2009)). Note that qualitatively Eqn. 4.8 and 4.9 smaller response time values are biased towards a higher proportion of transcription factors since, as we have argued earlier, these have a double impact in reducing the response time.

When all sensors form a single cluster, besides the expected increase in response time, the optimal points shifts slightly towards a higher number of transcription factors. Moreover, when the size of the cluster is not too big, it notably increases the time for a low number of transcription factors.

Despite the presence of a minimum, the response time remains similar for a broader range of ratios. This flat region is less accentuated when sensors are clustered. Between 10 and 45 TFs τ_{resp} is below a factor $\tau_{\text{resp}} + 1$. This suggests that the response time may be rather robust against fluctuations in the ratio of molecules.

The distribution for the first-passage times of τ_{resp}, shown in Fig. 4.4 (panel B and C), show a Gamma-like peaked distribution which arises from the sum of the two exponentially distributed diffusion processes. Comparing the distribution taken at three distinctive points (1, 30, 49), we note that the noise (η^2) is smaller at the optimal point. These lower noise levels are not only due to the sum of the two exponential processes, but also to a higher number of molecules.

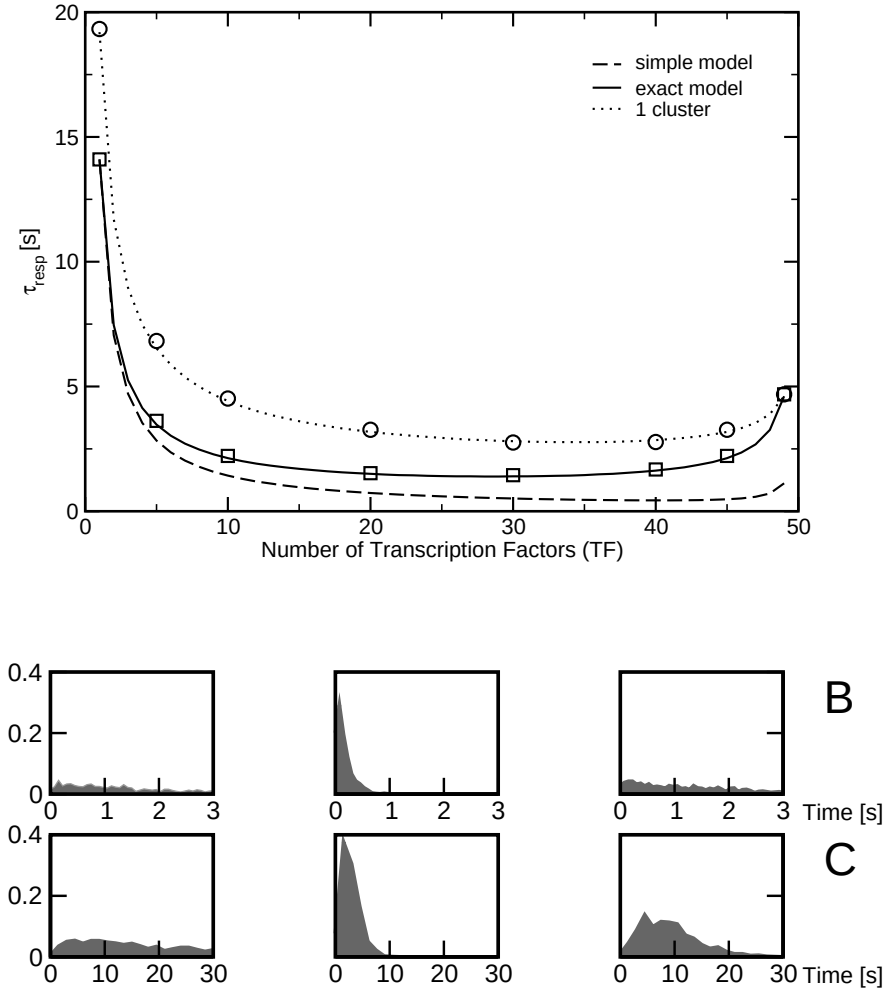

Figure 4.4: *Panel (A) response time optimality curve comparison between models and simulations. Note the bias of the curve and the presence of a minimum. Panel (B) distribution of the first-passage of the first (top) and second (bottom) process. Note the peaked-type of distribution, espicially noticeable in the second process. Simple model from Eqn. 4.8 (dashes) and the exact model from Eqn. 4.9 (solid line for scattered clusters and dots for one sensory cluster). $D = 5\,\mu m^2 s^{-1}$. Circles and squares are the results of numerical simulations.*

The smaller noise level indicates that the process is variable, in other words, that the delivery of the signal becomes more reliable.

