Imperial College of London
Department of Mathematics

Beyond the two-state model of switching in biology and computation

Abhishek Deshpande

July 2018

A thesis is submitted in fulfilment of requirements for the degree of
Doctor of Philosophy in Mathematics of Imperial College London

Declaration of Originality

The content of this thesis is a manifestation of my work. Any ideas/quotations that refer to other people's work have been appropriately acknowledged using standard referencing methods.

Copyright Declaration

Acknowledgements

With great pleasure, I would like to thank my supervisors Tom, Manoj and Nick for their guidance and support. Throughout the PhD, Tom has been my primary point of contact. He has been a constant source of inspiration and has graciously given his time to answer my questions. I have benefited immensely from the profundity of his scientific thought process and physical insight. He has also been my role model in various capacities as a researcher, advisor and friend.

Manoj has been my mentor and guide right from my undergraduate days. He has been very patient in catering to my needs and has always motivated me to work harder. He has taught me to think independently and has encouraged me to ask the right questions in research. Many of the discussions that I had with him have helped me improve my all-round research skills.

Nick has been my pillar of support, both at a professional and personal level. He has always encouraged me to strive harder to achieve my goals. He has been very understanding and supportive of me throughout the course of my PhD. I am really fortunate to have him as my supervisor.

I am really grateful to Prof. Richard Craster, the head of the math department at Imperial College London for supporting my PhD with a Roth scholarship. I would also like to thank Greg Pavliotis, under whom I had the opportunity to work on a few projects. He has magnanimously provided his time to answer my questions. Many of the techniques that I learnt through interacting with him still come in handy.

I thank Anderson Santos for all the administrative support and making my stay at Imperial comfortable. I also thank my friends and group mates who have made my PhD a memorable experience. I thank Jure Vogrinc for generously giving his time and mathematical insights during many of our interactions. Rory Brittain deserves a special mention for helping me understand some of the concepts in statistical physics and also assisting me with his programming prowess in Mathematica. I am also thankful to our entire group: Juvid Aryaman, Hanne Hoitzing, Matt Garrod, Till Hoffmann, Thomas Mcgrath,

Ferdinando Insalata, Alex de Figueiredo, Sarab Sethi, Bertalan Gyenes, Benjamin Ingledow, Luke Heaton, Antonia Godoy-Lorite, Sam Greenbury, Carl Lubba, Javier Cabello Garcia, Ismael Mullor Ruiz, Jenny Poulton and Wooli Bae for providing all the support and encouragement during the course of my PhD.

I thank my teachers, professors and friends at my Alma mater IIIT-Hyderabad and Singhania school for their perpetual support and cooperation. I also thank my parents, family members and my wife Aishwarya for supporting me at an emotional level.

Abstract

The thesis presents various perspectives on physical and biological computation. Our fundamental object of study in both these contexts is the notion of switching/erasing a bit. In a physical context, a bit is represented by a particle in a double well, whose dynamics is governed by the Langevin equation. We define the notions of reliability and erasing time-scales in addition to the work required to erase a bit for a given family of control protocols. We call bits "optimal" if they meet the required reliability and erasing time requirements with minimal work cost. We find that optimal bits always saturate the erasing time requirement, but may not saturate the reliability time requirement. This allows us to eliminate several regions of parameter space as sub-optimal.

In a biological context, our bits are represented by substrates that are acted upon by catalytic enzymes. We define retroactivity as the back-signal propagated by the downstream system when connected to the upstream system. We analyse certain upstream systems that can help mitigate retroactivity. However, these systems require a substantial pool of resources and are therefore not optimal. As a consequence, we turn our attention to insulating networks called push-pull motifs. We find that high rates of energy consumption are not essential to alleviate retroactivity in push-pull motifs; all we need is to couple weakly to the upstream system. However, this approach is not resilient to cross-talk caused by leak reactions in the circuit.

Next, we consider a single enzyme-substrate reaction and analyse its mechanism. Our system has two intermediate states (enzyme-substrate complexes). Our main question is "How should we choose binding energies of the intermediates to minimize sequestration of substrates (retroactivity), whilst maintaining a minimum flux at steady-state?". Choosing very low binding energies increases retroactivity since the system spends a considerable proportion of time in the intermediate states. Choosing binding energies that are very high reduces retroactivity, but hinders the progress of the reaction. As a result, we find that the the optimal binding energies are both moderate, and indeed tuned with each other. In particular, their difference is related to the free energy difference between the products and reactants.

Contents

Chapter 1

Introduction

The thesis focuses on the idea of manipulating a bit - the fundamental unit of computation that possesses two states. Physical and biological systems however, are more complex than their computational abstractions - they go beyond the canonical two state model involved in manipulating a bit. This raises the question of how the bit moves between the two states. The impact of these intermediate configurations on the properties of bit-like systems is the unifying idea of the thesis. We will study the manifestation of these properties in two separate contexts: (i) Physical bits - where we study the erasure of a bit represented by a particle in a double well obeying Langevin dynamics (ii) Biological bits - where we analyse the switching of a substrate by an enzyme, which converts it from an inactive (active) state to an active (inactive) state. To motivate this discussion, we will briefly outline the relevant topics and put them into historical context.

1.1 Physics of computation

Maxwell's Demon, Smoluchowski's resolution and Szilárd's engine: The story began in 1867 when James Maxwell wrote a letter [1] to Peter Tait describing a thought experiment in which a "finite being" observes the speed of gas molecules in a box containing a trapdoor in the centre. The finite being allows the faster than average molecules and

slower than average molecules to settle on opposite halves of the box by opening the trap door suitably. As a consequence, temperature difference is generated without any apparent work input, seeming to contradict the second law of thermodynamics [2, 3]. Figure 1.1 illustrates the thought experiment outlined by Maxwell.

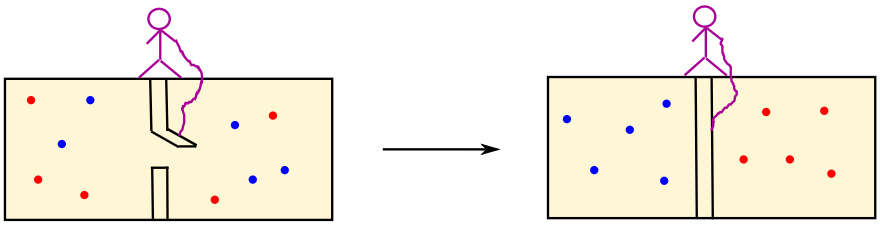


FIGURE 1.1: Maxwell's demon: A "finite being" can create temperature difference by opening the trapdoor suitably without any work input, to allow slow and fast molecules to settle in opposite halves of the box, seeming to contradict the second law of thermodynamics.

Later, Smoluchowski in 1914 suggested a modified experiment [4, 5], where he tried to represent the demon by a physical object, whose properties can be studied. More specifically, he imagined the trapdoor connected to a spring (which plays the role of the demon) such that the door allows only one way entry of molecules. Note that the gas molecules and spring are initially at different temperatures and hence out of equilibrium. One can exploit this temperature difference to power the gas molecules into one side of the trapdoor. Molecular collisions on the trapdoor gradually increase the energy of the spring. If the spring and gas molecules come into equilibrium, the door would start to open spontaneously allowing gas molecules to pass on both sides of the trapdoor. Thus, the second law of thermodynamics is never violated. Figure 1.2 shows Smoluchowski's construction of the spring-trapdoor mechanism.

In an attempt to explicitly bring measurement back into the story, Leo Szilárd in 1929 envisioned a thought experiment [6] using a model where calculations could be performed exactly. His model consists of a box containing a single gas molecule connected to a heat reservoir. The demon inserts a massless and frictionless partition in the centre of the box. If he measures which half of the box the particle is in, he can suitably attach a pulley that can lift a weight due to the collisions caused by the gas molecule on the

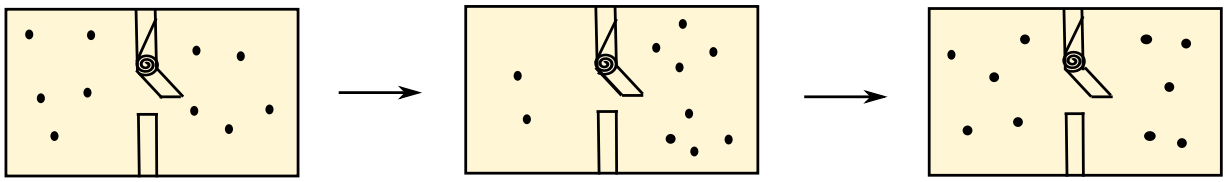


FIGURE 1.2: Smoluchowski's spring trapdoor: A trapdoor connected to a spring permits unidirectional flow of of gas molecules, building up pressure on one side of the box. When the spring heats up due to collisions by gas molecules and comes into equilibrium with the gas, the trapdoor starts allowing gas molecules to flow in either direction. The spring-trapdoor mechanism effectively acts as heat-engine, exploiting the initial temperature difference between the spring and gas molecules as a resource to channel gas molecules to one side of the box. The second law of thermodynamics is always preserved in this experiment.

partition. The entropy of the gas molecule is unchanged since it returns back to its initial state after the cycle. The entropy of the work reservoir also remains unchanged after this cycle. The heat drawn from the heat reservoir which is used to expand the gas, decreases the entropy of the heat reservoir. As a consequence, the total entropy of the universe (gas molecule + heat reservoir + work reservoir) appears to decrease, seeming to contradict the second law of thermodynamics. Szilárd posited that there has to be a cost for performing measurement. Though he did not define measurement explicitly, his observation played a crucial role in forming a mathematical bridge between information theory and thermodynamics. Figure 1.3 outlines the essential steps corresponding to the operation of the Szilárd engine.

Landauer's insight and Bennett's argument: Taking a cue from Szilárd's analysis, Brillouin [7] and Gabor [8] postulated that physical act of locating the position of the particle incurred an energy cost which saved the second law from being violated. Later Landauer [9] built a precise correspondence between logical and physical states of such systems. He discussed logically irreversible computations - where the outputs of the operations did not have a corresponding unique input. One such logically irreversible operation that he considered was the "erase" or RESET operation. It corresponds to taking a bit in a random state and setting it to a known state (usually "0"). Landauer's key insight was that the RESET operation requires a work input of at least $k_B T \log 2$. The RESET operation reduces the entropy of the bit by $k_B \log 2$. The input work is

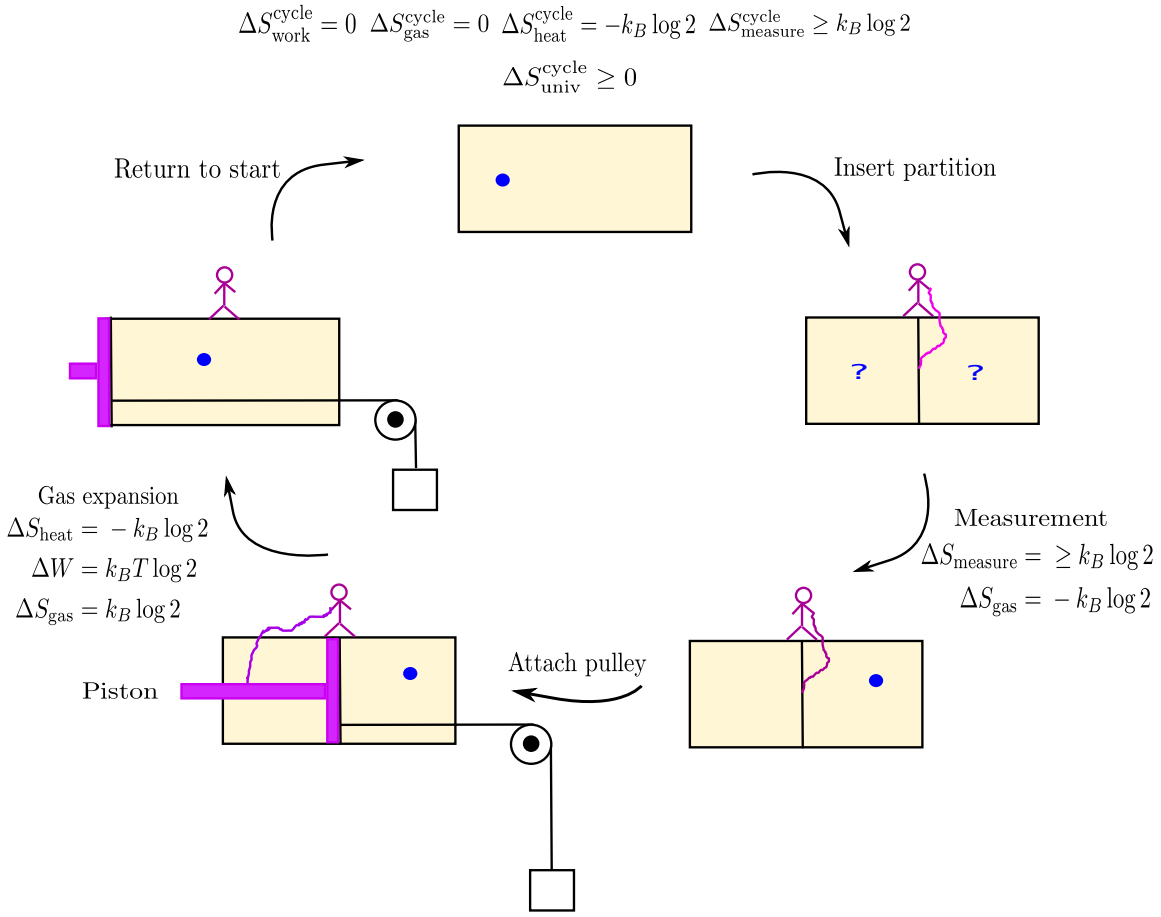


FIGURE 1.3: Szilárd engine: A cyclic protocol where the demon "measures" the position of the gas molecule, to extract work from it later. Szilárd postulated that the act of measurement requires at least $k_B \log 2$ entropy production, which compensates for the entropy reduction ($k_B \log 2$) that occurs when heat from the heat reservoir is taken to expand the gas.

converted into heat, which is deposited into the surroundings, increasing its entropy by at least $k_B \log 2$.

Bennett [10] proposed that Landauer's insight on physical consequences of bit manipulations was key to resolving the problem of Szilárd engine. Physically, he considered the same set-up as that of Szilárd, but with the gas molecule dimagentized and a ferromagnet in proximity that starts in a specific state, to be equivalent of the demon's memory. A measurement of the molecule would correspond to correlating the position of the diamagnetic gas molecule with the ferromagnetic material. He could show that measurement could be done in a logically reversible manner with no heat deposited into the environment. However, the measurement bit (ferromagnet) needed to be set to its initial state.

This act of resetting the demon's memory requires at least $k_B T \log 2$ of work input in accordance with Landauer's principle, which saves the second law from being violated. Later, both Lecerf [11] and Bennett [12] came up with a mechanism to run computers in a logically reversible fashion, with no heat exchange with the environment. This gave rise to the field of reversible computing. Subsequently, Fredkin and Toffoli [13, 14] came up with logic gates to perform arbitrary reversible operations.

Non-equilibrium thermodynamics: Modern resurgences in non-equilibrium thermodynamics inspired by advances in quantum mechanics and biological physics give a slightly different resolution to the Szilárd's problem in which the "resolution" of the paradox is not assigned to work done during a particular reset step; rather, it is seen that measurement creates a non-equilibrium state and this state must be paid for. They [15, 16] use the phrase "measurement" to include the following steps (i) correlating the demon's memory with data (the position of the gas particle) - the COPY step, (ii) de-correlating the memory with data - the DECORRELATE step and (iii) resetting the memory to a fixed value - the RESET step. For conceptual convenience, we can think of the demon's memory to be a hypothetical Szilárd engine different from the actual Szilárd engine. We will assume the demon's memory and the position of gas molecule to be binary random variables taking values 0 or 1.

Our universe shall consist of the demon's memory M, a gas molecule, also called data D (denoting the position of the gas molecule), a heat reservoir and a work reservoir. We will denote by $S(M)$, the entropy of the memory, $S(D)$, the entropy of the data, S_{heat}, the entropy of the heat reservoir, S_{work}, the entropy of the work reservoir and W, the available work in the work reservoir. Denoting by $S(M, D)$, the joint entropy of the combined system (demon's memory and the data), we note that $S(M, D) \leq S(M) + S(D)$ due to correlations between memory and data.

Assumption: Let's say that the demon's memory M starts with a fixed state "0".

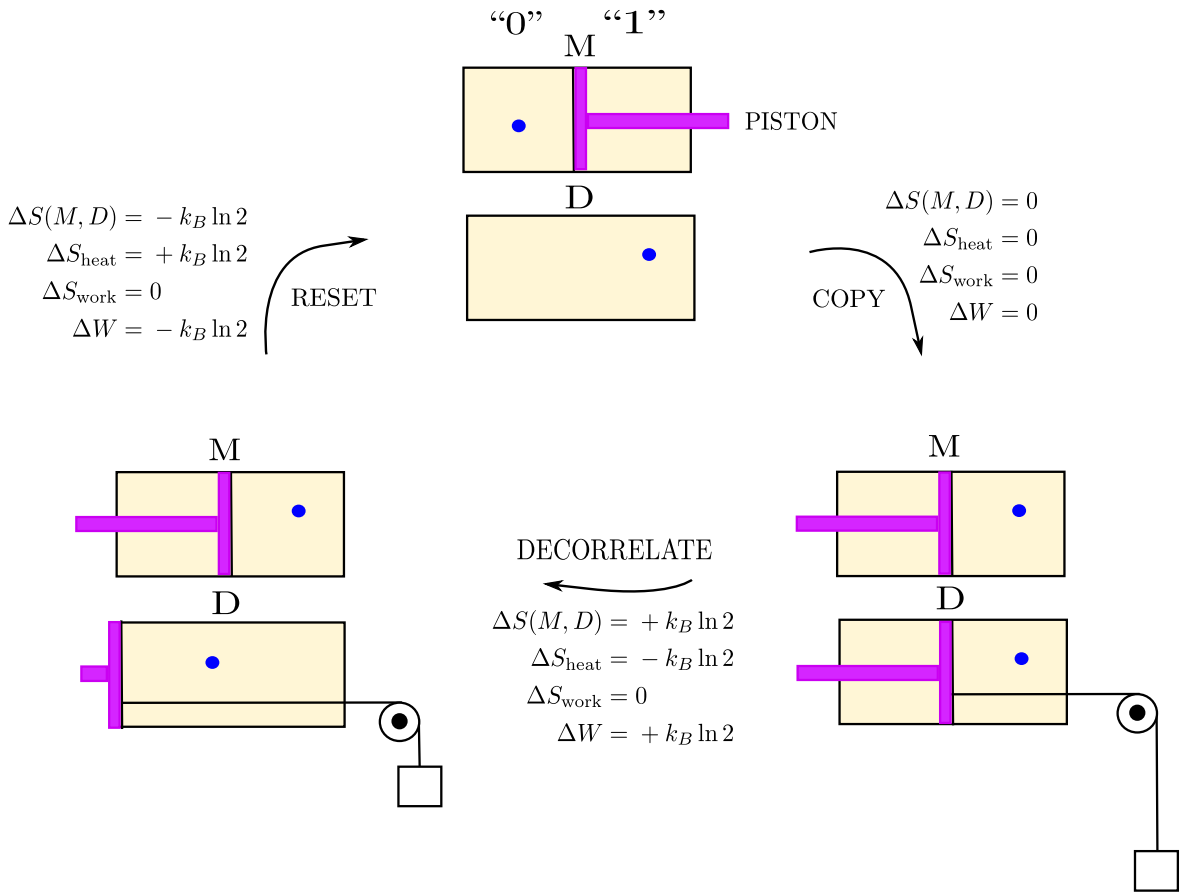


FIGURE 1.4: Non-equilibrium Szilárd engine: A schematic of the Szilárd engine, where the measurement procedure is split up into separate (i) COPY (ii) DECORRELATE and (iii) RESET steps. The universe consists of (a) The combined demon-memory system, (b) the gas molecule (c) a heat reservoir (d) a work reservoir. The schematic enlists the change in entropic costs for each component of the universe corresponding to various steps of the measurement protocol.

1. **Step 1 - COPY**: This corresponds to the following operation

$$M = 0 \text{ and } \mathrm{D} = \begin{cases} 0, & \text{w.p. } \frac{1}{2} \\ 1, & \text{w.p. } \frac{1}{2} \end{cases} \xrightarrow{\text{COPY}} \mathrm{M} = \mathrm{D} = \begin{cases} 0, & \text{w.p. } \frac{1}{2} \\ 1, & \text{w.p. } \frac{1}{2} \end{cases}$$

The entropy of the joint system before the copy step is $S(M,D) = S(M) + S(D) = 0 + k_B \log 2$ since the memory and data are uncorrelated. After the copy, we still possess one bit of information because the memory and data are now correlated, implying that $S(M,D) = k_B \log 2$. Therefore, $\Delta S(M,D) = 0$. There is no heat exchange with the heat reservoir, implying that $\Delta S_{\text{heat}} = 0$. The entropy of the work reservoir is unchanged, implying $\Delta S_{\text{work}} = 0$. Since there is no work input or extracted during this step, $\Delta W = 0$.