4.3.3 Correlated Numbers of HK and TF

In the previous section we considered that the two-component system was constrained by the total number of molecules allowed. The presence of the non-extreme optimal minimum is a direct consequence of such constraints. However, often HKs and TFs are expressed from the same gene and their numbers are positively correlated. For example, the number of transcription factors is proportional by a factor α to the number of histidine-kinases. Assuming that we may only have an integer number of molecules and at least 1, the relation is:

$$TF = \max(\,\mathrm{round}(\alpha\,\mathrm{HK}), 1). \tag{4.10}$$

Using this new constraint, it is clear, as observed in Fig. 4.5, that the more molecules the shorter the response time is, shifting the non-extreme optimum point to an extreme. The jagged lines arise due to the discreteness of the system and because we allow a minimum of one particle (a continuous version would result in a smooth curve). Nonetheless, when the ratio of molecules is optimal, $HK/TF = \alpha = 2/3$ a faster response time is achieved with fewer particles than when the ratio is non-optimal (grey solid line). Dashed and dash-point correspond to a α factor of 5 and 25 respectively. Black lines are for the unfavourable ratio, $1/\alpha$, where the number of histidine-kinases is larger than the response regulators.

As with the scaling studied in section 4.2, the optimal response time is proportionally inverse to the total number of molecules. The continuous version of Eqn. 4.10, shows excellent inverse scaling. That is, the inverse of τ_{resp} is a straight line with a slope that depends on the correlation factor α. However, we did not investigate the relation between them. The discrete values of the number of particles, as established in Eqn. 4.10, cause the scaling for non-optimal ratios to start to apply for larger values of molecule numbers. It is important to note that, again, just a few molecules achieve the most significant reduction in response time. So, after a certain value, adding more molecules to reduce, say 0.01 s, would not pay off in terms of energy investment. This would only be justified by other factors, such as inefficiency in phosphorylation or auto-dephosphorylation, not dealt with in this reduced model.

To illustrate the effects of a non-optimal ratio, in Table 4.1 we compare the number of molecules needed to match the optimal time for the case 30 TFs and 20 HKs ($\tau_{\mathrm{resp}} = 1.39\,\mathrm{s}$) for a number of correlation factors α. This emphasises, once more, the advantages of an optimal HK/TF ratio and the bias towards higher numbers of TFs than of HKs to reduce the response time with a number of molecules as low as possible.

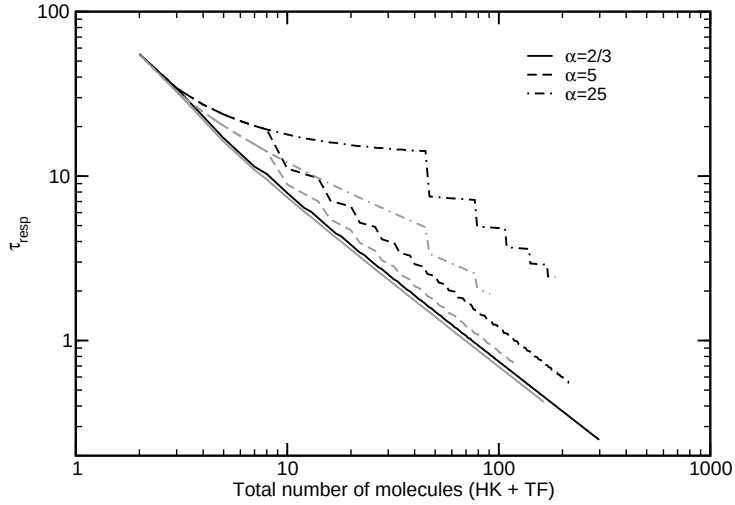

Figure 4.5: *Response time (τ_{resp}) for correlated numbers of particles and a number of α values shows an inverse proportionality to the number of molecules in the system. The jagged figures are consequences of the integer values used in calculating the proportion of molecules.*

Table 4.1: *Number of total molecules to obtain the same response time with 50 molecules with an optimal ration 2/3. Optimal response time 1.39 s.*

α	1	$\frac{2}{3}$	2	5	10
HK=(αTF)	51	50	51	63	80
TF=(αHK)	51	50	58	87	138

4.3.4 Assuming non-Diffusion-Limited Rates

In the basic model used previously, the rate of phosphorylation and activation are determined only by the diffusive transport of the transcription factors in the cytosol. As a result, the fixed relationship between those rates determines the shape of the response time curve, shown in Fig. 4.4. By relaxing the diffusion-limited behaviour of the system, the kinetic rates of phosphorylation and of activation may deviate from the diffusion-limited ratio causing the curve to change shape.