2. **Step 2 - DECORRELATE**: Knowing the data (i.e. which side the gas particle is on), one can connect a pulley to the partition suitably to lift a weight due to the expansion of the gas molecule. Symbolically, it corresponds to the following

$$M = D = \begin{cases} 0, & \text{w.p. } \frac{1}{2} \\ 1, & \text{w.p. } \frac{1}{2} \end{cases} \xrightarrow{DECORRELATE} M = \begin{cases} 0, & \text{w.p. } \frac{1}{2} \\ 1, & \text{w.p. } \frac{1}{2} \end{cases} \text{ and D} = \begin{cases} 0, & \text{w.p. } \frac{1}{2} \\ 1, & \text{w.p. } \frac{1}{2} \end{cases}$$

Just before the gas expansion, the data and memory are correlated and we have one bit of information implying that $S(M, D) = k_B \log 2$. After decorrelating the memory from data, we have two independent random bits giving a joint entropy $S(M, D) = S(M) + S(D) = 2k_B \log 2$. Therefore, $\Delta S(M, D) = k_B \log 2$. Heat is drawn from the heat reservoir to expand the gas, implying that $\Delta S_{\text{heat}} = -k_B \log 2$. The entropy of the work reservoir is unchanged, implying $\Delta S_{\text{work}} = 0$. The expansion of the gas is used to useful work; therefore $\Delta W = +k_B T \log 2$.

3. **Step 3 - RESET**: In this step, the demon's memory is set to state "0", by compressing the gas molecule of demon's Szilárd engine to half its volume. Symbolically,

$$M = \begin{cases} 0, & \text{w.p. } \frac{1}{2} \\ 1, & \text{w.p. } \frac{1}{2} \end{cases} \text{ and D} = \begin{cases} 0, & \text{w.p. } \frac{1}{2} \\ 1, & \text{w.p. } \frac{1}{2} \end{cases} \xrightarrow{\text{RESET}} \text{M} = 0 \text{ and D} = \begin{cases} 0, & \text{w.p. } \frac{1}{2} \\ 1, & \text{w.p. } \frac{1}{2} \end{cases}$$

Neglecting the internal details of how exactly one implements each step of the measurement protocol, Figure 1.4 gives a schematic of the key steps and their corresponding changes in entropic costs (given that the memory of the demon starts in state "0"). From the above analysis, we have the before resetting, $S(M, D) = 2k_B \log 2$. After resetting the memory bit to "0", we have essentially one bit of information giving a joint entropy $S(M, D) = k_B \log 2$. Therefore, the change in the entropy of the joint system $\Delta S(M, D) = -k_B \log 2$. The compression of gas in the demon's Szilárd engine puts $k_B \log 2$ heat into the heat reservoir. Hence, $\Delta S_{\text{heat}} = +k_B \log 2$. The entropy of the work reservoir is unchanged, implying

$\Delta S_{\text{work}} = 0$. The compression of gas molecule in the demon's Szilard engine requires $k_B \log 2$ of work input, implying that $\Delta W = -k_B T \log 2$. Table 1.1 summarises our observations.

	$\Delta S(M, D)$	ΔS_{heat}	ΔS_{work}	ΔW
COPY	0	0	0	0
DECORRELATE	$+k_B \ln 2$	$-k_B \ln 2$	0	$+k_B T \ln 2$
RESET	$-k_B \ln 2$	$+k_B \ln 2$	0	$-k_B T \ln 2$

Table 1.1: Changes in thermodynamic quantities for various steps in the measurement protocol when the bit starts in a fixed state "0".

We could also have started our measurement cycle with the bit in a purely random state instead of the fixed state "0". In this case, there is no need to reset the bit after decorrelating it from the data. The costs associated with the RESET step in the previous case become those associated to the COPY step in the present case. More precisely, we get Table 1.2.

	$\Delta S(M, D)$	ΔS_{heat}	ΔS_{work}	ΔW
COPY	$-k_B \ln 2$	$+k_B \ln 2$	0	$-k_B T \ln 2$
DECORRELATE	$+k_B \ln 2$	$-k_B \ln 2$	0	$+k_B T \ln 2$

Table 1.2: Changes in thermodynamic quantities for various steps in the measurement protocol when the bit starts in a random state.

The key observation is that correlations between non-interacting physical systems put them out of equilibrium - a state for which one needs to pay a cost. This behaviour is very similar to what catalysts do in a biological context. We will come back to this in the next section on biological computation, where we discuss catalytic systems in detail.

To summarise the history of the field, it is fair to say that Szilárd (rightly) posited (using the example of a single particle engine) that the act of "measurement" requires work input, which prevents the second law from being violated. Brillouin was perhaps confused, attributing the cost to the physical act of knowing the position of the particle. Landauer

postulated that performing logically irreversible computations necessarily increases entropy of environment. Bennett used Landauer's argument to say that the demon does not have to pay work to measure, but does have to pay work to reset. The fundamental point made by modern synthesis based on non-equilibrium thermodynamics is that the details of when work is put in– during measurement or reset is not important. The measured state is out of equilibrium and can be exploited to get work, but also must be paid for (either by external work or some other resource).

Reliability of information: In a slightly different context, the notion of reliability of information was introduced by von Neumann in his lectures [17] at the University of Chicago. He observed that there was a huge difference between the cost of neuron spiking and the thermodynamic estimate of $k_B T \log 2$, which is the minimal cost for typical manipulations of a bit. One possible explanation was that biological computations are required to be reliable i.e. be able store information for long periods of time, which results in extremely high energy costs. This view was further perpetuated by Swanson [18] who used ferromagnetic substances as memory devices to store information "reliably". Using the entire ferromagnetic material to make a bit was inefficacious since it did not result in many bits; while dividing the material into smaller pieces left it vulnerable to effects like calorific agitation and quantum tunnelling. Thus, he concluded that there was a size-efficiency trade-off for building bits – the optimal size for building efficient bits was neither too big nor too small; but in fact moderate.

Cost/speed trade-off to erasing a bit: It is important to note that the Landauer bound of $k_B T \log 2$ for erasing a bit holds in the quasi-static limit i.e. when the procedure is carried out infinitely slowly. However, in reality computations are performed fast and require substantial energy consumption – much more than the $k_B T \log 2$ estimate. Thus the time spent in the intermediate states is of paramount importance and we are coerced into looking *beyond the two state model of erasing a bit*, consonant with the title of the thesis. The idea of finite-time thermodynamics was kick-started in 1975 by Andresen,

Salamon and Berry, who published a series of papers [19, 20, 21, 22], analysing classical thermodynamic models like the heat-engine [23], and step-Carnot cycle [24] for their efficiency. Subsequently, this has evolved into the field of stochastic thermodynamics. The basic idea of stochastic thermodynamics starts with the following: Let $U(\lambda, x)$ denote the internal energy of the system as a function of coordinate x and a control parameter λ which is amenable to external manipulation. Then

$$dU = \frac{\partial U}{\partial \lambda} d\lambda + \frac{\partial U}{\partial x} dx \tag{1.1}$$

Sekimoto [25] recognises $đW = \frac{\partial U}{\partial \lambda} d\lambda$ as the differential work done in manipulating the potential of small systems at a rate faster than their characteristic time-scale and $đQ = \frac{\partial U}{\partial x} dx$ to be the heat dissipated into the environment. Formulating the first-law in this way allows us to define relevant thermodynamic quantities like work, heat, entropy on the level of fluctuating trajectories. Several fundamental results like the famous Jarzynski's equality [26] and Crooks fluctuation theorem [27] (which relate ensemble averages of functions of thermodynamic variables over trajectories) – have been established using tools from the thermodynamics of small systems. We refer the reader to a comprehensive review of this field by Udo Seifert [28]. Using these techniques, Sekimoto and Saga [29] have given a lower bound for the excess heat dissipated in a given time window for erasing a bit represented by a particle obeying Langevin dynamics. In particular, they obtain $Q_{\text{irr}} \Delta t \geq k_B T A_{\min}$, where Q_{irr} is the excess heat dissipated, Δt is the time window of operation, and $A_{\min}$ is the minimum value of action required for the protocol. Schmiedl and Seifert [30] have calculated optimal protocols for taking a system obeying overdamped Langevin dynamics from one equilibrium state to another. Using techniques from geometric optimization [31], Zulkowski, Sivak and Deweese [32] have extended this result for a system moving between two non-equilibrium steady-states. This has resulted in work by Aurell et al. [33] which gives an entropic lower bound for erasure in a finite time interval by using techniques from Monge-Kantorovich optimal mass transport theory. More recently, Zulkowski and DeWeese [34] have looked at erasure of

a bit represented by a particle in a square well potential obeying overdamped Langevin dynamics. Using the chemical master equation to model probability distributions, they obtain $Q/k_BT \approx \frac{\Delta S}{k_B} + \frac{4K}{\Delta t}$, where Q is the excess heat dissipated, ΔS is the difference in Shannon entropies of the initial and final states, K is the Hellinger distance between the initial and final states and Δt is the duration of the protocol. The notion of accuracy has also been explored by Giovanni et al. [35] where they study energy/speed/accuracy trade-offs for erasure in the setting of quantum dots. The common point that all these analyses drive home is that faster (and more accurate) computations require more work input, seeming to give an alternative resolution to von Neumann's question raised before.

It is interesting to note that bit erasure has been studied using arbitrary complex control strategies (i.e. arbitrary $\frac{\partial U}{\partial \lambda}$). This begs the question of whether these controls are physically reasonable? Taking this into account, we study the relationship between cost, speed and reliability for erasing a bit represented by a particle in a double well obeying Langevin dynamics. We pick a simple family of controls for our analysis, but many of our results generalise to other families. This forms the core idea for Chapter 2 of the thesis.

1.2 Biological computation

Natural computation: Examples of computations performed by natural systems include, but are not limited to DNA replication [36] – where information encoded in a single DNA molecule is passed on to two similar DNA strands, transcription of DNA to RNA [37] – a process where RNA polymerase catalyses the conversion of DNA to its complementary RNA strand, translation of RNA to proteins [38] and modulation of proteins by transcription factors, which bind to DNA and turn on/off certain gene sequences. A key property of such computations is that they are potentially autonomous i.e. (no external manipulation of potentials etc.). We are oblivious to the internal details of how biological systems accomplish certain tasks. This essentially gets around the question of: "How do

we implement our protocols?" linking back to the ideas mentioned in the previous section.

Yet another example of natural computation occurs in the context of signal amplification. The key idea here is that certain special proteins called enzymes trigger activation of a large number of molecules in the cascade which results in the output being amplified. This has been explored quite a while ago in two biological contexts [39] (i) blood clotting and (ii) vision. In the case of blood clotting [40], a special type of co-enzyme called Hageman factor trips the activation of a large numbers of pro-enzymes eventually resulting in the conversion of fibrinogen into fibrin which causes clotting. In case of vision [41], a few molecules of rhodopsin activate a large number of enzymes in their pathway resulting in an amplified signal. Cascades of enzymes have been known to produce switch-like behaviour [42]. One such example is the well known MAP cascade that consists of three components: (i) MAPKKK (ii) MAPKK and (iii) MAPK. This cascade is known to show ultra-sensitive behaviour. Further, it has been shown that the number of cascades has a multiplicative effect on the sensitivity [43].

Signal amplification has also been found to be useful in chemotaxis [44, 45, 46] - which is used by cells to sense concentrations of chemicals. It has been postulated that a large number of receptors "grouped" together can co-operatively amplify signals [47]. Therefore, bacteria like *E. coli* can detect even small changes in concentration gradients [48, 49]. The mechanism of chemotaxis in *E. coli* involves inter-conversion of the receptor-kinase system between its active and inactive forms [50] reminiscent of switching a bit. It also involves feedback mechanisms to effectively calculate a gradient [50] using a technique called "exact adaptation".

Engineered computation: Inspired by natural systems, Len Adleman [51] in 1994 first demonstrated how to solve the Hamiltonian path problem by using DNA strands. The idea was to exploit the massive parallelism generated by mixing DNA strands to simultaneously generate solutions to the Hamiltonian path problem (known to be NP-complete) in linear time. This kick-started the field of DNA-computing. Since then the field has

evolved to different branches like DNA self-assembly, a program initiated by Nadrian Seeman [52] where small DNA strands assemble themselves to form intricate shapes and structures. The field of DNA origami (a special case of self-assembly) has been pioneered by Paul Rothemund [53], where a very long strand of DNA folds to form complex architectures. Several interesting computations have been made using the tool of DNA strand displacement reactions [54, 55]. Notably, Qian and Winfree [56] have used DNA strand displacement reactions to compute the square root a 4 bit number – a first of a kind computation. It is worth mentioning that around 130 DNA strands were used in the circuit for this protocol.

Chemical Reaction Networks: The theory of chemical reaction networks (CRN) initiated by Horn and Jackson [57] provides a solid mathematical framework for analysing complex biological networks. The key ingredients of CRNs consist of a set of species, a set of complexes and a set of reactions from which the stoichiometric matrix is formed. Various questions like the persistence of network trajectories [58, 59, 60, 61, 62], stability of fixed points [63, 64, 65, 66] and properties of oscillations [67, 68, 69] can be answered using the powerful machinery of CRNs. The precise correspondence between chemical reaction networks and their implementation using DNA-strand displacement reactions was given by Winfree et al. [70]. Since then, a large community of researchers have been trying to come up with chemical reaction networks that can perform certain specific computations. David Soloveichik et al. [71] have designed a chemical reaction network that deterministically computes a semi-linear function. Further, they show that this computation is poly logarithmic in the size of the input CRN. This result has been generalised in a later paper [72] to the stochastic setting, where the computation happens with a non-zero probability of error. Luca Cardelli et al. [73] have designed a chemical reaction network such that the steady-state of the output molecule is a draw from an arbitrary probability distribution with finite support in the natural lattice. Separate initiations of their reaction network gives independent draws, but a single implementation only samples a single point. In a different paper, Cardelli and Csiksz-Nagy [74] have demonstrated that the

cell-cycle switch simulates approximate majority - a distributed computing algorithm. Recently, Gopalkrishnan [75] has come up with a scheme for designing a reaction network that calculates the maximum likelihood estimator for log-linear models. In a sequel, Gopalkrishnan et al. [76] build a stochastic reaction network scheme to perform information projection. It is worth noting that fundamental physical trade-offs (e.g. microscopic reversibility of physical reactions) have generally not been taken into consideration while analysing such networks. A recent review by Ouldridge [77] highlights the relevance of microscopic reversibility for the parity computation algorithm [78]. Ignoring microscopic reversible reactions effectively implies that one would need an infinite free energy drive for the computation to work.

Catalysis and Enzymes: Many of the natural and synthetic examples discussed in the previous section like DNA self-assembly, origami, allosteric activation of enzymes and their natural analogues involve relaxation to an equilibrium state specified by selective binding of molecules. Dominique Chu has studied generic properties of such systems. He defines entropy driven computation [79] as a discrete state continuous time Markov chain relaxing to equilibrium and shows that there is a trade-off between energy cost, accuracy and the system size. An alternative paradigm, particularly, common in natural systems discussed before like MAPK cascades, DNA self-replication and transcription is catalytic binding in which interaction is temporary, but the effect long-lived – reminiscent to the notion of non-interacting measurement of a particle in Szilárd's engine. Naïvely, a catalyst is a substance that alters the rate of a reaction without being consumed by it. Further, it also does not alter the equilibrium of the reaction system. However, catalytic systems use fuel molecules like ATP to pump the system [77] into a non-equilibrium steady-state. Ouldridge and ten Wolde [80] have established that enzymatic actions of certain catalytic systems called the "push-pull motifs" can be directly mapped to the notion of bit switching. They show that energetic costs derived for abstract bits are actually relevant bounds for biochemical circuits. Though they analyse the thermodynamics in detail, they still assume the two-state model for switching a bit. Catalysts also include enzymes that

are designed to act on specific substrates in various biochemical pathways. In particular, enzymes have been studied in the context of gene-metabolic feedback circuits by Oyarzún and Stan et al. [81, 82, 83]. In what follows in Chapters 3 and 4 of the thesis, we will study enzyme-substrate interactions in the context of retroactivity - an effect we introduce in the subsequent section.

Modularity and Retroactivity: The ability of individual components to retain their functionality upon interconnection is called modularity. It has been the object of study in biological systems; more specifically in biological evolution [84, 85], and the evolution of modularity itself [86]. The origin of modularity has been studied in detail by Wagner et al. [87] contemplating between whether it is a fallout of natural selection or mutation. It has also seen observed to contribute to system robustness [88, 89]. The review article by Pantoja et al. [90] gives a vivid history of the study of modularity in biological systems.

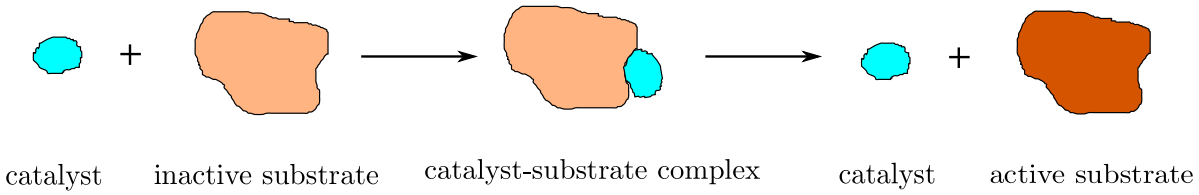


FIGURE 1.5: Mechanism of catalysis: A rough schematic of catalyst-substrate reaction. There is a non-zero time for which the catalyst is bound to the substrate and this catalyst-substrate intermediate gives rise to retroactive effects for the motif.

However, real systems do not exhibit perfect modularity. The breakdown of modularity leads to retroactivity [91] - an effect first observed in electrical systems. It essentially refers to the back-signal experienced by the upstream system when it passes on a signal to a downstream system. Catalysis is one way biological systems can try to reduce this effect since catalysts do not need to remain bound permanently with the substrate. However, even in catalysis there is a non-zero time for which the components are bound to each other. These intermediate states are analogous to the intermediate states involved in bit switching discussed in the previous section, and impart retroactivity to motifs. Figure 1.5 gives a rough sketch of the mechanism of catalytic interactions. We study the effect of retroactivity in two settings.

In the first one, given in Chapter 3, we analyse the "push-pull" motif, a network first studied in detail by Del Vecchio et al. [92]. The system essentially consists of two antagonist enzymes that act in opposite directions converting a substrate from one state to another. This is also the model studied by Ouldridge and ten Wolde [80], where they map it to a copying process using an ingenious transform. Maintaining the push-pull motif in a non-equilibrium steady-state requires energy consumption in the form of fuel molecules - an idea that has a parallel to the notion of energy consumption required to switch bits fast and reliably (as discussed in the previous section). Barton and Sontag [93] claim that it takes substantial energy consumption to reduce the retroactivity of this motif. One of our main contributions in this chapter is to refute the claim made by Barton and Sontag - we show that high rates of energy consumption are not required to attenuate retroactivity.

In the second setting given in Chapter 4, we study the mechanism of enzyme-substrate catalysis – a phenomenon ubiquitous in biological systems. Instead of considering the entire push-pull motif, we analyse the action of a single catalytic enzyme that converts a substrate from its inactive form to its active form. Such systems have been studied before under certain assumptions and are popularly known as the Michaelis-Menten and Briggs-Haldane dynamics. We generalise the assumptions made in these models to construct enzyme-substrate motifs that minimize retroactivity at fixed rate of reaction. This forms the basis of Chapter 4 in the thesis.

A more detailed breakdown of the contributions in each chapter is given below.

1. Chapter 2 introduces the idea of a bit as a particle in a double well obeying Langevin dynamics. We make precise what we mean by switching/erasing a bit. In addition, we introduce the notion of reliability (given by the reliability time-scale), speed of operation (characterised by the erasing time-scale) and the cost associated with erasing a bit. We also discuss the non-monotonic dependence of reliability and erasing time-scales as a function of the friction coefficient. We define optimal bits -

bits that meet the desired reliability and erasing time requirements with the lowest work cost within some family of controls. We analyse the geometry generated by such bits and rule out certain regions of parameter space as sub-optimal. The content of this chapter was published in [94].

2. Chapter 3 introduces the notion of retroactivity in the context of biological circuits. Our primary network of study is the push-pull motif. Maintaining this motif in a non-equilibrium steady-state requires energy consumption in the form of fuel molecules. Our main result is that it does not require high energy to reduce the retroactivity of this motif. We illustrate this effect by coupling the upstream enzyme weakly to the push-pull motif. However, making the coupling too weak leaves the motif exposed to cross-talk caused by leak reactions in the circuit. This work was recently published in [95].

3. Chapter 4 deals with designing efficient enzyme-substrate networks. More specifically, we ask: how to choose the binding energies of the intermediate enzyme-substrate complexes so that it minimizes the retroactivity of the motif at fixed overall rate of the reaction. This manuscript is in preparation.

4. Conclusion: In this section, we conclude by summarising our observations about switching in both physical and biological contexts.

5. The Appendix contains detailed calculations, proofs and other details related to various chapters in the thesis.

Chapter 2

Designing the Optimal Bit: Balancing Energetic Cost, Speed and Reliability

Abstract

We consider the challenge of operating a reliable bit that can be rapidly erased. We find that both erasing and reliability times are non-monotonic in the underlying friction, leading to a trade-off between erasing speed and bit reliability. Fast erasure is possible at the expense of low reliability at moderate friction, and high reliability comes at the expense of slow erasure in the underdamped and overdamped limits. Within a given class of bit parameters and control strategies, we define "optimal" designs of bits that meet the desired reliability and erasing time requirements with the lowest operational work cost. We find that optimal designs always saturate the bound on the erasing time requirement, but can exceed the required reliability time if critically damped. The non-trivial geometry of the reliability and erasing time-scales allows us to exclude large regions of parameter space as suboptimal. We find that optimal designs are either critically damped or close to critical damping under the erasing procedure.

2.1 Introduction

As discussed in the introduction, minimal costs of bit manipulation is a long standing topic of investigation, with deep connections to basic thermodynamics. In particular, certain information processing operations such as erasing a bit, or copying the state of one bit into another previously randomised bit have fundamental lower bounds on work input [6, 9, 10, 96, 97]. Practical devices, however, do not approach these bounds [98, 99]

and insights gained from thinking about the lower bound have not yet translated into more energy-efficient technology. A partial explanation is that man-made devices and biological cells need to operate on fast time-scales and hence cannot involve the quasi-static manipulations necessary to reach lower bounds [34, 100]. As remarked earlier in the introduction, an alternative suggestion from von Neumann is that the need to store information for long periods of time (**reliability**) leads to high-cost architectures [17]. We explore the interplay between reliability, speed and the energetic cost of bit operation. Equilibrium thermodynamic bounds such as the Landauer limit cannot account for these inherently kinetic phenomena. This general question of how to design fast, cheap and reliable bits has obvious technological relevance to the optimal design of low power computational devices [101, 102, 103] Additionally, since the discovery of the structure of DNA and the central dogma of molecular biology, it has become well accepted that information processing is at the heart of many natural phenomena. Many authors have explored information processing in biological systems, both to understand natural examples and design synthetic analogs [10, 104, 105, 106, 100, 77, 107, 108]. The question of the interplay between reliability, speed and cost are also relevant here, although under-explored. In this chapter, we explore the challenge of building fast, cheap and reliable bits, and provide a framework for its analysis in terms of reliability and erasure time-scales. We also take the first steps towards exploring the physics of the optimal design problem by considering a simple model: a particle in a 1-D potential, which is a quartic double-well potential in the device's "resting" state. We require that the bit be reliable, so that a particle equilibrated in either well stays in that well for a specified long time on average. Simultaneously, we require the implementation of an "erase" or "reset" operation using an external control, so that erasure is completed within a specified short amount of time. Our principal question is to find values for the design parameters which consist of the height of the double well, the friction coefficient, and the control parameters to guarantee these requirements without expending more energy than required. Our main contribution is an exploration of this design space, which demonstrates the previ-

ously under-appreciated role of friction. In particular, we identify a "Goldilocks zone" where the friction coefficient takes moderate values. This is somewhat counter-intuitive because historically friction has been viewed as a nuisance to computing, to be sent as low as possible [109, 110, 111, 112].

In Section 2.2, we describe the model which will provide intuition for our work. We formalize the time-scale over which the bit stores information through the notion of **reliability time**. In Section 2.2.3, we describe one simple family of control protocols for resetting a bit. We calculate the work done in erasing a bit for this form of control. We will use this particular control protocol to illustrate our subsequent ideas. In Section 2.2.3, we introduce the notion of **erasing time**. In Section 2.3, we consolidate from the literature the analytical forms and approximations for our two time-scales of interest, and confirm them with numerical simulations. We find that both the reliability and erasing time-scales are non-monotonic, roughly U-shaped functions of the friction coefficient. It follows that high reliability is obtained by setting the friction to a low or high value, whereas a low erasing time is favoured by an intermediate value of friction, implying a conflict between the two time scales for a given class of protocols. In Section 2.4, we investigate how this conflict feeds into the geometry of **optimal bits**: bits that fulfil the desired reliability and erasing time requirements with the minimum energy cost. We find and partially characterize a "Goldilocks zone" in design space where optimal bits reside. In Section 5, we discuss the robustness of our results when more freedom is allowed in the choice of design parameters, and the control protocol.

2.2 The Double-Well Bit

We will represent a device to store one bit of information by a particle in a symmetric bistable potential $U_{A,B}(x) = A\left(\frac{x^2}{B^2} - 1\right)^2$, where A is the height of the well and $\pm B$ are the coordinates of the minima of the right and left wells. We will refer to the device as a

whole by "a bit". The device reports "0" when the particle is in the left well, i.e., $x < 0$ and reports "1" otherwise (Figure 2.1a). The dynamics of the particle is described by the Langevin equation.

$$\begin{aligned} m\,dx &= p\,dt \\ dp &= -\gamma p\,dt - \partial_x U_{A,B}(x)\,dt + \sqrt{2m\gamma k_B T}\,dW \end{aligned} \tag{2.1}$$

where m is the mass of the particle, x is position, p is momentum, γ is the friction coefficient of the medium, $U_{A,B}(x)$ is the potential, k_B is Boltzmann's constant, and T is the temperature of the heat bath. The term $\sqrt{2m\gamma k_B T}\,dW$ represents the effect of noise from the surroundings. The Langevin equation is a stochastic differential equation, to be mathematically interpreted as a Stratonovich integral. For our case both the Ito and Stratonovich interpretations coincide [113, pp. 109] since the noise coefficient $\sqrt{2m\gamma k_B T}$ does not depend upon p. From [114, pp. 182], the generator for the Langevin equation 2.1 is

$$\mathcal{L} = \frac{p}{m}\partial_x - \left(\partial_x U_{A,B}(x)\right)\partial_p + \gamma\left(-p\partial_p + k_B T\partial_p^2\right) \tag{2.2}$$

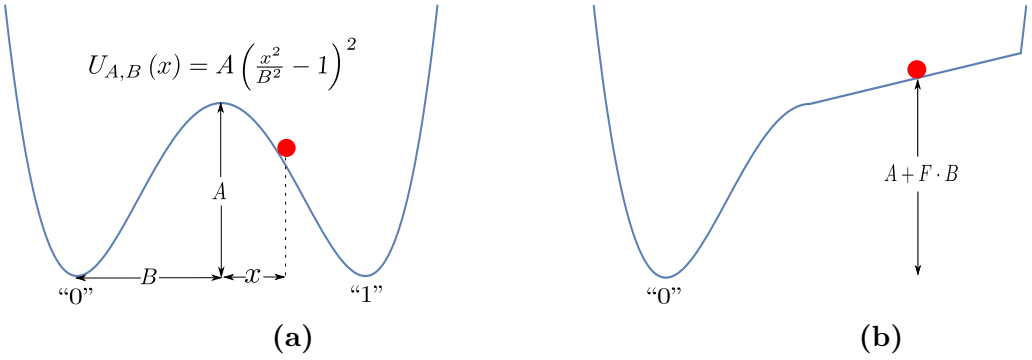


FIGURE 2.1: A bit as represented by a particle in a 1-D potential. Figure 2.1a: the bit in its resting state, with a barrier of height "A" separating particle locations that correspond to bit values of 0 or 1. Figure 2.1b: a control potential as in Example 2.2.1 is applied to erase the stored data.

The **Hamiltonian** of the system is $H(x,p) = U_{A,B}(x) + \frac{p^2}{2m}$. The **Gibbs distribution**

$$\pi(x,p) = \frac{e^{-H(x,p)/k_B T}}{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-H(x',p')/k_B T}\,dx'\,dp'} \tag{2.3}$$

is approached as the system relaxes to equilibrium. Convergence to $\pi(x,p)$ happens

exponentially fast at a rate given by the first non-zero eigenvalue of the generator $\mathcal{L}$ [115].

2.2.1 Reliability

A device to store information should be able to store it with high fidelity for a specified long period of time. We introduce the **reliability time** to represent the time-scale over which our device can store data. Specifically, we define the reliability time τ_r as the expected first passage time for the particle to cross the barrier of the resting-state potential of the bit, given the Gibbs distribution $\pi(x,p)$ (Equation 2.3) as the initial distribution. That is,

$$\tau_r := \mathbb{E}\left[\inf\{t \geq 0 \mid x(t) = 0\}\right] \tag{2.4}$$

where the expectation is over trajectories $(x(t), p(t))$ distributed as specified by Equation 2.1 from the initial condition $(x(0), p(0)) \sim_{\text{law}} \pi(x,p)$. Note that τ_r is also the first passage time to cross the barrier for a bit prepared with a Gibbs distribution, but confined to either the left-hand well $\pi_0(x,p)$ or right-hand well $\pi_1(x,p)$.

$$\pi_0(x,p) = \begin{cases} 2\pi(x,p) & \text{if x} < 0 \\ 0 & \text{otherwise} \end{cases}, \quad \pi_1(x,p) = \begin{cases} 2\pi(x,p) & \text{if x} > 0 \\ 0 & \text{otherwise} \end{cases}. \tag{2.5}$$

Intuitively, once a typical particle has had enough time to reach the top of the barrier, the data stored is no longer reliable.

2.2.2 Setting information

A device intended to store information must provide functionality to load, or set this information into the device. Setting information is a two-bit operation. A common use case is when a reference bit and the bit to be set are initially at some arbitrary values. We require that after the SET operation the reference bit is unchanged whereas the bit to

be set now holds a copy of the reference bit. This is the operation that Szilard [6] refers to as "copying" (in contrast, Landauer [9, 116] chooses to reserve the word "copying" for the operation where the bit to be set is initially already known to be in the state "0").

Note that in the operation of setting information, or copying in the sense of Szilard, initially the two bits are uncorrelated and unknown whereas after the operation they are still unknown but correlated. Thus implementing this operation requires decreasing the entropy of the system. Since it is easier to study a one-bit system rather than a two-bit system, we will investigate a one-bit proxy for the task of decreasing the entropy of the system, which is the task of **erasing** a bit.

Erasing involves taking a device whose initial state is maximally unknown into a known reference state, usually "0." Somewhat counter intuitively, given the name, erasing increases the information we know about the system. What is erased is not information but randomness. It helps to keep in mind the example of erasing a blackboard where some random state with chalk marks is reset to the "all clear" state.

2.2.3 Erasing

The example that follows describes a simple family of control potentials to implement the erasing operation for our device, which will form the basis of our analysis.

Example 2.2.1. Our control potentials are described by a single parameter $F \in \mathbb{R}_{>0}$ as follows.

$$V_F(x) := \begin{cases} A + F \cdot x - U_{A,B}(x) & \text{if } x \geq 0 \text{ and } A - U_{A,B}(x) + F \cdot x \geq 0, \\ 0 & \text{otherwise.} \end{cases} \tag{2.6}$$

One control potential from this family is illustrated in Figure 2.1b. We chose such a simple class of controls to make a full understanding feasible, setting a framework for analysing more complex protocols. We also note that arbitrary variation of a physical

potential in reality is highly non-trivial; experimental studies in which complex time-dependent potentials have been applied in fact use highly dissipative mechanisms to generate them [117, 118].

The Langevin equation in the presence of control is

$$\begin{aligned} m\,dx &= p\,dt \\ dp &= -\gamma p\,dt - \partial_x U_{\text{effec}}(x)dt + \sqrt{2m\gamma k_B T}\,dW \end{aligned} \tag{2.7}$$

, where $U_{\text{eff}}(x) = U_{A,B}(x) + V_F(x)$. More explicitly,

$$U_{\text{eff}}(x) := \begin{cases} A + F \cdot x & \text{if } x \geq 0 \text{ and } A - U_{A,B}(x) + F \cdot x \geq 0, \\ U_{A,B}(x) & \text{otherwise.} \end{cases} \tag{2.8}$$

Note that the control potential, as defined, is not differentiable at the boundary of the region in which it is non-zero. In practice, we assume that $\partial_x V_F$ changes rapidly but continuously in a small vicinity around these points.

In this work, we will consider variation of A, F and γ at fixed m, B, and T. In this case, m specifies the natural mass scale, B the natural length scale and $k_B T$ the natural energy scale; the natural time scale is then $\sqrt{mB^2/k_B T}$. Henceforth, all numerical quantities will be reported using reduced units defined with respect to these natural scales, although m, B and $k_B T$ will be retained within formulae.

Operational view of Erasing

The speed of bit operations is of practical importance: a useful bit must be reliable on much larger time-scales than those required to set or switch it. The control is switched on at time 0 and switched off at an appropriately-chosen time τ. The time τ is chosen beforehand, and does not depend on details of individual trajectories – a trajectory-dependent control would require measurement and feedback that itself would need ac-

counting for [119, 120, 121, 122, 123, 124]. We could declare erasing as completed and switch off the control as soon as a majority of trajectories are expected to be in the left well. However, many of these "erased" bits would have high energies compared to typical bits drawn from the equilibrium distribution in the left well, $\pi_0(x,p)$. Thus they could rapidly return to the right well after a very short stay in the left well. So we insist on a more stringent condition. We require that the time τ should be large enough so that the majority of bits are in the target well, with an expected next passage time close to the reliability time.

One way to guarantee that the next passage time is high is by insisting on mixing, in the sense that the initial distribution $\pi(x,p)$ comes close to a distribution of particles thermalised in the left-hand well, $\pi_0(x,p)$. If this happens, we can guarantee that the expected next passage time will be equal to, or close to, the expected first passage time. However, we found this criterion too stringent for the following reason. At the end of the erasing protocol, it is not necessary that the distribution is close to $\pi_0(x,p)$ – only that the particles tend to relax to this distribution much faster than they cross back into the right-hand well, and thus have barrier passage times representative of particles initialised with $\pi_0(x,p)$. Nonetheless, we show in Section 6.2.1 of Appendix A that using such a criterion preserves the qualitative features reported below (in particular, the scaling of erasure time with friction in the high and low friction limits).

Instead, we define an **erasure region** in well "0" as all points (x,p) with total energy $H(x,p) \leq A - 3k_BT$ where A is the barrier height. We look for the mean first passage time to reach the erasure region for particles initiated in well "1" and take this quantity to be representative of the erasing time-scale. The choice of $3k_BT$ criterion is somewhat arbitrary, but has been used before by Vega et al. [125] to study atom-surface diffusion. As we show in Section 6.2.2 of Appendix A, using $4k_BT$ makes no qualitative difference to our conclusions. This metric has the merit that it provides a clear computable criterion for erasing. Below, we demonstrate that particles within the $3k_BT$ erasure region do

indeed have expected next passage times close to the reliability time, as required.

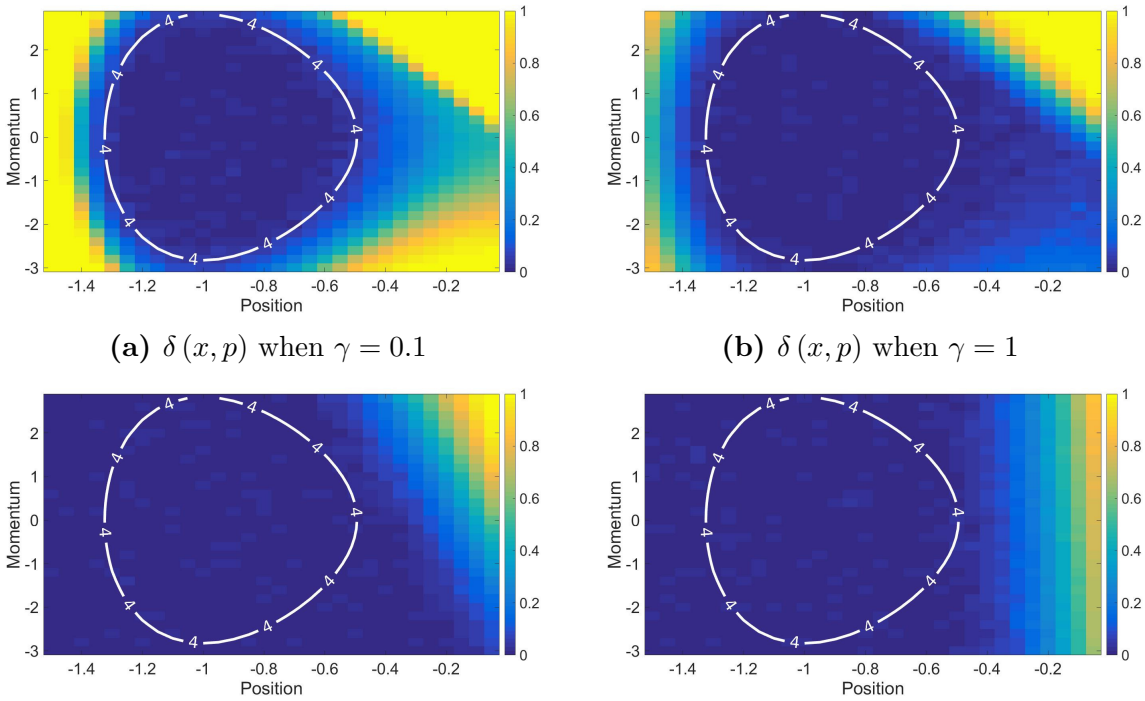
For a range of well parameters, we used the Langevin A algorithm from [126] [refer to Section 6.1 of Appendix A for the integrator set-up and validation], to estimate $\tau(x,p)$, the average barrier crossing time for particles initialised at position x with momentum p in the left well, for a grid of points (x,p). The average reliability time for a given well can be approximated in terms of an ensemble average of $\tau(x,p)$ as follows:

$$\tau_r \approx \frac{\sum_{x,p} \tau(x,p)\, \mathrm{e}^{-H(x,p)/k_BT}}{\sum_{x,p} \mathrm{e}^{-H(x,p)/k_BT}}. \tag{2.9}$$

The deviation $\delta(x,p) := |1-\frac{\tau(x,p)}{\tau_r}|$ for every point (x,p) in the grid is plotted in Figure 2.2, for a range of friction parameters at well height $A = 7$. It is clear from the figure that, for all values of friction, the points with total energy $H(x,p) \leq A - 3k_BT$ have reliability times close to τ_r. The same is true of other well heights A. This is because such particles typically undergo thermal mixing before they can escape the well. Once mixed, their next escape over the barrier will be on a time-scale of the order of τ_r.

Despite the robustness of this result to the value of the friction, the heatmaps in Figure 2.2 are friction-dependent. When γ is low, the particle diffuses very slowly in energy space, and it is the challenge of diffusing within this energy space that prohibits escape from the well. As a result the heatmap corresponding to $\gamma = 0.1$ (Figure 2.2a) follows the shape of constant energy contours. As friction starts increasing (e.g. in Figure 2.2b and 2.2c), diffusion in momentum-space becomes more rapid, but diffusion in position-space slows down. Once γ becomes very high (e.g. $\gamma = 100$ in Figure 2.2d), the behaviour of the heatmap is essentially determined by the initial position of the particle; those close to the barrier and with $U_{A,B}(x)$ sufficiently close to A can escape easily, but the momentum is irrelevant. Using the total energy $H(x,p)$ as a criterion ensures that we account for all the regimes of friction.

Since we are interested in the typical time scale of transferring particles to a different



(a) $\delta(x,p)$ when $\gamma = 0.1$ **(b)** $\delta(x,p)$ when $\gamma = 1$

(c) $\delta(x,p)$ when $\gamma = 10$ **(d)** $\delta(x,p)$ when $\gamma = 100$

FIGURE 2.2: For particles initiated with $H(x,p) \leq A - 3k_BT$, well escape times are close to τ_r. Heat maps show the fractional deviation in expected escape time $\delta(x,p)$ from the well-thermalised average τ_r, as a function of initial position x and momentum p. The labelled contours correspond to a well height $A = 7$ with energy $H(x,p) = A - 3k_BT = 4k_BT$. These heat maps are representative of the situation for other barrier heights $A \geq 5k_BT$.

well from the existent well, we will sample initial points only from the right well. We define the **erasing time** τ_e as the expected time to hit the erasure region, given that the particle started in the right-hand well:

$$\tau_e = \mathbb{E}\left[\inf\{t \geq 0 \mid x(t) < 0 \text{ and } H(x(t), p(t)) \leq A - 3k_BT\}\right] \tag{2.10}$$

where $(x(t), p(t))$ is the solution to Equation 2.7 with the initial condition $(x(0), p(0)) \sim_{\text{law}} \pi_1(x,p)$. Given this definition, τ_e indicates a typical time scale over which the control must be applied to successfully erase a large fraction of the bits. In practice, the control would be applied for a period $\tau > \tau_e$ to achieve high accuracy. We will use τ_e as an indicative time scale of control operation for the purposes of our analysis. It is useful to decompose the erasing time τ_e as the sum of two times: the **transport time** and **mixing time**.

- **Transport time (τ_t)**: The time taken by the particle to reach well "0" given that it is initially distributed according to $\pi_1(x, p)$.

$$\tau_t = \mathbb{E}\left[\inf\{t \geq 0 \mid x\left(t\right) \leq 0\}\right] \tag{2.11}$$

where $x(t)$ is the solution to Equation 2.7 with the initial condition $(x(0), p(0)) \sim_{\text{law}} \pi_1(x, p)$.