In Fig. 4.6 we show that increasing $\tau_{\rm phos}$, for instance, ten times (dotted line), gives rise to a more symmetric response curve while shifting the optimal point to a slightly lower number of TF (26/24). When increasing $\tau_{\rm act}$ by the same factor, but keeping $\tau_{\rm phos}$ in the original formulation, the curve becomes

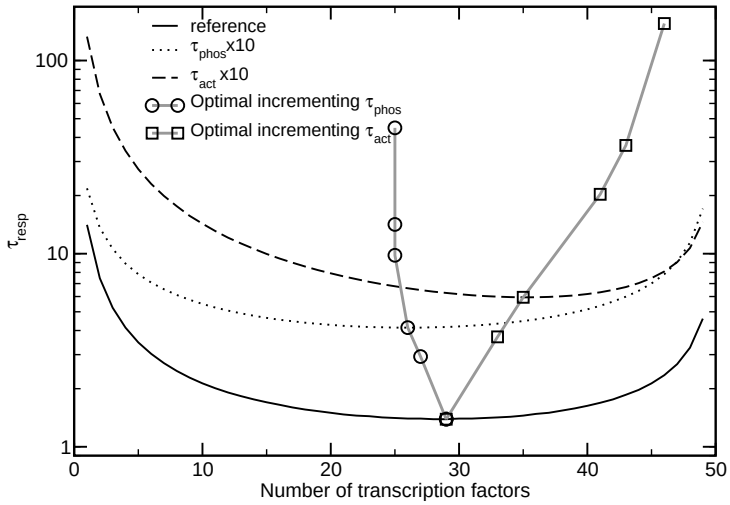

Figure 4.6: *Comparison of the signalling response time when the reaction time of the τ_{phos} is increased ten times (dotted line) and τ_{act} is increased 10 times (dashed line). The optimal point moves, and the highest impact of the response time τ_{resp} is due to an increment in τ_{act}. Note that, incrementing τ_{phos} makes the curve more symmetric.*

less symmetrical, especially on the left side, with a fourfold increase in response time, while on the other extreme the increase is approximately only twofold. We should note that the central flat area of the optimality curve is roughly conserved despite the bias. Now the optimal point is shifted towards the right, at 36/14. Increasing further the phosphorylation time, (5, 10, 50 and 100 times the diffusion limited rate, ascending order in Fig. 4.6, and circles) this process has less effect on the total response time than by an equivalent increase of the activation time (shown are squares). The observed convergence of the optimal response ratio towards 25/25 is well illustrated from the simple model shown in Eqn. 4.8, where the first term dominates the response time and its minimum occurs when $N_{HK} = N_{TF}$. However, increasing the activation time causes the optimal response time to shift towards 49, that is, using the totality of available transcription factors. However this convergence is much slower than the previous case, and $\tau_{\rm act}$ needs to be at approximately 3600 times larger than the original $\tau_{\rm phos}$. Once more, this limit is clearly deduced from the simple model, rather than the more accurate, and convoluted, one.

4.4 Discussion

We have studied the influences that structural factors, such as spatial organisation of sensors and ratio of molecules, have on the response time in two-component signalling systems. Importantly, we have relied on a generic model

as to shed light on the principal mechanism driving these systems. To this goal, rather than modelling a fully realistic system with detailed molecular-level descriptions, such as, reaction rates, dephosphorylation of active response regulators by scavenging molecules, autodephosphorylation and a short-lived activation of the sensors, we have opted, as a first step, to account only for the *essential* mechanisms involved in the signalling process. The first-passage paradigm allow us to analyse the system in the pure diffusion limited regime, which represents the lower limit of how quickly signalling may function (instant binding of molecules upon collision). Additionally, it is an explicit and insightful method to model first-time events, in this case, the first activation of a gene expression in response to an external signal.