- **Mixing time (τ_m)**: The time taken by the particle to mix sufficiently inside the well. This is the time starting from when the particle first reaches well "0" to when it first hits the erasure region.

$$\tau_m = \tau_e - \tau_t \tag{2.12}$$

Cost of erasing

In this section, we calculate the work done in erasing a bit. From Sekimoto's expression [25, 127], for a protocol applied for a time τ and to a region $I = \{x \geq 0 \mid A - U_{A,B}\left(x\right) + F \cdot x \geq 0\}$,

$$\langle W \rangle := \int_0^\tau \int_{x \in I} \frac{\partial V_F\left(x, t\right)}{\partial t} p(x, t)\, dx\, dt, \tag{2.13}$$

where $p(x, t)dxdt$ is the probability that the particle is between position $(x, x + dx)$ in the time interval $(t, t + dt)$. There are two potential sources of work that appear in our calculation.

1. **When we begin the erasure protocol by switching on the control to lift the particle**: We first give an intuitive argument for the expression of work, and then justify it both analytically and numerically. The particle's initial potential energy is approximately $k_B T/2$ on average, due to the equipartition theorem, and after the control is switched the average potential energy is $A + F \cdot B$ for a particle in the right

well, since the particle is localised around $x = B$, and still $k_BT/2$ for a particle in the left well. Therefore, the work done in this step is $W = (A + F \cdot B - k_BT/2)/2$. We now justify that this approximation is accurate for the values of A and F that we consider, and we will use this as the form of work for the rest of the chapter.

Since we switch on the control instantaneously, $\frac{\partial V_F(x,t)}{\partial t} = V_F(x,t)\,\delta(t)$. From Equation 2.13, and letting $I = \{x \geq 0 \mid A - U_{A,B}(x) + F \cdot x \geq 0\}$, we get

$$\begin{aligned} \langle W \rangle &= \int_0^{\tau} \int_I V_F(x,t)\, p(x,t)\, \delta(t)\, dx dt \\ &= \int_I (A - U_{A,B}(x) + F \cdot x)\, p(x,0)\, dx \end{aligned} \tag{2.14}$$

Since $p(x,0) = \frac{e^{-\frac{U_{A,B}(x)}{k_BT}}}{\int_{-\infty}^{\infty} e^{-\frac{U_{A,B}(x)}{k_BT}}}$, we can rewrite the expression of work as

$$\langle W \rangle = \frac{\int_I (A - U_{A,B}(x) + F \cdot x)\, e^{-\frac{U_{A,B}(x)}{k_BT}}\, dx}{\int_{-\infty}^{\infty} e^{-\frac{U_{A,B}(x)}{k_BT}}\, dx} \tag{2.15}$$

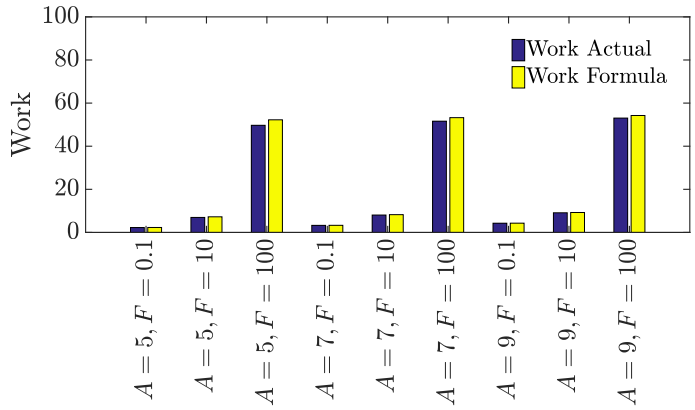


FIGURE 2.3: Work when the control is switched on against various values of A and F, comparing the full expression Eq. 2.15 and the approximate result Eq. 2.20.

Since $I \subseteq [0, \infty)$ and $(A - U_{A,B}(x) + F \cdot x)\, e^{-\frac{U_{A,B}(x)}{k_BT}}$ is negligible as $x \to \infty$, replac-

ing the upper limit of integration by ∞ is reasonable. Hence the integral becomes

$$\langle W \rangle \approx \frac{\int_0^\infty (A - U_{A,B}(x) + F \cdot x)\, \mathrm{e}^{-\frac{U_{A,B}(x)}{k_B T}}\, dx}{\int_{-\infty}^{\infty} \mathrm{e}^{-\frac{U_{A,B}(x)}{k_B T}}\, dx} \tag{2.16}$$

Since our potential $U_{A,B}(x) = A(\frac{x^2}{B^2} - 1)^2$ is symmetric about $x = 0$, we get that $\int_{-\infty}^{\infty} \mathrm{e}^{-\frac{U_{A,B}(x)}{k_B T}}\, dx = 2\int_0^{\infty} \mathrm{e}^{-\frac{U_{A,B}(x)}{k_B T}}\, dx$. The expression of work then becomes

$$\langle W \rangle \approx \frac{\int_0^\infty (A - U_{A,B}(x) + F \cdot x)\, \mathrm{e}^{-\frac{U_{A,B}(x)}{k_B T}}\, dx}{2\int_0^{\infty} \mathrm{e}^{-\frac{U_{A,B}(x)}{k_B T}}\, dx} \tag{2.17}$$

When $A \gg k_B T$, we can safely assume that the particle only samples the region of space near the bottom of the well. As a result, we can approximate the potential corresponding to the region $x > 0$ by a harmonic oscillator and set the lower limit of integration in both the numerator and denominator to $-\infty$. As we show in Equation 2.24, $U_{A,B}(x) \approx \frac{4A(x-B)^2}{B^2}$ for $x > 0$. Therefore, Equation 2.17 becomes

$$\begin{aligned} \langle W \rangle &\approx \frac{\int_{-\infty}^{\infty} (A - \frac{4A(x-B)^2}{B^2} + F \cdot x)\, \mathrm{e}^{-\frac{4A(x-B)^2}{k_B T B^2}}\, dx}{2\int_{-\infty}^{\infty} \mathrm{e}^{-\frac{4A(x-B)^2}{k_B T B^2}}\, dx} \qquad (2.18) \\ &= \frac{A}{2} - \frac{\int_{-\infty}^{\infty} \frac{4A(x-B)^2}{B^2}\, \mathrm{e}^{-\frac{4A(x-B)^2}{k_B T B^2}}\, dx}{2\int_{-\infty}^{\infty} \mathrm{e}^{-\frac{4A(x-B)^2}{k_B T B^2}}\, dx} + \frac{F\int_{-\infty}^{\infty} (x-B)\, \mathrm{e}^{-\frac{4A(x-B)^2}{k_B T B^2}}\, dx}{2\int_{-\infty}^{\infty} \mathrm{e}^{-\frac{4A(x-B)^2}{k_B T B^2}}\, dx} + \frac{F \cdot B}{2} \end{aligned}$$

Let $y = x - B$. As $x \to -\infty, y \to -\infty$ and as $x \to \infty, y \to \infty$. Therefore, the expression of work becomes

$$\begin{aligned} \langle W \rangle &\approx \frac{A}{2} - \frac{\int_{-\infty}^{\infty} \frac{4Ay^2}{B^2}\, \mathrm{e}^{-\frac{4Ay^2}{k_B T B^2}}\, dx}{2\int_{-\infty}^{\infty} \mathrm{e}^{-\frac{4Ay^2}{k_B T B^2}}\, dx} + \frac{F\int_{-\infty}^{\infty} y\, \mathrm{e}^{-\frac{4Ay^2}{k_B T B^2}}\, dx}{2\int_{-\infty}^{\infty} \mathrm{e}^{-\frac{4Ay^2}{k_B T B^2}}\, dx} + \frac{F \cdot B}{2} \qquad (2.19) \\ &\approx \frac{A}{2} - \frac{\frac{1}{4} AB\sqrt{\pi}(\frac{k_B T}{A})^{3/2}}{B\sqrt{\pi \frac{k_B T}{A}}} + 0 + \frac{F \cdot B}{2} \end{aligned}$$

Therefore, our expression for work becomes

$$\langle W \rangle \approx \frac{\left(A + F \cdot B - \frac{k_B T}{2}\right)}{2} \tag{2.20}$$

justifying our physical intuition. The accuracy of this expression compared to Eq. 2.15 is illustrated in Figure 2.3.

2. **At the end of the protocol when we switch off the control**: We show that in our family of controls, there is negligible energy recovered when the control is switched off, since the probability of the particle being in the region in which the control is applied is small. More generally, the question of whether energy might be recovered from small systems and stored efficiently is a complex one, despite the optimism shown in previous discussions of erasing. Indeed, current technology does not attempt to recover any energy from bits.

   To illustrate this more formally, we assume that the control is switched off after a time τ sufficiently large compared to τ_e so that the proportion of particles remaining on the right hand side of the well is determined by the Boltzmann factor. The work that we could then in principle recover is given by the following expression:

$$\langle W_{rec} \rangle \approx \frac{\int_I \left(A - U_{A,B}(x) + F \cdot x\right) \mathrm{e}^{\frac{-(A+F\cdot x)}{k_B T}} \, dx}{\int_{-\infty}^{0} \mathrm{e}^{\frac{-U_{A,B}(x)}{k_B T}} \, dx + \int_I \mathrm{e}^{\frac{-(A+F\cdot x)}{k_B T}} \, dx} \tag{2.21}$$

   Recall that $I = \{x \geq 0 \mid A - U_{A,B}(x) + F \cdot x \geq 0\}$. This implies that $I = [0, x^*]$, where $A - U_{A,B}(x^*) + F \cdot x^* = 0$. Using Equations 2.20 and 2.21, we will calculate the fraction of recovered work i.e., $W_{rec}^f = \frac{\langle W_{rec} \rangle}{\langle W \rangle}$. Figure 2.4 precisely calculates this quantity. As is evident from the figure, the fraction is almost negligible and reaches its maximum value at low A and F.

Observation 2.2.2. Work is an increasing function of well height A at fixed F and γ. This follows immediately from the expression of work $W = (A + F \cdot B - k_B T/2)/2$.

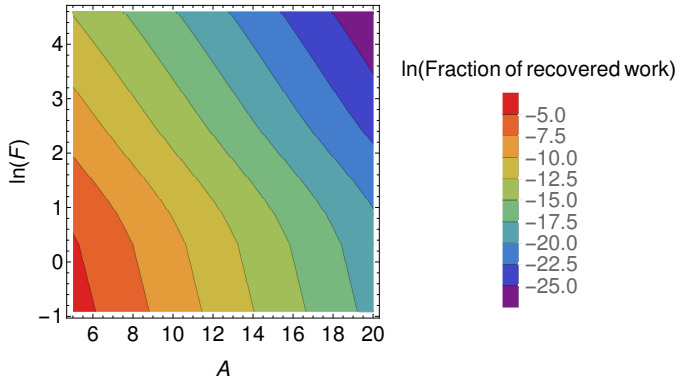


FIGURE 2.4: Negligible energy can be recovered for our family of controls.

2.3 Friction-based trade-offs for reliability and erasing

We explore the behaviour of the reliability and erasing time-scales as functions of the friction coefficient. We find that both these time-scales are non-monotonic, roughly U-shaped functions of the friction coefficient. A high reliability time requirement is favoured by a very low or very high friction; whereas a low erasing time requirement is helped by the choice of a moderate value of friction. Since a bit designer would seek reliable bits (needing high or low friction) that can be erased fast (needing intermediate friction), this yields a friction-based trade-off between reliability and speed of erasure.

2.3.1 Reliability Time

Our definition of reliability time (Equation 2.4) is very similar to the classic problem of escape rates from one-dimensional wells (Fig 2.5), as applied in transition state theory to understand chemical reactions. In a famous paper [128], Kramer found analytic expressions for the escape rate k from a well by calculating the flux of particles between a source on one side of the barrier (x_A) and a sink at the other side (x_B). Kramer's expressions

apply separately to the regimes of low friction, moderate to high friction and very high friction.

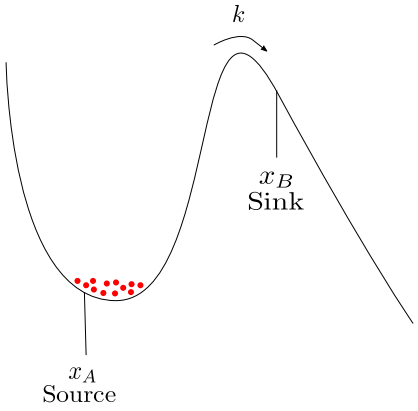


FIGURE 2.5: The escape of particles from a one-dimensional well. Kramer [128] considered a source of particles at the bottom of the well, and estimated the rate of escape to a sink on the far side of a barrier.

Later the groups of Melnikov and Meshkov [129] and Pollack, Grabert and Hänggi [130] gave formulae that interpolate accurately over all values of friction (see review in [131]). We illustrate the key ideas used in the derivation of Melnikov and Meshkov [129]. The starting point is to describe the dynamics in the low friction regime by a Fokker-Planck equation in energy-action variables. Subsequently, this leads to an integral equation with the Green function as its kernel. The integral equation is then solved using the Wiener-Hopf method by explicitly calculating certain Fourier transforms. Refer to [129] for a detailed derivation of these facts. In particular, we shall use the following formula [129, Equations 6.1a and 6.3] to estimate analytical forms of the escape rate for our bistable system.

$$\begin{aligned} k &= \frac{\omega_0}{2\pi}\left[\sqrt{1+\frac{\gamma^2}{4\omega_b^2}}-\frac{\gamma}{2\omega_b}\right] g\, \mathrm{e}^{-A/k_BT}, \text{where} \\ \ln g &= \frac{1}{2\pi}\int_0^{\frac{\pi}{2}} \ln\left[1-\exp\left(\frac{-\gamma\, I(A)}{4k_BT\cos^2 x}\right)\right] dx. \end{aligned} \tag{2.22}$$

Here, ω_b is the angular frequency at barrier height, ω_0 is the angular frequency at the bottom of the well and $I(A)$ is the action for barrier height A. In order to calculate these

parameters for our system $U_{A,B}(x) = A(\frac{x^2}{B^2} - 1)^2$, we need $\partial_x U_{A,B}(x) = \frac{4Ax\left(x^2 - B^2\right)}{B^4}$ and $\partial_{xx} U_{A,B}(x) = \frac{4A\left(3x^2 - B^2\right)}{B^4}$. Then, we have the following

1. ω_b: We can approximate the region near the barrier by an inverted harmonic oscillator. By Taylor expanding the potential about the point $x = 0$, we get

$$\begin{aligned} U_{A,B}(x) &\approx U(0) + \partial_x U_{A,B}(x)\Big|_{x=0} x + \frac{\partial_{xx} U_{A,B}(x)\Big|_{x=0} x^2}{2} \\ &\approx A - \frac{2Ax^2}{B^2} = A - \frac{m\omega_b^2 x^2}{2} \end{aligned} \tag{2.23}$$

   Therefore we have $\omega_b = \sqrt{\frac{4A}{mB^2}}$.

2. ω_0: We can approximate the region near the bottom of the well by a harmonic oscillator. By Taylor expanding the potential about the point $x = B$, we get

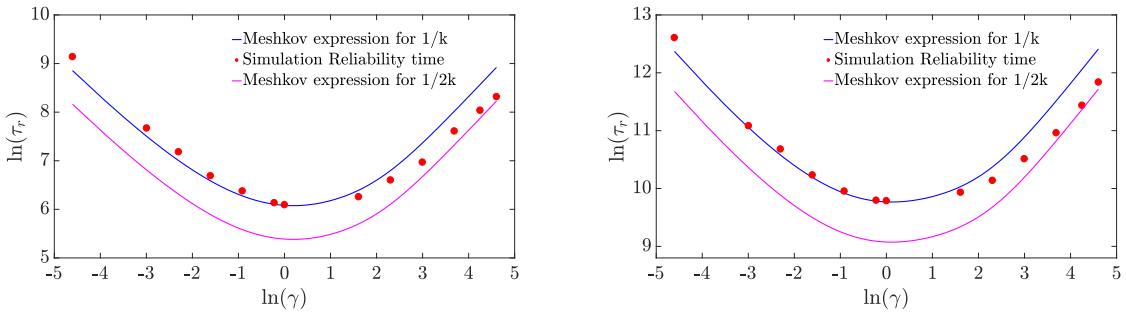
$$\begin{aligned} U_{A,B}(x) &\approx U(B) + \partial_x U_{A,B}(x)\Big|_{x=B} (x - B) + \frac{\partial_{xx} U_{A,B}(x)\Big|_{x=B} (x - B)^2}{2} \\ &= \frac{4A(x - B)^2}{B^2} = \frac{m\omega_0^2 (x - B)^2}{2} \end{aligned} \tag{2.24}$$

   Therefore we have $\omega_0 = \sqrt{\frac{8A}{mB^2}}$.

3. $I(A)$: Consider a particle of mass m with a starting velocity $v = 0$, moving along a constant energy surface with energy A. The particle starts at $x = 0$ and moves to $x = \sqrt{2}B$ and returns back to $x = 0$. The action for this round trip is given by $I(A) = \oint p dx = 2\sqrt{2m} \int_0^{\sqrt{2}B} \sqrt{A - A\left(\frac{x^2}{B^2} - 1\right)^2} dx = \frac{8B\sqrt{mA}}{3}$.

We plot the analytical prediction of $1/k$ given by Eq. 2.22 in Fig. 2.6 for two values of well height A, as a function of friction γ. This prediction is compared to average first passage time for particles to reach the top of the barrier from an initial Boltzmann distribution within a single well. The two quantities differ at large γ because Kramer's definition does not treat a particle that crosses the barrier but then immediately crosses back as having

"escaped", whereas our definition of reliability in terms of a first passage time treats such particles as no longer being reliable. In the underdamped regime, immediate recrossings are rare and hence τ_r and $1/k$ coincide; in the overdamped regime, particles that reach the barrier top have a 50% chance of returning and so $\tau_r = 1/2k$. As can be seen from Fig. 2.6, τ_r smoothly interpolates between $\frac{1}{k}$ and $\frac{1}{2k}$, with the small numerical factor providing only a minor correction to the underlying physics of the analytical expression in Eq. 2.22. Note that all stochastic simulations here and elsewhere in the chapter were performed by using the Langevin A integrator [126] and averaging the results over many initial conditions.



(a) $A = 6$ **(b)** $A = 10$

FIGURE 2.6: The reliability time $\tau_r \propto \frac{1}{\gamma}$ in the low friction regime, $\tau_r \propto \gamma$ in the high friction regime and is minimum at moderate friction. Simulation results are compared to the inverse of escape rate from a single well $(1/k)$ and $(1/2k)$, as predicted by Eq. 2.22. Here, and elsewhere in the chapter, error bars are omitted when comparable to data points.

The Melnikov-Meshkov expression predicts an almost-exponential scaling of $1/k$ with barrier height A, which is reproduced by τ_r and expected from the Arrhenius rate law [132]. Note that both $1/k$ and τ_r are non-monotonic in friction γ, with long reliability times in the underdamped and overdamped limits. This behaviour results from the need for particles to diffuse in both position and energy in order to reach the top of the barrier from an initial state thermalized within a single well. At high friction, particles rapidly sample different kinetic energies due to strong coupling with the environment, but move slowly in position space and hence take a long time to cross the barrier. At low friction, particles can move rapidly but their energy remains effectively constant over short time periods. They only cross the barrier when they have eventually gained enough total en-

ergy. Intermediate friction, when neither process is excessively slow, gives the shortest τ_r. This behaviour is typical of equilibrating systems in which an initial out-of-equilibrium condition (particles are guaranteed to be on one side of the well and not the other) relaxes towards an equilibrium state (particles on both sides of the barrier), and is thus insensitive to the details of our bit design.

2.3.2 Erasing time

As noted earlier, the erasing time is composed of two parts: the transport time defined in Eq. 2.11 and the mixing time defined in Eq. 2.12. We now present analytical estimates of these times and compare them with numerical solutions.

Transport time

Using Equations 2.7 and 2.8 for the region in which control is applied i.e. $x \geq 0$ and $A - U_{A,B}(x) + F \cdot x \geq 0$, the Langevin equation becomes

$$\begin{aligned} m\,dx &= p\,dt \\ dp &= -\gamma p\,dt - F dt + \sqrt{2m\gamma k_B T}\,dW \end{aligned} \tag{2.25}$$

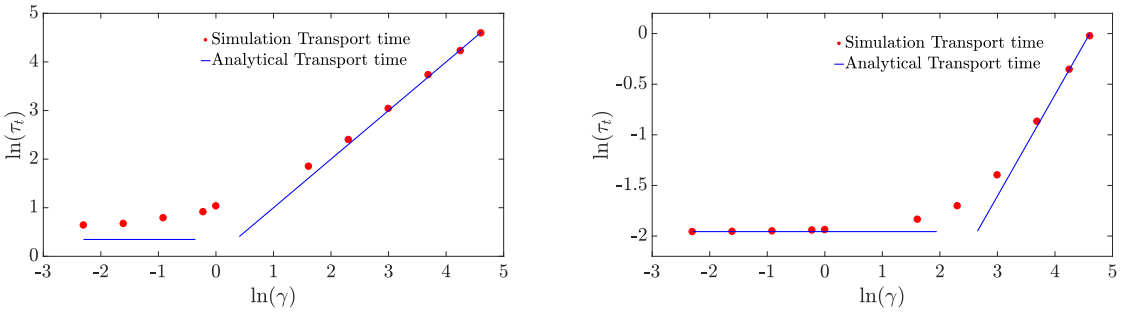
Dividing by dt throughout and taking ensemble averages, we obtain(after noting that the operations of averaging and differentiation with respect to time commute)

$$\begin{aligned} m\frac{d\langle x\rangle}{dt} &= \langle p\rangle \\ \frac{d\langle p\rangle}{dt} &= -\gamma\langle p\rangle - F \end{aligned} \tag{2.26}$$

For the sake of rough calculation, we can approximate the potential for the region $x > 0$ by a harmonic oscillator. Using Equation 2.24, we get $U_{A,B}(x) \approx \frac{4A(x-B)^2}{B^2}$, a harmonic oscillator symmetric about $x = B$. Therefore, on expectation, the particle starts at $x = B$

and the average distance travelled by the particle during the period of transport time is B.

1. **Low friction regime**: When friction is very low, Equation 2.26 reduces to $\frac{d\langle p\rangle}{dt} = -F$. Therefore, the particle travels with an average acceleration $\langle a\rangle = -\frac{F}{m}$. Using the laws of kinematics, we get $B = \langle u\rangle\tau_t + \frac{1}{2}\langle a\rangle\tau_t^2$. Plugging in $\langle u\rangle = 0$ as the average initial velocity, we get $\tau_t \approx \sqrt{2mB/F}$.



(a) $A = 10, F = 1$ **(b)** $A = 10, F = 100$

FIGURE 2.7: The transport time obtained from simulations approximates the analytical estimates of $\tau_t \approx \sqrt{2mB/F}$ in the low friction regime, $\tau_t \approx mB\gamma/F$ $(\propto \gamma)$ in the high friction regime.

2. **High friction regime**: Since friction is very high, we can assume that the net average force on the particle is 0 i.e. $\frac{d\langle p\rangle}{dt} = 0$. Plugging in $\frac{d\langle p\rangle}{dt} = 0$ in Equation 2.26, we get $\gamma m\langle v\rangle = \gamma\langle p\rangle = -F$. Therefore, the particle travels with an average velocity of $\langle v\rangle = -\frac{F}{m\gamma}$. The time taken to travel a distance B is $\tau_t \approx mB\gamma/F$.

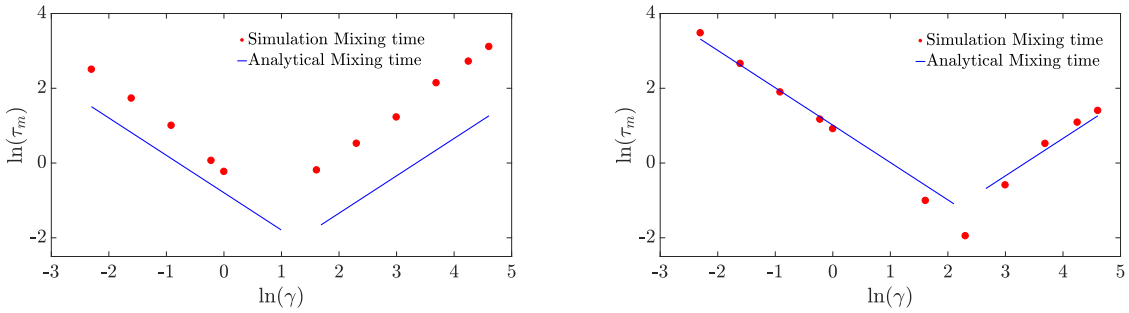
We thus expect the transport time to be constant in the underdamped regime and increase linearly with friction in the overdamped regime. Figure 2.7 illustrates that this scaling is observed in Langevin simulations, and that numerical values are in reasonable agreement with these estimates. The largest quantitative deviations occur at low force and low friction (e.g. $F = 1$ in Figure 2.7a), when the diffusion of the particle on the slope contributes significantly to τ_t. This results in a simulation transport time larger than the analytical estimate.

Mixing time

Similar to the transport time, analytical estimates of the mixing time can be obtained in the limits of high and low friction.

1. **Low friction regime**: For purposes of approximate calculation we treat the well "0" as a harmonic oscillator. Deterministically, the energy of a harmonic oscillator decays exponentially in the underdamped regime. Therefore we have $E(t) = E_0 \, \mathrm{e}^{-\gamma t}$, where E_0 is the initial energy of the particle when it first reaches $x = 0$ and $E(t)$ is the energy of the particle at time t. In the underdamped regime, a particle starting at B arrives at position $x = 0$ with energy $E_0 \approx A + F \cdot B$. Thus solving for $E(\tau_{\mathrm{mix}}) = A - 3k_B T$,

$$\tau_{mix} \approx \frac{1}{\gamma} \log \frac{A + F \cdot B}{A - 3k_B T} \tag{2.27}$$



(a) $A = 10, F = 1$ **(b)** $A = 10, F = 100$

FIGURE 2.8: Evidence from simulation that the mixing time $\tau_{mix} \approx \frac{1}{\gamma} \log \frac{A+F \cdot B}{A-3k_B T} \left(\propto \frac{1}{\gamma} \right)$ in the low friction regime, $\tau_{mix} \approx \frac{mB^2 \gamma}{2\sqrt{2}A} \, (\propto \gamma)$ in the high friction regime and is minimised at moderate friction.

2. **High friction regime**: A sensible estimate of the behaviour can be obtained by explicitly modelling the diffusion of the particle near the barrier top. In the overdamped limit, the criterion of reaching a total energy of $E(\tau_{\mathrm{mix}}) = A - 3k_B T$ is equivalent to reaching a point d which has potential energy of $A - 3k_B T$, since momenta are sampled arbitrarily rapidly in this limit. To proceed, we consider

the typical time required to reach an absorbing barrier at d starting from $x = 0$, assuming a sufficiently large F that we can treat $x = 0$ as a reflecting barrier. Starting from the overdamped stochastic differential equation

$$m\gamma\, dx = -\partial_x U_{A,B}(x)\; dt + \sqrt{2m\gamma k_B T}\, dW \tag{2.28}$$

with generator $\mathcal{L} = \frac{k_B T}{m\gamma} \mathrm{e}^{\frac{U_{A,B}(x)}{k_B T}} \partial_x \mathrm{e}^{\frac{-U_{A,B}(x)}{k_B T}} \partial_x$, we apply the standard methods outlined in Pavliotis [114, (7.1), pp. 239], which leads to the following system of equations for the average mixing time $\tau_{\mathrm{mix}}(x)$ as a function of the initial position x

$$\begin{aligned} \frac{k_B T}{m\gamma} \mathrm{e}^{\frac{U_{A,B}(x)}{k_B T}} \partial_x \mathrm{e}^{\frac{-U_{A,B}(x)}{k_B T}} \partial_x \tau_{mix}(x) = & -1, d < x \leq 0. \\ \tau_{mix}(x) = 0, & x = d. \end{aligned} \tag{2.29}$$

We can solve Equation 2.29 using appropriate limits to get

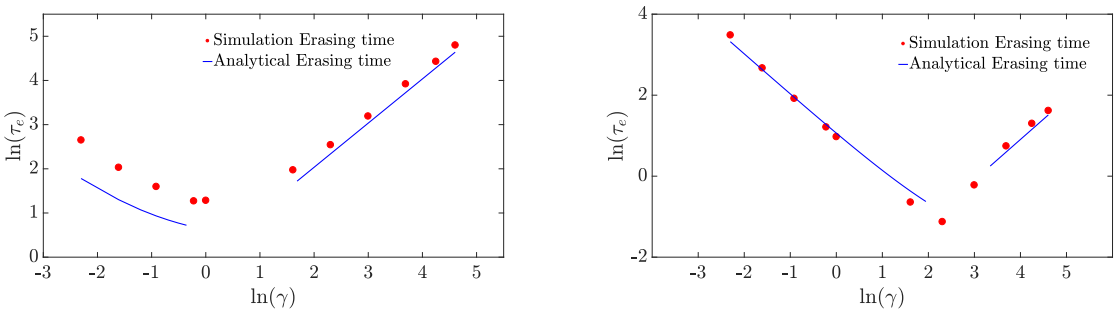
$$\tau_{mix}(x) = \frac{m\gamma}{k_B T} \int_0^{\sqrt{\frac{3k_B T}{2A}} B} \int_0^q \mathrm{e}^{\frac{U(q)-U(r)}{k_B T}}\, dqdr, \tag{2.30}$$

where we have approximated the potential near the barrier as an inverted harmonic oscillator to estimate $d = \sqrt{\frac{3k_B T}{2A}} B$. Repeating this approximation within the integral, we obtain

$$\tau_{mix}(x) \approx \frac{m\gamma}{k_B T} \int_0^{B\sqrt{\frac{3k_B T}{2A}}} \int_0^q \mathrm{e}^{\frac{2A\left(r^2 - q^2\right)}{B^2 k_B T}}\, dqdr \approx \frac{mB^2\gamma}{2\sqrt{2}A}. \tag{2.31}$$

Equations 2.27 and 2.31 predict that the mixing time will scale as $1/\gamma$ in the low friction limit and as γ in the high friction limit. In the first case, mixing within the well is limited by the rate at which the particle can reduce its total energy, whereas in the second it is determined by the speed with which the particle can diffuse in position space to a configuration with lower potential energy. We plot simulation results for the mixing time, along with the analytic predictions, in Fig. 2.8, confirming this scaling and the resultant non-monotonicity. Quantitatively, simulation results deviate from the crude analytic predictions at low force (e.g. $F = 1$ in Figure 2.8a), when it is no longer reasonable to treat $x = 0$ as either a reflecting barrier or a steep side of a harmonic well.

Instead, excursions of the particle back onto the slope occupying the region $x > 0$ lead to much larger mixing times. Nonetheless, the scaling and non-monotonicity in friction are preserved. Combining τ_{trans} and τ_{mix} gives τ_e, plotted in Fig. 2.9. Analytically, the erasing time is given as:



(a) $A = 10, F = 1$ **(b)** $A = 10, F = 100$

FIGURE 2.9: Evidence from simulation that the erasing time $\tau_e \approx \sqrt{\frac{2mB}{F}} + \frac{1}{\gamma}\log\frac{A+F\cdot B}{A-3k_BT}$ in the low friction regime, scaling as $1/\gamma$, and $\tau_e \approx \frac{mB\gamma}{F} + \frac{mB^2\gamma}{2\sqrt{2}A}$ in the high friction regime, scaling as γ. The erasing time is minimised at moderate friction.

1. **Low friction regime**:

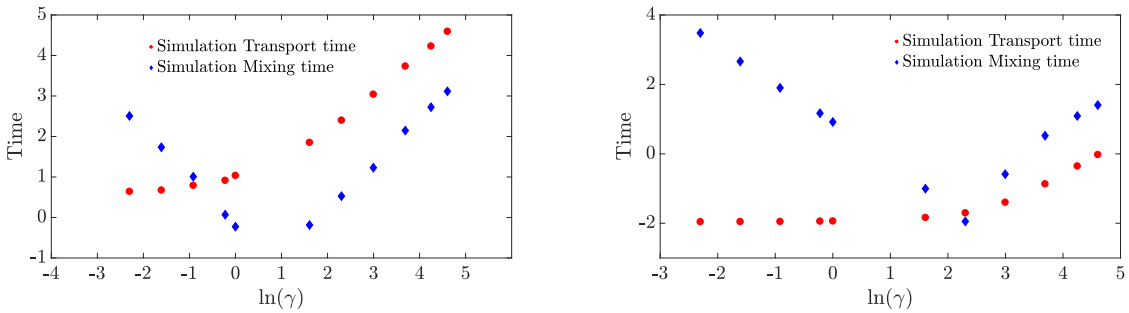
$$\tau_e \approx \sqrt{\frac{2mB}{F}} + \frac{1}{\gamma}\log\frac{A + F\cdot B}{A - 3k_BT}. \tag{2.32}$$

2. **High friction regime**:

$$\tau_e \approx \frac{mB\gamma}{F} + \frac{mB^2\gamma}{2\sqrt{2}A} \tag{2.33}$$

Like reliability, erasing time is large in the underdamped and overdamped limits, and minimized at intermediate values of friction. The physical cause is the same as before; our erasing protocol involves setting the system into a non-equilibrium state, and waiting for the system to relax towards an equilibrium in the perturbed potential. This process requires the system to diffuse in energy space and also explore configuration space, and is therefore favoured by intermediate friction. Specifically, if the friction is too low, the particle oscillates and slowly loses energy to be confined within the desired well. If the friction is too high, both the transport and mixing times increase as the particle's

movement through space is so slow. The relative importance of these effects can be seen in Fig. 2.10. We note that the value of the damping γ that minimises τ_e is quite sensitive to F (Fig. 2.10). Fundamentally, a larger F means the challenge of moving in position-space is made easier, and a greater loss of energy is needed to reach equilibrium. Therefore a higher friction coefficient is optimal. As with the reliability time, further analysis is possible but not necessary for the conclusions we wish to draw. Once again, the key point is the trade-off between high and low friction, which is not specific to our control. The trade-off is likely to be quite generic since any protocol will necessary push the system out of equilibrium, and will require particles to be typically confined within the target well before the control is removed.



(a) $A = 10, F = 1$ **(b)** $A = 10, F = 100$

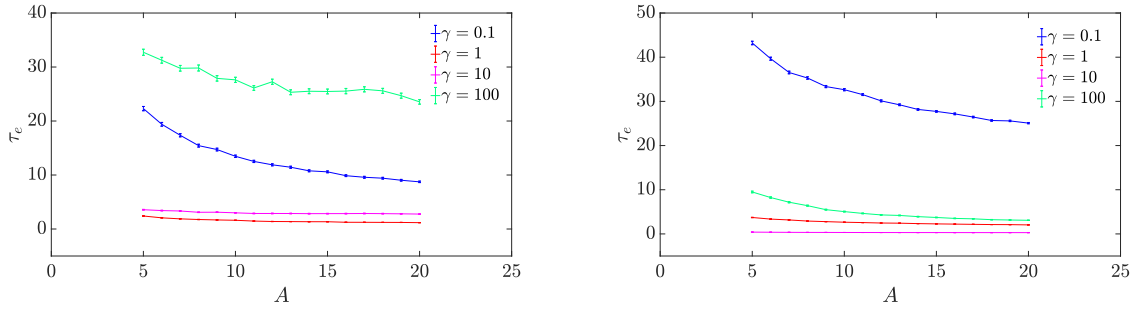
FIGURE 2.10: Comparison of transport and mixing times. Transport time dominates the mixing time for low force at high friction.

Both erasing and reliability times exhibit a trade-off in friction, being minimised by intermediate values. This fact sets up a second trade-off between designing bits with extreme values of friction to optimise reliability, or moderate values of friction to optimise erasing. The consequences of this secondary trade-off will be explored in Section 2.4.

2.3.3 Additional dependencies of the erasing time

A larger value of A implies a steeper descent into the target left-hand well, making mixing faster. We therefore expect that the mixing time and hence the erasing time monotonically decreases with A.

Observation 2.3.1. The erasing time is a strictly decreasing function of well height A at fixed F, γ. This can be seen from the analytic expressions of erasing time (Equations 2.32 and 2.33) backed up with numerical simulations (Figure 2.11).



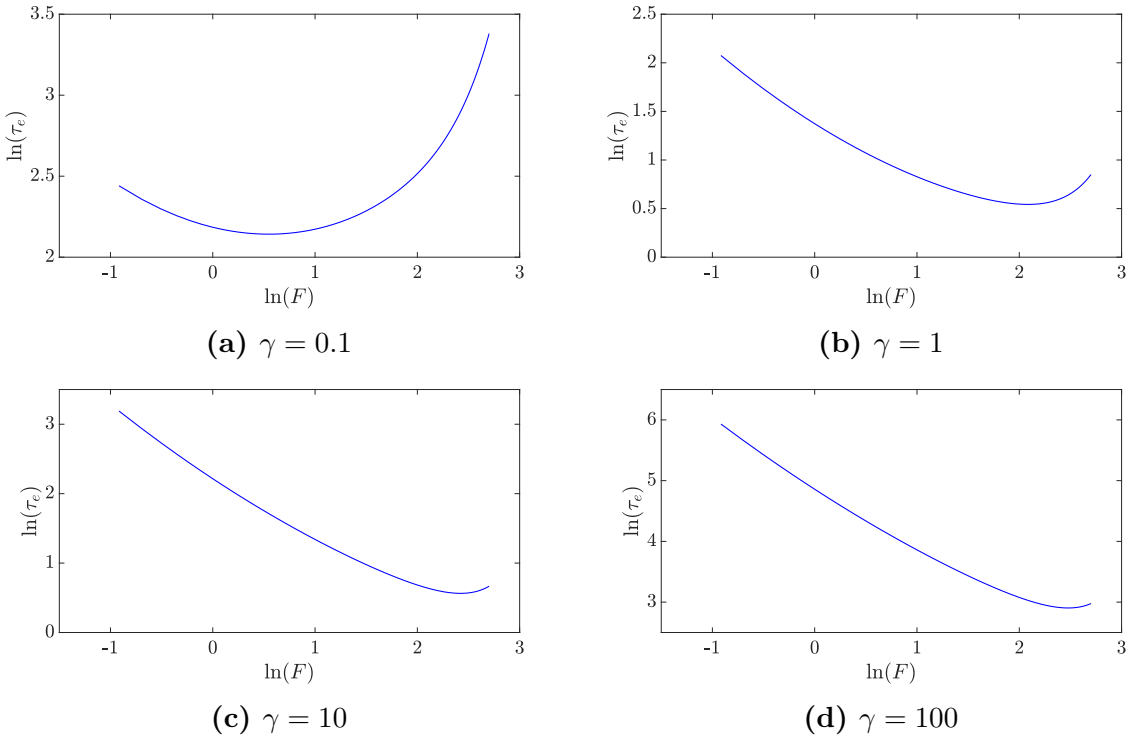
(a) $F = 5$ **(b)** $F = 100$

FIGURE 2.11: Evidence that the erasing time is a strictly decreasing function of well height across a range of F and γ. Other values of F and γ show similar behaviour.

By contrast, erasing time shows a non-monotonic dependence on F at fixed A, γ. Applying too little force leads to slow transport, and doesn't effectively trap the particle within the target well. But applying too much force supplies the particle with too much energy, which must subsequently be lost during the mixing period. We illustrate the non-monotonicity of τ_e on F at fixed $W = A + F$ and γ in Figure 2.12, in which simple regression formulae have been fitted to the simulation data to enable interpolation at fixed W and γ (see Section 6.3 of Appendix A). As friction increases, the force required to provide the particle with excess energy increases, leading to minima at higher values of F.

We make the following observation which will be used in the subsequent sections.

Observation 2.3.2. We have found no evidence of multiple local minima of erasing time in a level set of work for our control family. Refer to Section 6.4 of Appendix A for characteristic plots showing the minima of erasing time in a level set of work. Physically this is unsurprising since the non-monotonicity in τ_e with γ and F mentioned above arise from fairly simple trade-offs, producing curves with single minima.



(a) $\gamma = 0.1$ **(b)** $\gamma = 1$

(c) $\gamma = 10$ **(d)** $\gamma = 100$

FIGURE 2.12: For a fixed value of W and γ, the erasing time is a non-monotonic function of F and is minimum at moderate F. This is illustrated at work $W = 20$ for various values of γ.

As with the reliability time, a more detailed analysis of the dependence of τ_e on other parameters, and even the shape of the control, is possible. However, these details are likely to be difficult to generalise, and are not necessary for the conclusions we draw in the subsequent sections.

2.4 Design of Bits

We are now ready to study the question of how to design good bits. A design involves choosing parameters A, F, γ for a bit to satisfy requirement specifications in terms of speed of erasing and reliability, without expending more work than required. The most general formulation of our problem would require us to also allow the length scale B, the temperature T and the mass m to vary, as well as allowing arbitrary controls. Such a formulation would appear to make the problem even more challenging, so it seems

prudent in a first analysis to restrict our analysis to the variables A, F, and γ. Our restricted analysis is not without value since the underlying technology in any given construction typically does not allow arbitrary variation. Our numerical analysis with Example 2.2.1 will guide us in our assumptions and analysis, but our results will hold in greater generality. We will construct our proofs based on general assumptions, and subsequently explain how these assumptions are met by our control family.

We introduce the following terms.

1. The design of a bit is completely specified by the **design triple** (A, F, γ). **Design Space** $(\mathcal{DS})$ is the space of all design triples (A, F, γ).

2. A **requirement specification** is a tuple $(t_r, t_e) \in \mathbb{R}^2_{>0}$ denoting the reliability and erasing time that we require of the bit. **Requirements Space** $(\mathcal{RS})$ is the space of all requirement specifications.

3. **Erasing time** $\tau_e : \mathcal{DS} \to \mathbb{R}_{>0}$ takes a design triple (A, F, γ) to the time required for erasing the corresponding bit under the control protocol specified by F. **Reliability time** $\tau_r : \mathcal{DS} \to \mathbb{R}_{>0}$ takes a design triple (A, F, γ) to the reliability time of the corresponding bit. Note that τ_r is constant as a function of F since it is a property of the dynamics in the absence of control.

4. **Work** $W : \mathcal{DS} \to \mathbb{R}_{>0}$ represents the expected work done by the control in erasing the corresponding bit. We will assume that W is constant as a function of γ, as is the case in Example 2.2.1.

5. A design (A, F, γ) is **feasible** for a requirement (t_r, t_e) iff both $\tau_r(A, F, \gamma) \geq t_r$ and $\tau_e(A, F, \gamma) \leq t_e$. A (t_r, t_e)-feasible design (A, F, γ) is (t_r, t_e)-**optimal** iff the work $W(A, F, \gamma)$ is minimum among all (t_r, t_e)-feasible designs.