We have supported the analytical development of section 4.3.1 by stochastic simulations of the first-passage processes, and more extensively analysed in the forthcoming publication Dobrzynski et al. (2009). The goal of these simulations was to support the applicability of the analytical model, since it relies on assumptions such as memory-less diffusion processes (Poissonian noise $\eta^2 \simeq 1$), and that histidine-kinases activities are independent of each other. It has been shown recently by Holcman and Schuss (2008) that if the membrane sensors were close enough to each other, the first passage of transcription factor would be affected non-linearly, breaking our assumption. Instead of using GMP or any other advanced tool, we made use of direct, explicit in space stochastic simulations. We needed to reduce uncertainties related to advanced features used in those methods in order to compare results against the analytical formulation. These simulations needed detailed spatial features, such as size of particles and spherical cap sensors, which are not fully supported directly by methods like GMP, MesoRD or Smoldyn. However, to study these systems with more accurate kinetics those methods are readily available.

In this chapter we addressed issues regarding two constraints imposed on the number of molecules used in two-component signalling. In the first, the total number of molecules is constrained and the numbers of histidine-kinases and transcription factors are inversely proportional, or inversely correlated. The second imposes no constraints on the total number, but correlates positively to the number of molecules. It is found that a ratio $HK/TF = 2/3$ gives, for small number of particles, nearly optimal signalling response, while keeping the number of molecules low. In Dobrzynski et al. (2009) we show that this optimal ratio is a very good approximation for a system with more than 20 particles, and for lower values the ratio is no larger than 1. With as few as 50 molecules, the response time is approximately 1/60th of the time required for expressing the first protein ($\approx 1\,\mathrm{min}$). Doubling the number of molecules, whilst keeping an optimal ratio, would only half that time (≈ 0.7), which in view of the time scale of other processes is probably not significant.

The assumption that the kinetics of the system are driven exclusively by diffusion of transcription factors helps to establish the lower bound of the response time in signalling. One may argue that when the kinetics are not diffusion limited, but reaction limited, the optimal point shifts. In Fig. 4.6) it is demonstrated that, in spite of the fact that the ratio for an optimal response time is dependent

on the ratio between the kinetic rates of phosphorylation and activation, the HK/TF optimal ratio is always ≤ 1. This could not have been observed a priori by inspecting the simple model in Eqn. 4.8, but by using the more accurate model in Eqn. 4.9, the simple model illustrates well the limits for the optimal ratio.

Increasing the activation time is not only affected by the affinity between the promoter and the transcription factor. Elf et al. (2007) reports, using recently developed single-molecule imaging in *in vivo* conditions, that transcription factors spend a reasonable time bound to non-specific DNA sites before reaching for the final target. Only then do they obtain an apparent diffusion coefficient, $D = 0.4\,\mu m^2 s^{-1}$, which is one order of magnitude smaller than of that of a transcription factor when it does not bind non-specifically to the DNA. Changing the diffusion coefficient in the model did not affect the optimal response time location (results not shown). Using normal diffusion for cytosolic proteins seems justified by the same experiments and in Elowitz et al. (1999). Anomalous diffusion has been mainly observed in eukaryotic cells where the cytoplasm is highly crowded (Kues et al., 2001 ; Wachsmuth et al., 2000), for membrane molecules (Nicolau et al., 2007), or for larger molecules such as mRNA (Golding and Cox, 2006). In bacteria, further experiments should shed more light onto the consequences of the dynamics of nonspecific 1D DNA sliding transcription factors in their diffusion and kinetics.

When considering ratios between histidine-kinases and transcription factors, we have to be aware of three modelling assumptions: that the reactions of the model are irreversible; that the external stimuli activate all HKs at once; and that the transport of the signal is perfect and the signal is not lost during it. We should refer to these numbers in the model as the number of molecules being activated, which may be lower than the number of molecules, both HKs and TFs, present in the system but remain left unused. This fact is important to account for when considering the degree of external signal present that activates the sensors. If the signal is abundant, it is likely that all sensors will become active, more or less, simultaneously. However, if the strength (or amount) of the signal is low, then a smaller fraction of sensors may be activated. This suggests that a two-component signalling system stimulated by a weak signal, may need to synthesize more HKs, as to avoid working with an unfavourable ratio, that is, in the right extreme of the response time optimality curve. Conversely, if during the transport of the signal, the transcription factor loses its phosphryl group, additional TFs would need to be present in the system to guarantee that a fraction of them actually deliver the signal.