6. Inspired by the observation that non-trivial minima of erasing time at fixed work exist for our family of protocols (Section 2.3.3), we define the notion of **trapped**

bits. A design (A, F, γ) is **trapped** iff for all designs (A', F', γ') with $W(A, F, \gamma) = W(A', F', \gamma')$, the erasing time $\tau_e(A, F, \gamma) \leq \tau_e(A', F', \gamma')$. A design (A, F, γ) is **uniquely trapped** iff for all designs (A', F', γ') with $W(A, F, \gamma) = W(A', F', \gamma')$, the erasing time $\tau_e(A, F, \gamma) \leq \tau_e(A', F', \gamma')$ with equality iff $(A, F, \gamma) = (A', F', \gamma')$. A design (A, F, γ) is **locally trapped** iff there exists a neighbourhood of (A, F, γ) consisting of bits (A', F', γ') with $W(A, F, \gamma) = W(A', F', \gamma')$ such that the erasing time $\tau_e(A, F, \gamma) \leq \tau_e(A', F', \gamma')$. More informally, a trapped design has the lowest erasing time within a level set of work; a trapped design is unique if it is the *only* design within that level set of work to have the minimal erasing time; and a locally trapped design has the minimal erasing time within a local neighbourhood of designs of equal work.

7. A requirement specification (t_r, t_e) is **unsaturated** iff there exists a (t_r, t_e)-optimal design (A, F, γ) such that either $\tau_r(A, F, \gamma) > t_r$ or $\tau_e(A, F, \gamma) < t_e$. A feasible requirement specification that is not unsaturated is called **saturated**.

Throughout this section, we will assume that τ_e, τ_r, and W are continuous functions. We will state and prove the main results related to the properties of the optimal design. We first claim that an optimal design always saturates the bound on the erasing time constraint. Further, if the optimal bit is not locally trapped, then it also saturates the bound on the reliability time constraint.

Claim 2.4.1 (Saturation of timescales)**.** Let us assume that it is possible to locally decrease work at fixed reliability time (This is generally possible since one can perturb the control parameters to reduce work; but reliability time does not depend on the control parameters). Fix requirement specifications $(t_r, t_e) \in \mathcal{RS}$. Suppose (A, F, γ) is a (t_r, t_e)-optimal design. Then

1. $\tau_e(A, F, \gamma) = t_e$.

2. If the design (A, F, γ) is not locally trapped, then $\tau_r(A, F, \gamma) = t_r$.

Proof. 1. For contradiction suppose that $\tau_e(A, F, \gamma) < t_e$. Since τ_e is continuous and it is possible to locally decrease work at fixed reliability time, there exists a design (A', F', γ') with $W(A', F', \gamma') < W(A, F, \gamma)$ that is (t_r, t_e)-feasible contradicting the optimality of the design (A, F, γ).

2. For contradiction suppose that $\tau_r(A, F, \gamma) > t_r$. Since the design (A, F, γ) is not locally trapped and τ_r is continuous, there exists a design (A_0, F_0, γ_0) requiring work $W(A_0, F_0, \gamma_0) = W(A, F, \gamma)$, erasing time $\tau_e(A_0, F_0, \gamma_0) < \tau_e(A, F, \gamma) \leq t_e$ and maintaining reliability time $\tau_r(A_0, F_0, \gamma_0) \geq t_r$. Thus the design (A_0, F_0, γ_0) is (t_r, t_e)-optimal contradicting Claim 2.4.1.1 that the optimal design saturates the bound on the erasing time constraint.

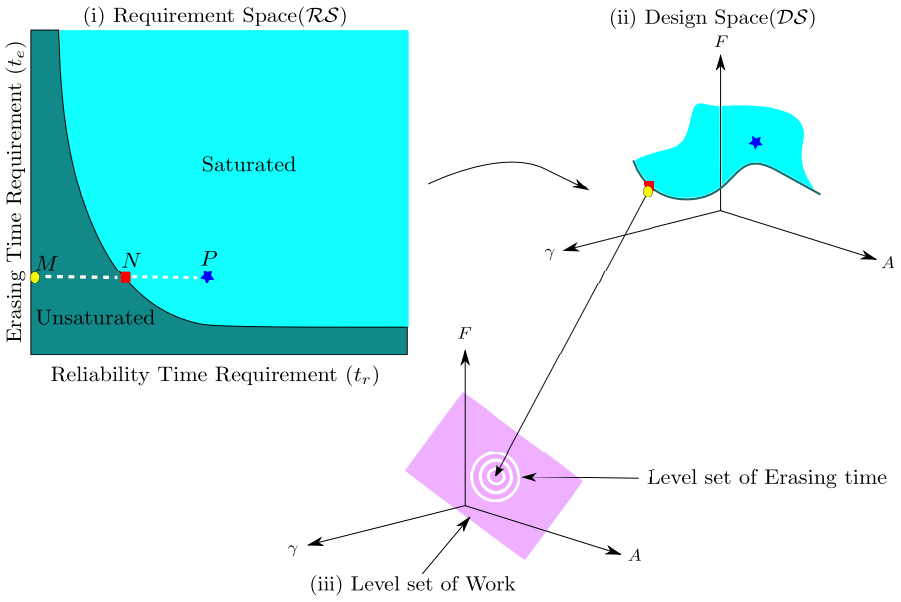


FIGURE 2.13: An illustration of the mapping of requirement specifications to optimal designs. The requirement space is divided by a curve corresponding to the reliability and erasing times of trapped designs. Points M, N in (i) Requirement Space($\mathcal{RS}$) having the same erasing time requirement get mapped to the same optimal bit in (ii) Design Space($\mathcal{DS}$): a trapped design with τ_r and τ_e equal to the requirements at N. The requirement specifications represented by points like P having the same t_e but greater t_r than N are mapped to distinct points in design space. (iii) A representation of a level set of work W within $\mathcal{DS}$, illustrating that the optimal designs to which unsaturated requirements are mapped minimize erasing time among all designs requiring the same work.

The next claim provides insight into the geometry of optimal designs. In particular, it states that under mild assumptions the requirement space is divided into two regions by a boundary given by the reliability and erasing times of trapped designs. Requirements with $t_r < t'_r$ and $t_e = t'_e$, where (t'_r, t'_e) is a requirement on the dividing line, are unsaturated, while other requirement specifications are saturated. To prove this claim, we first make the following observation that the erasing time of trapped designs is a strictly decreasing function of the work.

Observation 2.4.2. The erasing time of trapped designs is a strictly decreasing function of work for our family of protocols (Example 2.2.1). In other words, if (A_0, F_0, γ_0) and (A^*, F^*, γ^*) are trapped designs with $W(A^*, F^*, \gamma^*) > W(A_0, F_0, \gamma_0)$ then $\tau_e(A^*, F^*, \gamma^*) < \tau_e(A_0, F_0, \gamma_0)$.

Proof. Since W is a continuous increasing function of A (Observation 2.2.2), one can choose $A' > A_0$ such that $W(A', F_0, \gamma_0) = W(A^*, F^*, \gamma^*)$. Noting that increasing well height at fixed F and γ decreases erasing time (Observation 2.3.1), we get $\tau_e(A', F_0, \gamma_0) < \tau_e(A_0, F_0, \gamma_0)$. Using the fact that (A^*, F^*, γ^*) is a trapped design, we get $\tau_e(A^*, F^*, \gamma^*) \leq \tau_e(A', F_0, \gamma_0) < \tau_e(A_0, F_0, \gamma_0)$ establishing the claim.

Claim 2.4.3 (Saturated and Unsaturated Requirements)**.** Assume that the erasing time of trapped designs is a strictly decreasing function of the work by Observation 2.4.2 and that as before it is always possible to decrease work at fixed reliability time. Let (A^*, F^*, γ^*) be a trapped design such that $\tau_e(A^*, F^*, \gamma^*) = t_e$.

1. If $t_r \leq \tau_r(A^*, F^*, \gamma^*)$ then (A^*, F^*, γ^*) is (t_r, t_e)-optimal.

2. If $t_r < \tau_r(A^*, F^*, \gamma^*)$ then (t_r, t_e) is unsaturated.

3. Make the additional assumption that locally trapped designs are uniquely trapped (as noted for our family of protocols (Example 2.2.1) in Observation 2.3.2).

   If $t_r \geq \tau_r(A^*, F^*, \gamma^*)$, then (t_r, t_e) is saturated.

Proof. 1. Since $\tau_r(A^*, F^*, \gamma^*) \geq t_r$ and $\tau_e(A^*, F^*, \gamma^*) = t_e$, the design (A^*, F^*, γ^*) is (t_r, t_e)-feasible. Suppose that the design (A^*, F^*, γ^*) is not (t_r, t_e)-optimal. Then there exists a (t_r, t_e)-feasible design (A', F', γ') such that $W(A', F', \gamma') < W(A^*, F^*, \gamma^*)$. Let (A_0, F_0, γ_0) be a trapped design with $W(A_0, F_0, \gamma_0) = W(A', F', \gamma') < W(A^*, F^*, \gamma^*)$. Then $\tau_e(A_0, F_0, \gamma_0) > \tau_e(A^*, F^*, \gamma^*)$ since the erasing time of trapped bits is a strictly decreasing function of work. Using the fact that (A_0, F_0, γ_0) is a trapped design, we get $\tau_e(A', F', \gamma') \geq \tau_e(A_0, F_0, \gamma_0) > \tau_e(A^*, F^*, \gamma^*) = t_e$, a contradiction since (A', F', γ') is a (t_r, t_e)-feasible design.

2. Immediate from Claim 2.4.3. 1.

3. For contradiction suppose that the requirement (t_r, t_e) is unsaturated. Then by Claim 2.4.1. 1 there exists a (t_r, t_e)-optimal design (A_0, F_0, γ_0) such that $\tau_r(A_0, F_0, \gamma_0) > t_r$ and $\tau_e(A_0, F_0, \gamma_0) = \tau_e(A^*, F^*, \gamma^*) = t_e$. Since locally trapped designs are uniquely trapped, using Claim 2.4.1. 2, we get that the design (A_0, F_0, γ_0) must be uniquely trapped. Noting that uniquely trapped bits are trapped and using the fact that the erasing time of trapped designs is a strictly decreasing function of work, we get $W(A_0, F_0, \gamma_0) = W(A^*, F^*, \gamma^*)$. This implies that $(A_0, F_0, \gamma_0) = (A^*, F^*, \gamma^*)$, a contradiction since $\tau_r(A_0, F_0, \gamma_0) > t_r \geq \tau_r(A^*, F^*, \gamma^*)$.

The claims about saturation/unsaturation of times-scales can also be proved using KKT conditions(a standard tool from optimization theory [133, 134]), as we show below.

2.5 KKT conditions

Consider the optimization problem of finding the design with the lowest work that is (t_r, t_e)-feasible.

Problem 2.5.1.

$$(A^*, F^*, \gamma^*) = \arg \inf_{A,F,\gamma} W(A, F)$$
$$t_r - \tau_r(A^*, \gamma^*) \leq 0$$
$$\tau_e(A^*, F^*, \gamma^*) - t_e \leq 0$$

In order to state the KKT conditions, we will need the notion of a regular point. The following definition will make this precise.

Definition 2.5.2 (**Regular point**)**.** Let $Sat(x^*)$ denote the set of gradients of the constraints that are saturated at the point x^*. Then x^* is regular iff $Sat(x^*)$ does not form a linearly dependent set.

Theorem 2.5.3 (**KKT conditions**)**.** *Let (A^*, F^*, γ^*) be a local optimum of 2.5.1 and a regular point. Then by [133, (12.1), pp. 95], there exists $\lambda_1^*, \lambda_2^* \in \mathbb{R}_{\geq 0}$ such that*

1. $\nabla W(A^*, F^*, \gamma^*) - \lambda_1^* \nabla \tau_r(A^*, \gamma^*) + \lambda_2^* \nabla \tau_e(A^*, F^*, \gamma^*) = 0.$
2. $\lambda_1^*(t_r - \tau_r(A^*, \gamma^*)) = 0$ *and* $\lambda_2^*(\tau_e(A^*, F^*, \gamma^*) - t_e) = 0.$

Given this powerful theorem 2.5.3, we are now ready to prove the the same result that we obtained earlier but with the machinery of KKT conditions.

Lemma 2.5.4. Let us assume that it is always possible to locally decrease work at fixed reliability time. Let (A^*, F^*, γ^*) be a local optimum of 2.5.1. Then either

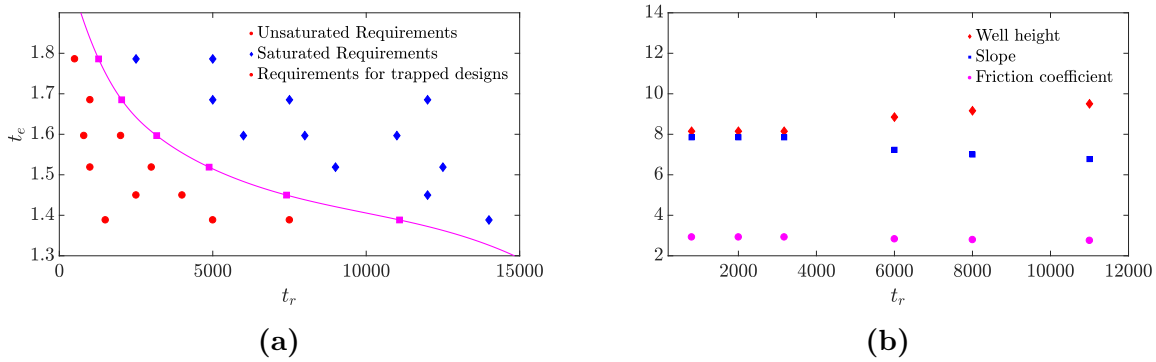
1. The design (A^*, F^*, γ^*) saturates the bound on both constraints i.e. $\tau_r(A^*, \gamma^*) = t_r$ and $\tau_e(A^*, F^*, \gamma^*) = t_e$ or
2. The design (A^*, F^*, γ^*) saturates the bound on the erasing time constraint i.e. $\tau_e(A^*, F^*, \gamma^*) = t_e$ but does not saturate the bound on the reliability time constraint i.e. $\tau_r(A^*, \gamma^*) > t_r$ and is locally trapped.

Proof. Consider an optimal design (A^*, F^*, γ^*) such that either it does not saturate the bound on the reliability time constraint i.e. $\tau_r(A^*, \gamma^*) > t_r$ or it does not saturate the bound on the erasing time constraint i.e. $\tau_e(A^*, F^*, \gamma^*) < t_e$. Then we have the following cases:

- Case 1: The design (A^*, F^*, γ^*) saturates the bound on the erasing time constraint, but does not saturate the bound on the reliability time constraint i.e. $\tau_r(A^*, \gamma^*) > t_r$ and $\tau_e(A^*, F^*, \gamma^*) = t_e$. This implies that $\lambda_1^* = 0$. Since only one constraint is active, (A^*, F^*, γ^*) is a regular point. Hence, by Theorem 2.5.3 on KKT conditions, there exists $\lambda_2^* > 0$ such that $\nabla W(A^*, F^*, \gamma^*) + \lambda_2^* \nabla t_e(A^*, F^*, \gamma^*) = 0$. This implies that (A^*, F^*, γ^*) is a stationary point of erasing time in the level set of it's work $W(A^*, F^*, \gamma^*)$. The fact that this stationary point is actually a local minimum follows from Claim 2.4.1. 2.
- Case 2: The design (A^*, F^*, γ^*) saturates the bound on the reliability time constraint, but does not saturate the bound on the erasing time constraint i.e. $\tau_r(A^*, \gamma^*) = t_r$ and $\tau_e(A^*, F^*, \gamma^*) < t_e$. This implies that $\lambda_2^* = 0$. Since only one constraint is active, (A^*, F^*, γ^*) is a regular point. Hence, by Theorem 2.5.3 on KKT conditions, there exists $\lambda_1^* > 0$ such that $\nabla W(A^*, F^*, \gamma^*) = \lambda_1^* \nabla t_r(A^*, \gamma^*)$, a contradiction since $\frac{\partial W}{\partial F} \neq 0$ but $\frac{\partial \tau_r}{\partial F} = 0$.
- Case 3: The design (A^*, F^*, γ^*) does not saturate the bound on either constraints i.e. $\tau_r(A^*, \gamma^*) > t_r$ and $\tau_e(A^*, F^*, \gamma^*) < t_e$. Since no constraint is active we have $\nabla W(A^*, F^*, \gamma^*) = 0$, which is not possible.

A more intuitive picture of the results can be understood from Figure 2.13. In this figure, we illustrate how finding an optimal design subject to a specification maps a point in the requirement space to a point in the design space. For a trapped design (A^*, F^*, γ^*), requirements with $t_r < \tau_r(A^*, F^*, \gamma^*)$ and $t_e = \tau_e(A^*, F^*, \gamma^*)$ are unsaturated

and get mapped to the same design (A^*, F^*, γ^*) (claims 2.4.3. 2 and 2.4.3. 1). If the design (A^*, F^*, γ^*) is uniquely trapped, then requirements with $t_r \geq \tau_r(A^*, F^*, \gamma^*)$ and $t_e = \tau_e(A^*, F^*, \gamma^*)$ are saturated (claim 2.4.3. 3).



(a) **(b)**

FIGURE 2.14: Illustration of the division of Requirement Space($\mathcal{RS}$) into saturated and unsaturated regions by requirements that correspond to trapped designs. (a) Squares show requirements (t_e, t_r) that are saturated by trapped designs for the family of protocols we consider. Numerical optimization shows that requirements to the left of the locus defined by these points are unsaturated (circles), whereas requirements to the right are saturated (diamonds). (b) A plot of the optimal designs for points from (a) at $t_e = 1.5967$. It is clear that for requirements $t_r \leq 3173$, lying to the left of the trapped-design locus in (a), optimal design parameters are identical whereas they are distinct for $t_r > 3173$.

Figure 2.14 illustrates these results for our example family of controls (Example 2.2.1). As discussed in Section 6.3 of Appendix A, we have implemented simple regression to fit the functions $\tau_e(.)$ and $\tau_r(.)$ to our simulation results. We then identified trapped designs using numerical minimisation, plotting the requirement specifications saturated by these designs. For each trapped bit (A^*, F^*, γ^*), we randomly selected requirements with $t_e = \tau_e(A^*, F^*, \gamma^*)$, but with t_r either greater than equal to or less than $\tau_r(A^*, F^*, \gamma^*)$, and used numerical optimization techniques to search for the optimal designs. The results support our analysis; requirements with $t_r < \tau_r(A^*, F^*, \gamma^*)$ are unsaturated, and those with $t_r \geq \tau_r(A^*, F^*, \gamma^*)$ are saturated. Furthermore, as we show in Figure 2.14 (b), unsaturated requirements at fixed t_e all map to the same trapped design.

2.5.1 Optimal friction for simple controls

In Section 2.3, we demonstrated that both reliability and erasing times are non-monotonic in friction, with short erasing times favoured by moderate values of friction, and long reliability times favoured by extreme values. In what follows, we give a precise quantification of the resultant trade-off in finding the friction of an optimal bit. The analysis is significantly simplified for our family of controls, in which work $W \approx \frac{\left(A+F\cdot B-\frac{k_B T}{2}\right)}{2}$ is independent of the friction coefficient. Let us introduce the following terms. Fix an A and F. Then,

1. γ^e_{crit} is the friction coefficient that minimizes erasing time as a function of friction coefficient γ at fixed A and F, i.e., for all $\gamma' \in \mathbb{R}_{>0}$, we have:

$$\tau_e(A, F, \gamma^e_{\text{crit}}) \leq \tau_e(A, F, \gamma'). \tag{2.34}$$

   We call the design $(A, F, \gamma^e_{\text{crit}})$ **critically damped**.

2. γ^r_{crit} is the friction coefficient that minimizes reliability time as a function of friction coefficient γ at fixed A and F, i.e., for all $\gamma' \in \mathbb{R}_{>0}$, we have:

$$\tau_r(A, F, \gamma^r_{\text{crit}}) \leq \tau_r(A, F, \gamma'). \tag{2.35}$$

It is easy to note that trapped bits are also critically damped. In Figure 2.15 we show illustrative curves of the erasing and reliability times as a function of friction coefficient γ at fixed A, F. These curves have single minima at γ^e_{crit} and γ^r_{crit}, respectively. Also shown on this graphs are regions of friction space that can be eliminated from consideration for optimal bits. To eliminate extreme values of friction, we note that the design must have a minimal finite A to be a well-defined two-state system in the resting state. For our bit, it is $A_{\min} \approx 3$. In the next claim we precisely describe which regions of friction can be eliminated.

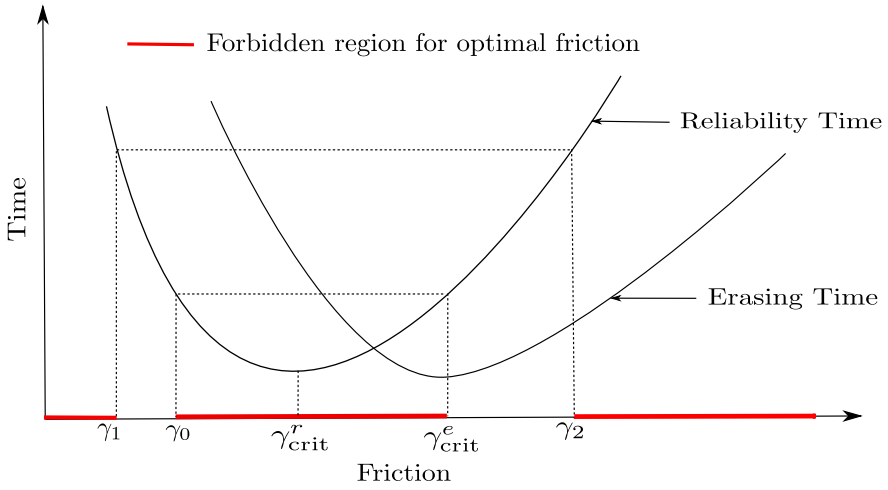


FIGURE 2.15: Regions of friction-space can be eliminated from the search for optimal bits for our class of controls. As a result, the optimal friction is either critical damping, or lies somewhere within two regions of moderate friction. Illustrative curves of τ_e and τ_r at fixed A, F indicate these regions.

Claim 2.5.5. [Forbidden regions for optimal friction] Assume that both τ_e and τ_r have a single-well-defined minimum and tend to infinity as γ tends to zero or infinity. Let (A, F, γ) be a (t_r, t_e)-optimal design. (Refer to Figure 2.15 for notational convenience)

1. Let γ_0 be such that $\tau_r(A, F, \gamma_0) = \tau_r(A, F, \gamma^e_{\text{crit}})$.

   (a) If $\gamma^e_{\text{crit}} > \gamma^r_{\text{crit}}$, then $\gamma \notin (\gamma_0, \gamma^e_{\text{crit}})$.

   (b) If $\gamma^e_{\text{crit}} < \gamma^r_{\text{crit}}$, then $\gamma \notin (\gamma^e_{\text{crit}}, \gamma_0)$.

   i.e. the friction of the optimal bit does not reside in the central red region in Figure 2.15.

2. Let $A_{\min}$ be the minimum height for a bit to be reliable and let $\gamma_1 < \gamma_2$ be such that $\tau_r(A_{\min}, F, \gamma_1) = \tau_r(A_{\min}, F, \gamma_2) = t_r$. If (A, F, γ) is not locally trapped, then $\gamma \notin (0, \gamma_1) \cup (\gamma_2, \infty)$ i.e. the friction of the optimal bit does not arise from the extreme red regions in Figure 2.15.

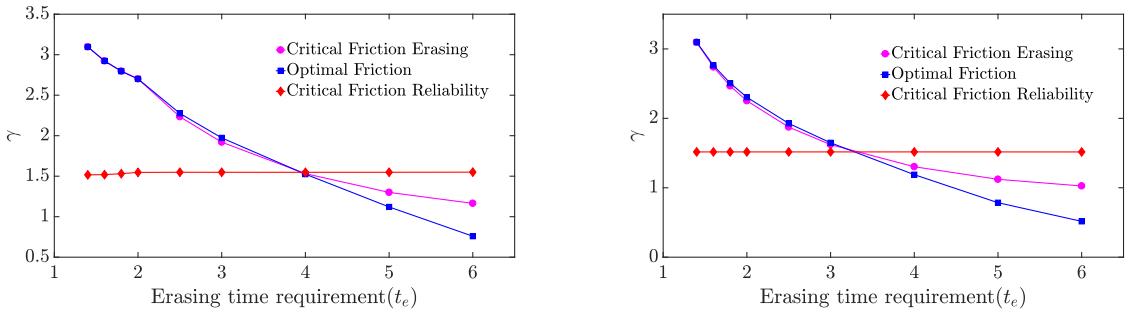
Proof. 1. We prove it for the case when $\gamma^e_{\text{crit}} > \gamma^r_{\text{crit}}$, the other case proceeds in identical fashion. For contradiction, assume that $\gamma \in (\gamma_0, \gamma^e_{\text{crit}})$. Then, due to the single

minima in both τ_e and τ_r, and the fact that τ_r tends to infinity as γ tends to zero or infinity, there exists a design (A, F, γ') with $\gamma' > \gamma_0$ and $\tau_r(A, F, \gamma') = \tau_r(A, F, \gamma) \geq t_r$, but $\tau_e(A, F, \gamma') < \tau_e(A, F, \gamma) \leq t_e$. The design (A, F, γ') is (t_r, t_e)-optimal since it is (t_r, t_e)-feasible and has $W(A, F, \gamma') = W(A, F, \gamma)$, contradicting Lemma 2.4.1. 1 that the optimal bit saturates the bound on the erasing time constraint.

2. For contradiction, suppose that $\gamma < \gamma_1$ or $\gamma > \gamma_2$. Then since $A \geq A_{\min}$ and the reliability time increases with well height and more extreme values of γ, either $\tau_r(A, F, \gamma) \geq \tau_r(A_{\min}, F, \gamma) > \tau_r(A_{\min}, F, \gamma_1) = t_r$ or $\tau_r(A, F, \gamma) \geq \tau_r(A_{\min}, F, \gamma) > \tau_r(A_{\min}, F, \gamma_2) = t_r$, contradicting claim 2.4.1. 2 that an optimal design that is not locally trapped saturates the bound on the reliability time constraint.

For clarity, let us assume initially that $\gamma^e_{\text{crit}} > \gamma^r_{\text{crit}}$ (equivalent arguments hold for the alternative). We see that optimal designs reside either at γ^e_{crit}, or lie within two regions at moderate friction, as illustrated in Figure 2.15. Interestingly, one region is adjacent to γ^e_{crit}, whereas the other is not. It is not easy to see how designs in one region ($\gamma_1 \leq \gamma \leq \gamma_0$) as in Figure 2.15 can outperform those in the other region ($\gamma^e_{\text{crit}} < \gamma \leq \gamma_2$). Indeed, when we performed numerical optimisation on the regression-based fits to our simulation data, we only observed optimal bits that are either critically damped or lie in the allowed region adjacent to critical damping.

This is illustrated in Figure 2.16, where we plot the optimal friction as a function of erasing time requirement at fixed reliability time requirement, for two values of reliability time requirements. We also plot γ^e_{crit} and γ^r_{crit} for comparison. At low erasing time requirements, designs reside at γ^e_{crit}. At slightly higher erasing time requirements, the designs become saturated and the optimal friction lies adjacent to γ^e_{crit} in the region $\gamma^e_{\text{crit}} < \gamma \leq \gamma_2$. Eventually, γ^e_{crit} crosses γ^r_{crit}. At the crossing point, we have $\gamma = \gamma^e_{\text{crit}} = \gamma^r_{\text{crit}}$. At higher values of erasing time requirements, γ still occupies the region adjacent



(a) Reliability time requirement$(t_r) = 500$ **(b)** Reliability time requirement$(t_r) = 10000$

FIGURE 2.16: Optimal friction is either critical damping, or lies within a small region adjacent to critical damping, for our family of controls. We plot friction for optimal designs (A, F, γ) against erasing time requirements (t_e) for a fixed value of reliability time requirement (t_r), alongside γ^e_{crit} and γ^r_{crit}. Note that A and F are not fixed, but determined by the optimisation procedure alongside the optimal friction for each requirement (t_r, t_e). The data was obtained from numerical optimisation and minimisation based on regression fits to simulation data.

to γ^e_{crit}, which is now $\gamma_1 \leq \gamma \leq \gamma^e_{\text{crit}} < \gamma^r_{\text{crit}}$.

2.6 Conclusions

We have explored the question of the design of optimal bits. Previously, authors [135, 136, 137, 34, 138] have focused on designing optimal protocols that minimize work input when implementing a finite-time operation on a given system, as highlighted in the introduction. Our approach differs in considering that bits need to have two distinct functionalities: retain data for long periods of time and allow rapid switching or erasing. Moreover we consider optimising over system parameters such as the intrinsic friction as well as the external control. Our fundamental observation is that friction plays a non-trivial role in the design of bits. Both switching/erasing and the eventual degradation of data involve relaxation towards equilibrium from a non-equilibrium distribution. This process is fastest at intermediate values of the friction, but slow in the overdamped and underdamped regimes. The best bit designs have high reliability times and low switching/erasing times, which implies an inherent trade-off in bit design between extreme values of friction that favour high reliability, and moderate values of friction that favour

rapid switching or erasing.

We have explored the consequences of the biphasic role of friction for a simple class of controls. The existence of non-trivial minima of erasing time in the level set of work leads to the generation of trapped designs. These designs are optimal for reliability requirements smaller than their own reliability time leading to unsaturated requirements. The result of the trade-off between extreme values of friction that maximize reliability time and moderate values of friction that minimise erasing times is that optimal designs are either critically damped or occupy a region of moderate friction close to critical damping.

Our work opens up a new perspective on the design of efficient computational devices showing that: *the best designs are likely to be neither underdamped nor overdamped.* This observation is particularly important as some authors have considered friction to be inherently problematic for computation. The role of friction is suppressed when bits are modelled as discrete two-state systems [9, 100, 139], since this approximation assumes rapid equilibration within the discrete states. Friction also has links back to the IBM project [109, 140] that was initiated in the 1970's, which used Josephson junctions (consisting of two superconducting materials segregated by an insulating layer) to build logical devices. Current flows through this junction due to the tunnelling of electrons across the insulator at virtually no resistance (friction). There is a critical value of current above which the material stops acting as a superconductor. This "switching" of junction state from superconducting to non-superconducting is analogous to the switching of a bit in our case. In fact, authors [110, 111] compare the switching in Josephson circuits to Kramer's escape rate problem in the underdamped regime. Further, Klein and Mukherjee [111] compare the performance of noise switching rates among several multi-junction SQUIDS(Superconducting Quantum Interference Devices).

The main issue for building devices out of Josephson junctions came in the construction of fast memory chips. Firstly, the cycle time of these devices was restricted by

"punchthrough"- a phenomenon in which the voltage of the device could not be reset after its cycle. Secondly, the value of critical current was lower bounded due to the presence of thermal noise that caused unwarranted switching, analogous to the noise-induced switching of a particle in a bistable well. Thirdly, semi-conductor technology was not that far behind, and the current super-conductor project would take atleast a couple of years to be put into practice. As a consequence, the IBM project was finally abandoned [141] in 1983.

It is important to note the we have only considered a simple family of controls to motivate our analysis and illustrate our findings. This family is not optimal - it was chosen for it's simplicity and ease of analysis. Moreover, there is some arbitrariness in the definition of both the erasing and reliability times. We are not claiming to have derived numerical corrections to the minimal cost of erasing a bit, for example, or the specific work costs (substantially larger than $k_B T \ln 2$) which are not that informative. Rather, it is the qualitative results, which hold for a much broader class of controls which are important. The non-monotonic role of friction in both the erasing and reliability time-scales is a generic physical phenomenon that extends beyond the details of our implementation, and implies a competition between the goals of fast manipulation and long reliability times. Relatively weak assumptions – that it is always possible to decrease work at fixed reliability time and that the minimal erasing time decreases with increased work imply that erasing time requirements are always saturated by optimal bits and that trapped designs lead to unsaturated reliability time requirements respectively. Other results rely more on the simplicity of the control family: the existence of only one local minimum of erasing time at fixed work simplifies the question of whether a requirement specification is saturated. The fact that work is independent of friction simplifies the task of eliminating certain values of friction as sub-optimal.

Explicit exploration of a broader class of controls, including those with more complex variation over time, and varying parameters such as particle mass and distance between

wells, are possible directions for future work. It is not immediately clear whether minima in erasing time at fixed work cost will become more or less prominent features of the optimisation landscape when the complexity of the system is increased, for example. In particular, raising or lowering the barrier between metastable states is a common idea [100, 34, 117, 118]. Lowering the barrier during erasing potentially allows for faster erasing at fixed reliability time and lower work cost. If said barriers could be raised and lowered arbitrarily far and quickly, it may be possible to circumvent any conflict between high reliability and low erasure time. However, real physical systems are not generally this flexible. Indeed, in order to apply a complex time-dependent control to a small colloid, experimenters typically use optical feedback traps [117, 118], which are not true potentials and rely on the continuous input of energy to apply forces and perform feedback control. For true physical protocols that permit finite raising and lowering of barriers between metastable states, we expect that our findings would still apply to a family of protocols with optimal barrier manipulation. An alternative direction would be to consider similar effects in systems with inherently quantum mechanical behaviour.

The ideas of reliability, speed and cost have been studied in the context of bit-erasure for a two-state Markov chain [139]. Here, the authors come up with a cost of $\log 2 - \log 1 - e^{\frac{\tau_e}{\tau_r}}$ for fast and reliable bit-erasure. Our work differs from theirs in many ways. At a fundamental level, we are essentially interested in questions that involve a relaxation time –a notion that does not carry over directly to the discrete state limit. In addition, the cost function used by the authors in [139] is a Kullback-Leibler divergence – which is not a typical work function. This makes it hard to map the results obtained in our setting to the two-state Markov chain problem.

It is important to observe that a significant fraction of our results arise as a manifestation of going beyond the two-state model of a bit. In particular, the crucial role played by friction resulting in a trade-off between high reliability and low erasing time-scales emerges as a consequence of the presence of a continuum of intermediate states separating the

two states of the bit. The notion of relaxing to equilibrium within a state (as defined in erasing time) becomes relevant precisely due to the same reason. In addition, our discussion about careful consideration of control strategies draws its pertinence from exploring bit manipulations beyond its two-state canonical model.

Chapter 3

High rates of fuel consumption are not required by insulating motifs to suppress retroactivity in biochemical circuits

Abstract

Retroactivity arises when the coupling of a molecular network $\mathcal{U}$ to a downstream network $\mathcal{D}$ results in signal propagation back from $\mathcal{D}$ to $\mathcal{U}$. The phenomenon represents a breakdown in modularity of biochemical circuits and hampers the rational design of complex functional networks. Considering simple models of signal-transduction architectures, we demonstrate the strong dependence of retroactivity on the properties of the upstream system, and explore the cost and efficacy of fuel-consuming insulating motifs that can mitigate retroactive effects. We find that simple insulating motifs can suppress retroactivity at a low fuel cost by coupling only weakly to the upstream system $\mathcal{U}$. However, this design approach reduces the signalling network's robustness to perturbations from leak reactions, and potentially compromises its ability to respond to rapidly-varying signals.

3.1 Introduction

In the second section of the introduction, we discussed and outlined some of the mechanisms by which biological systems perform computation. Performing complex biological computations requires the use of multiple connected components or pathways, as in electrical circuits. The possibility of designing electrical circuits to function as well-defined

modules is crucial to the engineering of complex circuitry with many interconnected components. Ideal modules have clearly-defined inputs, and produce outputs specified by those inputs and the internal structure of the module, independent of the broader context within which they are embedded [142].

Biological systems appear to exhibit modularity in some contexts, and it has been suggested that this modularity is evolutionarily advantageous [143, 86, 85], or contributes to system robustness [89, 144]. For humans seeking to engineer biological systems without the benefit of billions of years of highly parallel evolution, modularity can be hugely advantageous. In the context of biochemical networks, the concept of modularity was first put

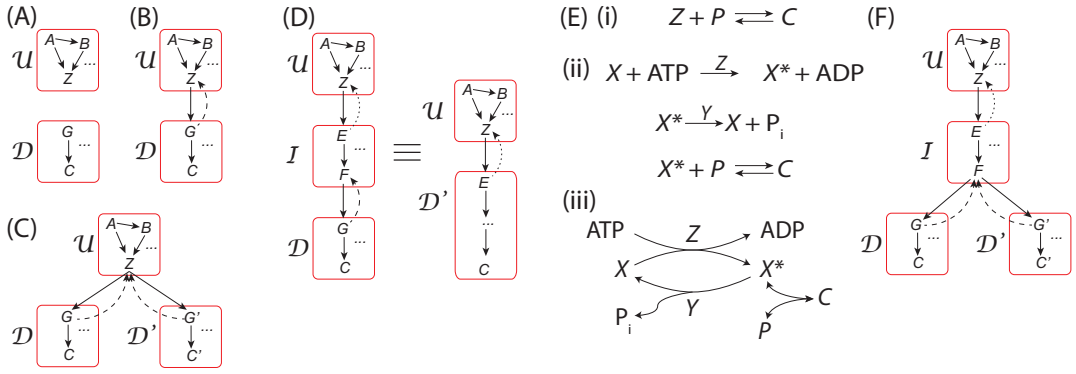


FIGURE 3.1: A schematic representation of the concept of retroactivity. (A) Subsystems $\mathcal{U}$ and $\mathcal{D}$ evolve separately according to internal dynamics (the arrows in these diagrams indicate arbitrary interactions). (B) By coupling $\mathcal{U}$ and $\mathcal{D}$ via a molecular interaction between Z (an output of $\mathcal{U}$) and G (an input of $\mathcal{D}$), a signal is propagated. However, in general the coupling also induces changes in Z due to retroactivity (dashed line). (C) This retroactivity is particularly problematic when adding an additional downstream system $\mathcal{D}'$ that also couples to $\mathcal{U}$. (D) Retroactivity can potentially be reduced using an insulator $\mathcal{I}$ between $\mathcal{U}$ and $\mathcal{D}$. In effect, $\mathcal{U}$ couples to a compound downstream motif $\mathcal{D}'$. (E) A specific example of an insulating circuit. Retroactivty is very high for propagation of a signal by direct binding, as in (E.i). Alternatively, Z can act as a catalyst for the phosphorylation of an intermediate species, $X \rightarrow X^*$, that when phosphorylated binds to P, thereby reducing the retroactivity experienced by Z, as in (E.ii). Note that this insulating "push-pull" motif requires turnover of ATP to function, and involves an antagonistic phosphatase Y. (E.iii) A graphical representation of the reactions in (ii). (F) If the insulating circuit $\mathcal{I}$ itself couples to multiple downstream subsystems, retroactivity experienced by $\mathcal{I}$ may be relevant.

on a solid footing by Hartwell [145], Lauffenburger [146] and Weiss et.al. [147, 148]. The phenomenon of "retroactivity" [149, 150, 151, 152, 93], illustrated formally in Fig. 3.1 A,B has been shown to cause a breakdown in modularity. Here, an upstream system $\mathcal{U}$ consisting of a set of molecular species and reactions is coupled to a downstream system $\mathcal{D}$ via a species Z, which is part of $\mathcal{U}$. The coupling causes a change in the output of $\mathcal{D}$

(here represented by a species C), passing on a signal from $\mathcal{U}$ to $\mathcal{D}$. However, in general $\mathcal{U}$ and in particular Z is also affected by the coupling, implying the propagation of an unintended signal back from $\mathcal{D}$ to $\mathcal{U}$. Thus the meaning of the basic concepts of "upstream" and "downstream" is corrupted, and the ability to logically design circuits with well-defined inputs and outputs is compromised.

The presence of retroactivity is particularly problematic when coupling to a subsystem $\mathcal{U}$ with pre-existing connections (Fig. 3.1C) due to the fan-out effect [153] - an effect which causes signal reduction/delay due to the presence of excess downstream connections than required. Such a situation could arise from a change in network topology from human intervention or natural evolution, or due to dynamic changes in molecular abundance within a cell. In these circumstances, strong retroactivity would lead to the new "downstream" subsystem having an undesirable influence on the other "downstream" subsystems, and vice-versa.

Having identified the possibility of undesired retroactive interactions, several questions present themselves. (i) Most immediately, how should retroactivity be quantified? (ii) Given a suitable metric, are certain designs of $\mathcal{U}$ and $\mathcal{D}$ more prone to retroactive effects? (iii) Is it possible to design insulating motifs $\mathcal{I}$, as illustrated in Fig. 3.1D, that suppress retroactivity between $\mathcal{U}$ and $\mathcal{D}$? (iv) Does suppression of retroactivity necessarily imply an increased fuel consumption, and are there trade-offs associated with, for example, the accuracy of signal propagation?

In the last decade, several groups have considered these questions. In particular, Del Vecchio et.al. [151, 152, 154] have proposed the relative change in Z due to the introduction of $\mathcal{D}$ as a potential metric for retroactivity. Later, Barton and Sontag [93, 155] proposed two alternative metrics to quantify retroactivity, namely the distortion and competition effect. The distortion captures the change in C relative to an idealised system with no retroactive effect, while the competition effect quantifies the consequence for an existing downstream subsystem when a new one is attached to U. Particular attention

has been paid to retroactivity in the context of the binding of transcription factors to DNA [151, 152, 156, 93, 154, 50]. Certainly, the passing of signals via binding is naturally retroactive, since it intrinsically requires sequestration of the upstream molecule. Moreover, signal propagation by direct binding occurs in contexts other than transcription factor binding [157], and is widely used as a way to transmit signals in engineered nucleic acid systems both *in vitro* and *in vivo* [56, 158].

Sub-networks described as "insulators" ($\mathcal{I}$) have been proposed to mitigate retroactivity by connecting $\mathcal{D}$ to $\mathcal{U}$ indirectly (see Fig. 3.1D) [151, 93, 154, 155]. The key component underlying these insulators is catalysis, as described in the Introduction. Put simply, by acting as a catalyst molecule, the Z is able to influence the downstream reactions without being sequestered indefinitely [159, 80].