Another alternative to detect weak signals, yet produce a quick and robust response, is the mechanism found in the chemotaxis cluster sensor (Shimizu et al., 2003 ; Sourjik, 2004). Sensors interact in a co-operative manner so the activation of one sensor may facilitate the autophosphorylation of a neighbouring sensor. This situation is roughly equivalent to the one cluster case shown in Fig. 4.4, where the cluster becomes fully activated, or not at all. Under this situation our model predicts, with similar arguments to those found in Berg and Purcell (1977) ; il Lim and Yin (2005), that the response time is slower.

Consequently, this clustered organisation requires more molecules to transport the signal. In the chemotaxis case, the receiver system is not a single promoter site, but rather the flagellar motors, which in contrast to the centre location of the promoter, these are located on the membrane and towards the opposite side of the cell of which the sensor is located (see section 3.3 for more details on signal transport in chemotaxis).

In the model shown in Eqn. 4.9 we have assumed rather simple diffusion and kinetics. Direct measurements of these response times and of the active or inactive state of various molecules under *in vivo* conditions may not yet be feasible. Rapid progress in single molecule microscopy technology, such as that used in Elf et al. (2007), may allow some properties of the response time to be measuredalbeit indirectly–at single molecule resolution (Yu et al., 2006 ; Marshall et al., 2008).

A few Words on Experimental Validation

Validating the theoretical results of the two search processes and obtaining their distributions remains a difficult task. The most straight forward approach would consist in timing the individual and fast events, for each binding event. Tracking a single transcription factor for its two binding events faces more difficulties due to the limitations of bleaching and over exposure of the fluorescent tags, together with the multiple particles present in the system. Recent single molecule microscopy techniques have been applied to the study of single protein synthesis (gene expression) (Cai et al., 2006), mRNA levels (Golding et al., 2005) or single transcription factor binding to the chromosome (Elf et al., 2007).

The optimality curve, as shown in Fig. 4.4, would require a control over the number of proteins present in the system. In a natural systems it is more easily achieved for the correlated case, as described in section 4.3.3, than for the fixed number of TFs and HKs as in section 4.3.2. Then, the timing can be in principle inferred indirectly from the gene expression synthesis.

4.5 Conclusions

From the modelling performed in this section we conclude that the signalling time, measured as the response time to an external signal, shows some degree of robustness against uncorrelated numbers of molecules and that significantly short times can be achieved by small numbers of molecules, both histidine-kinases (HK) and transcription factors (TF). The signs for robustness arise from the flatness of the optimality curve and the fact that only extreme ratios have a major impact on the response time. Under more realistic conditions, for example by taking into account reaction rates and diffusion coefficients, the response time is bound to increase. However, this increase can be, in principle, compensated by just adding few molecules to the system without recurring to a great expense in energy. It seems then that correlation is a central issue in the efficiency of these signalling systems: in prokaryotic organisms correlation is not uncommon. In fact, the operon organisation of most genes offers such a feature at a very

low energetic cost and guaranteeing that one single signal produces both (or more) molecules. Contrast such system with the single gene regulation typical of eukaryotic cells, in which two genes would require two correlated signals to produce a correlated product instead of just one signal required by prokaryote cells. In evolution terms, it is a very good choice to delegate to an architectural design the responsibility to correlate protein synthesis instead of using two systems to regulate one signal with consequent minimisation of energy utilisation and ultimately of space.

Besides the operon grouping, the optimality curve also suggests that a proper ordering of the genes might help to achieve a closer to optimal ratio between the number of histidine-kinases and transcription factors. Because the transcription and translation processes are imperfect, an upstream gene (closer to the transcription factor binding site) is more likely to be fully expressed than a downstream gene. Thus, gene order, together with operon organisation, might be a product of evolution in order to achieve efficient, robust and near optimal signalling. In certain occasions, like in chemotaxis, a mixture of genes found in the same operon and on different ones are used in order to bring the system to work efficiently. In that case, a special cluster and hypersensitive sensors are needed which would inevitably shift the optimal ratio between HKs and TFs. Regarding the expression processes, and more importantly how often these fail, we should note that they also play a role in regulating the expression levels at first instance without the intervention of additional regulating processes. We have already mentioned a number of other factors that affect signalling times, but in this chapter we have focused in the structural ones.

In the following chapter we explore the impact that gene order and expression processes have in the ratio of expressed (intermediate) products for prokaryotic two-component signalling systems.