The push-pull motif illustrated in Fig. 3.1E.ii is a common catalytic motif in natural signalling systems [160] and is also known by the name of "futile" cycle in literature [161]. The molecule X is catalytically switched between its two states X and X^* by the antagonistic enzymes Z and Y (respectively, a kinase and a phosphatase if the modification of X is phosphorylation, as in Fig. 3.1E.ii). This mechanism is cognate to the notion of switching a bit between two states given by X and X^*, as stated in the Introduction. The output of X^* is then sensitive to the relative concentrations of Z and Y. By this mechanism, a signal encoded in the concentration of Z can be propagated without permanent binding of Z to a downstream substrate. Previous work has shown that such a push-pull motif can function as an effective insulator between $\mathcal{U}$ and $\mathcal{D}$ [151, 93], allowing information in the concentration of Z to be propagated via X^* to a downstream system with only limited sequestration of Z.

Catalysts cannot alter the equilibrium point of a reaction. Thus if $[X^*]$ is to be sensitive to $[Z]$, the system must be driven out of equilibrium by the turnover of biochemical fuel molecules [159]. In the case of phosphorylation-based signalling, the system is driven out of equilibrium by the coupling of phosphorylation/dephosphorylation cycles to the

breakdown of ATP into ADP and inorganic phosphate P_i, as shown in Fig. 3.1E.ii. The ATP molecules, which have a high free energy, are the chemical fuel. On a fundamental level, this fuel consumption (breakdown of ATP) allows the X molecules to "remember" the fact that they interacted with either Y or Z most recently, even though the interaction has ended [162, 80].

For the catalytic reactions in Fig. 3.1E.ii to proceed, complexes between Z and X must exist for a finite time [163]. Therefore, although the push-pull insulator can reduce retroactivity, some of the Z molecules are still sequestered at any given point and so some retroactivity remains. Thus, the presence of an intermediate state (catalyst-substrate complex) imparts retroactivity to the motif – yet another example where going beyond the two-state model of bit switching gives rise to interesting behaviour. Barton and Sontag [93] explored the question of whether this residual retroactivity could be suppressed, concluding that substantial energy consumption in the form of a high turnover of chemical fuel molecules was required. In a subsequent paper [155], they considered a slightly modified direct binding system in which the insulator itself acts catalytically on the downstream system $\mathcal{D}$, reaching the same conclusion.

In this chapter, we revisit the resource costs of retroactivity suppression, as first highlighted in Ref [93], by considering the simplest steady-state setting. In Section 3.3, we consider whether the design of the upstream system $\mathcal{U}$ can mitigate retroactive effects when long-lived binding of Z is necessary for signal propagation. We observe that a constant turnover of Z due to continuous production and decay can itself mitigate retroactivity in certain circumstances. However, this turnover is associated with a large resource cost if it is to be more rapid than the time-scale of signal variation.

As a result, we turn our attention in Section 3.5 to the analysis of insulating push-pull motifs that can potentially reduce retroactivity with limited protein production costs. Our main finding is that a higher rate of fuel consumption is not required to produce better insulators. In general, both fuel consumption and retroactivity can be reduced simply by

decreasing the coupling strength of both the upstream molecule Z and the phosphatase Y shown in Fig. 3.1E.ii to the X/X^* molecule, whilst maintaining the steady-state output of the system. In Section 3.6, we generalise this result to account for microscopic reversibility in the catalytic reactions that was neglected in Ref. [93]. We find that a large chemical driving force (a large free energy stored in each ATP molecule) is necessary to propagate strong signals, but not to suppress retroactivity, and that it is still possible to reduce both retroactivity and energy consumption to low levels by reducing the coupling of Z and Y to the push-pull network.

Although high free-energy consumption is not necessary to suppress retroactivity, we postulate that a certain amount is indeed important for faithful signal transduction. In particular, in Section 3.7 we show that weak coupling of Z and Y to the push-pull network renders the system as a whole vulnerable to unintended leak reactions. We also hypothesize that high turnover of fuel molecules is necessary to accurately track time-dependent inputs to $\mathcal{U}$.

3.2 Methods

We will work in the limit of a large copy number of molecules, in which case the reaction networks can be modelled deterministically by mass-action ordinary differential equations. All our calculations related to retroactivity and energy consumption will be performed after the decay of initial transients, as in previous works [151, 93, 164, 90, 165]. Previously, some authors have focussed on systems driven by time-dependent variation of parameters within $\mathcal{U}$, such as sinusoidally varying birth and death rates of Z [151, 93]. In particular, Barton and Sontag [93] analyse a certain network called the push-pull motif having a relatively rapid variation in the upstream signal. They define two metrics to quantify the amount of retroactivity in this system, namely the distortion and competition effect. The distortion captures the difference between the actual output $[C_{\text{real}}(t)]$ and the output

$[C_{\text{ideal}}(t)]$ of a hypothetical system in which the downstream system responds to Z as if binding were occurring, but the population of Z is unaffected by these reactions (and thus there is no retroactivity). The distortion metric is given by

$$\mathscr{D} = \frac{1}{\sigma_{[C_{\text{ideal}}]}} \langle |[C_{\text{ideal}}(t)] - [C_{\text{real}}(t)]| \rangle \,. \tag{3.1}$$

Here, $\sigma_{[C_{\text{ideal}}]}$ is the standard deviation of the ideal signal corresponding to the hypothetical system, and the angled brackets indicate an average over time. They also define the competition metric $\mathcal{C}$ as

$$\mathscr{C} = \frac{1}{\sigma_{[C]}} \left\langle \left| \frac{\partial [C(t)]}{\partial [P'_{\text{tot}}]} \right|_{[P'_{\text{tot}}]=0} \right| \right\rangle , \tag{3.2}$$

where $[P'_{\text{tot}}]$ is the total concentration of a binding site for a second downstream subsystem $\mathcal{D}'$.

On the basis of these metrics, Barton and Sontag argued that *producing better insulators(low retroactivity) requires substantial energy consumption.* However, these conclusions are a direct result of the particular choice of retroactivity metrics, which we believe are poorly-justified. Firstly, although the presence of retroactive terms in the dynamical equations does influence the output of a system, it is unclear why the deviation of the output from a particular hypothetical "ideal" system should quantify retroactivity. For a start, one could write down other "ideal" systems in which the retroactive terms were removed. But more importantly, the metric $\mathscr{D}$ does not quantify the back-action felt by $\mathcal{U}$. For example, it is large if $\mathcal{U}$ is completely decoupled from any insulator and downstream network. Such a system does a poor job of propagating a signal, but doesn't exhibit retroactivity in any meaningful sense.

The competition metric $\mathscr{C}$ comes closer to the spirit of retroactivity, in that it quantities the effect of one downstream subsystem on another via $\mathcal{U}$. However, minimising $\mathscr{C}$ with respect to the parameters in $\mathcal{D}$ and/or $\mathcal{I}$, rather than the newly-added subsystem, does not minimise retroactivity due to the $\mathcal{I}+\mathcal{D}$ subsystem. Instead, it involves making $\mathcal{I}+\mathcal{D}$

insensitive to the presence of the new downstream subsystem – which can be achieved, for example, by coupling to $\mathcal{U}$ very strongly, so that the introduction of a new downstream system has essentially no effect. Such a design would be highly retroactive in the sense that $\mathcal{I} + \mathcal{D}$ strongly influences $\mathcal{U}$, but would have a low value of $\mathscr{C}$.

We therefore believe that the "optimal" systems found by Barton and Sontag do not minimise retroactivity. Instead they identified subsystems $\mathcal{I}$ that allow rapid tracking of $\mathcal{U}$, and are relatively insensitive to the introduction of parallel downstream subsystems, due to strong coupling between $\mathcal{U}$ and $\mathcal{I}$ that results in high fuel turnover.

In our case, we will assume slow signal variation that implies that our systems reach steady states – as in Ref. [156]. We make this simplifying assumption to deconvolve the distinct problems of tracking a time-dependent signal and suppressing retroactivity.

Taking our cue from Refs. [151, 152, 154] we use the retroactivity metric

$$\mathcal{R} = \left| 1 - \frac{[Z^{\mathrm{ss}}]}{[Z^{\mathrm{ss}}_{\mathcal{D},\mathcal{I}\to\emptyset}]} \right|, \tag{3.3}$$

where $[Z^{\mathrm{ss}}]$ denotes the concentration of free Z in steady state in the actual system, $[Z^{\mathrm{ss}}_{\mathcal{D},\mathcal{I}\to\emptyset}]$ is the steady state concentration of free Z when it is not connected to the downstream system ($\mathcal{D}$ and $\mathcal{I}$ are absent, $\mathcal{D},\mathcal{I} \to \emptyset$). We find this metric to be more natural than the ones proposed by Barton and Sontag [93], since it directly quantifies the back action on the upstream system. Using this metric, we shall show that high rates of energy consumption are not required to reduce retroactivity in push-pull motifs.

Throughout the thesis, we will use the following reduced units: concentrations will be measured relative to $[C_0] = 10^{-6}\mathrm{M}$. Uni-molecular rate constants will be measured relative to $k_0 = 1s^{-1}$ units and bimolecular rate constants relative to $\frac{k_0}{[C_0]} = 10^6\mathrm{M}^{-1}\mathrm{s}^{-1}$.

In addition to the retroactivity on $\mathcal{U}$, one could also consider the retroactivity experienced by $\mathcal{I}$. This effect was termed "retroactivity to the output" by Del Vecchio *et al.* [151], who considered the design of compound motifs for minimising both input and

output retroactivity simultaneously [166, 154]. However, for a simple insulator intended to prevent retroactivity on $\mathcal{U}$, any retroactivity experienced by $\mathcal{I}$ is irrelevant – in the same way that the retroactivity internal to the compound motifs of Del Vecchio *et al.* is ignored [154]. We therefore restrict our analysis to the retroactivity experienced by the upstream system $\mathcal{U}$. If, however, an insulating circuit needs to couple to multiple downstream subsystems (see Fig. 3.1F), the retroactivity experienced by $\mathcal{I}$ may be relevant. In this case, the properties of those downstream interfaces could be analysed in the same way as we consider the retroactivity on $\mathcal{U}$.

3.3 Dependence of retroactivity on the upstream subsystem

In this section, we will consider three basic alternatives for the internal dynamics of the upstream subsystem $\mathcal{U}$, and explore the consequences for retroactivity. We will first illustrate the problem using a simple choice of $\mathcal{D}$ and the coupling between $\mathcal{U}$ and $\mathcal{D}$. We shall then seek to generalise the results. For our illustrative downstream subsystem $\mathcal{D}$, we consider the inter-conversion of molecular species P and C. We shall take a direct coupling between upstream and downstream subsystems via binding of Z to P to produce C. This setting is an extremely common motif for passing on a signal in biology, provided that the complex C has properties that are distinct from those of P. For example, Z could be a transcription factor that binds to a promoter P, triggering or suppressing translation [167], or a receptor that recruits proteins to the cell membrane when active [157]. Similar motifs are widespread in nucleic acid nanotechnology [56, 158].

1. **Fixed total concentration of Z:** In the simplest case, there is a fixed and finite pool of Z molecules, $[Z_{\text{tot}}] = [Z] + [C]$. This description would approximate a setting in which a pool of Z molecules are suddenly activated or released due to an

external signal, and the overall system reaches a steady state response prior to the deactivation or recapture of Z at the end of the signalling period. A fixed $[Z_{\text{tot}}]$ during the signalling period is consistent with an *in vitro* setting in which there is no net turnover of components, or to an *in vivo* setting in which protein production and decay or dilution of molecules is slow compared to signal dynamics. In this case, we need only solve for the steady state of

$$Z + P \underset{k_{\text{off}}}{\overset{k_{\text{on}}}{\rightleftharpoons}} C \tag{3.4}$$

subject to $[Z_{\text{tot}}] = [Z] + [C]$ and $[P_{\text{tot}}] = [P] + [C]$.

2. **Constant birth/death dynamics:** In this case, we assume that Z molecules undergo a rapid birth/death process in addition to binding to P. Specifically, we imagine that Z molecules are produced and degraded at rate constants k and δ, respectively, as well as binding to P to form the complex C. Thus the system

$$\begin{aligned} \emptyset &\underset{\delta}{\overset{k}{\rightleftharpoons}} Z \\ Z + P &\underset{k_{\text{off}}}{\overset{k_{\text{on}}}{\rightleftharpoons}} C \end{aligned} \tag{3.5}$$

reaches steady state subject to the constraint $[P_{\text{tot}}] = [P] + [C]$. Such a description would approximate an *in vivo* response to a variation of k and δ on a time-scale slower than protein production and decay or dilution. This system was previously analysed in the stochastic setting by Ghaemi *et al.* [156].

3. **Active and inactive forms of Z:** Finally, we consider a setting in which Z exists in both inactive (Z_0) and active (Z) forms, as well as in complex with P. We assume that Z_0 is incapable of forming a complex, and that the total population $[Z'_{\text{tot}}] = [Z_0] + [Z] + [C]$ is fixed. Such a setting would correspond to a situation similar to case (1), but when only a fraction of the the Z molecules are activated

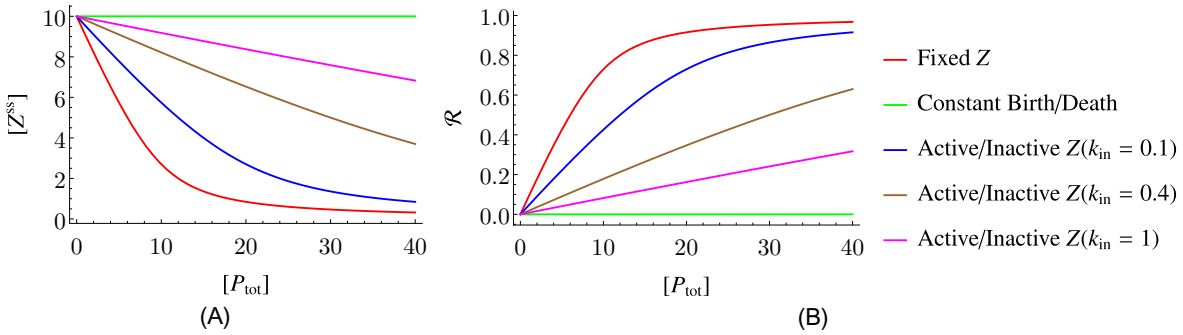


FIGURE 3.2: Replenishing the pool of Z from large reservoirs or by rapid production and decay of components can suppress retroactivity. To illustrate this, we plot the steady-state concentration of Z and the retroactivity metric as a function of $[P_{\text{tot}}]$ for three different upstream subsystems $\mathcal{U}$. Parameters of the system: $k = 10, \delta = k_{\text{on}} = k_{\text{off}} = 1, k_{\text{active}} = 0.1, [Z_{\text{tot}}] = \frac{k}{\delta}, [Z'_{\text{tot}}] = \frac{k_{\text{active}}+k_{\text{inactive}}}{k_{\text{active}}}[Z_{\text{tot}}]$.

or released in response to an external signal. In this case we solve

$$\begin{aligned} Z_0 &\underset{k_{\text{in}}}{\overset{k_{\text{ac}}}{\rightleftharpoons}} Z \\ Z + P &\underset{k_{\text{off}}}{\overset{k_{\text{on}}}{\rightleftharpoons}} C \end{aligned} \tag{3.6}$$

for the steady state subject to the constraints $[Z'_{\text{tot}}] = [Z] + [Z_0] + [C]$ and $[P_{\text{tot}}] = [P] + [C]$. Here, k_{in}, and k_{ac} are first-order rate constants,

The steady-state concentration $[Z^{\text{ss}}]$ is the output signal of $\mathcal{U}$; in the limit $[P_{\text{tot}}] \to 0$ ($\mathcal{D}$ absent), the three alternatives for $\mathcal{U}$ all produce the same signal if

$$\frac{k}{\delta} = [Z_{\text{tot}}] = \frac{k_{\text{ac}}}{k_{\text{ac}} + k_{\text{in}}}[Z'_{\text{tot}}]. \tag{3.7}$$

In Figure 3.2 we show $[Z^{\text{ss}}]$ and the retroactivity metric $\mathcal{R}$ as we increase $[P_{\text{tot}}]$, given equal $[Z^{\text{ss}}]$ for $[P_{\text{tot}}] \to 0$. Figure 3.2 demonstrates that retroactivity is highly sensitive to the internal details of $\mathcal{U}$. Clearly, the system with fixed $[Z_{\text{tot}}]$ shows the strongest retroactivity; the system with constant birth and death of Z shows no retroactivity; and the system with active and inactive forms of Z interpolates between these two limits. Constant birth-death dynamics is analogous to having an infinite pool of Z to draw upon (formally, Z is coupled to a chemostat [168]). On average, a Z molecule gets replenished every time it is consumed after binding to P. As a result, this system has zero retroactivity and $[Z^{\text{ss}}] = \frac{k}{\delta}$ irrespective of $[P_{\text{tot}}]$ – this fact was previously noted in the stochastic setting

by Ghaemi *et al.* [156]. The case with fixed $[Z_{\text{tot}}] = [Z] + [C]$ has the highest retroactivity since there is nothing to replenish Z once it binds to P. The setting with active and inactive forms of Z implies a finite buffer upon which to draw; for low $[P_{\text{tot}}]$, most of the sequestration of Z can be compensated for by conversion of Z_0 into Z, but as $[P_{\text{tot}}]$ grows, this buffer gets depleted. Consequently this third case is moderately retroactive, interpolating between the regimes of fixed $[Z]+[C]$ and constant birth-death dynamics. In particular when $k_{\text{inactive}} \gg k_{\text{active}}$, this intermediate case approaches constant birth-death dynamics. Refer to the case of Active/Inactive forms of Z in Section 7.1 of Appendix B for a proof of this fact. More generally, Section 7.1 gives analytic expressions for each of the regimes considered here. As shown by Ghaemi *et al.* [156] in the setting of non-deterministic dynamics, introducing birth-death dynamics for Z is a general approach to buffering against the influence of downstream systems $\mathcal{D}$. In a wide range of steady-state contexts, this buffering eliminates retroactivity. In particular, $[Z] = k/\delta$ necessarily holds if the reaction network obeys detailed balance [169, 170, 171, 57]. However, a constant decay rate of the complex C_1 and a constant production of P is sufficient to compromise this perfect buffering. Explicitly, the system

$$\begin{aligned} \emptyset &\underset{\delta}{\overset{k}{\rightleftharpoons}} Z, & Z + P &\underset{\beta_2}{\overset{\beta_1}{\rightleftharpoons}} C_1 \\ C_1 &\xrightarrow{\alpha_1} \emptyset, & \emptyset &\xrightarrow{\gamma_1} P. \end{aligned} \tag{3.8}$$

has the following set of differential equations:

$$\begin{aligned} [\dot{Z}] &= k - \delta[Z] - \beta_1[Z][P] + \beta_2[C_1] \\ [\dot{C_1}] &= \beta_1[Z][P] - \beta_2[C_1] - \alpha_1[C_1] \\ [\dot{P}] &= -\beta_1[Z][P] + \beta_2[C_1] + \gamma_1. \end{aligned} \tag{3.9}$$

In steady-state, we have $[Z^{\text{ss}}] = \frac{k-\gamma_1}{\delta} \neq \frac{k}{\delta}$ implying non-zero retroactivity. Increasing γ_1 increases retroactivity, as more Z molecules are consumed by the downstream system and never released. Even in the case of perfect buffering, implementing a system that produces

and degrades components on a time scale that is fast compared to signal variation would be extremely expensive. Turning over a single protein molecule, for example, costs a cell on the order of thousands of ATP fuel molecules [172]. The alternative of having a very large but fixed pool of molecules from which to create Z is also costly; energy (and in the cell, space) needs to be devoted to these molecules, and in vivo the large population would need to be maintained against a background dilution/decay, which is more costly than maintaining a small population. Similar arguments apply to maintaining a large pool of Z that bind to P only weakly, or more complex $\mathcal{U}$ subsystems that replenish Z. In summary, although the use of a large or bottomless supply of Z can suppress retroactivity, the inherent cost of this strategy is related to the cost of producing a large number of molecules, which is generally high. In Section 3.5, we consider the alternative of using an insulating push-pull motif, in which retroactivity is suppressed without excessive production of the signalling species. Instead, an energy-consuming circuit involving catalytic molecular modification is used to reduce sequestration of Z. This modification consumes chemical fuel molecules such as ATP, which are far less costly than, for example, entire proteins. It is because of the relatively low cost of post-translational modification that we focus on the push-pull motif, rather than alternative insulators incorporating protein production and degradation [151]. We now recall explicit formulae for basic notions in thermodynamics like the free energy of a reaction and the binding energy of a complex.

3.4 Free energy of a reaction

Consider the following reaction

$$A + B \underset{k_2}{\overset{k_1}{\rightleftharpoons}} C \tag{3.10}$$

Then,

$$\Delta G = \Delta G_0 + k_B T \ln \frac{[C]}{[A][B]} \tag{3.11}$$

where ΔG is the net free energy difference between the products and reactants, ΔG_0 (also called the **binding energy** of C) is the intrinsic free energy difference between products and reactants. In the present case, $\Delta G_0 = \ln \frac{k_2}{k_1}$. Therefore, the equation of free energy becomes

$$\Delta G = k_B T \ln \frac{k_2}{k_1} + k_B T \ln \frac{[C]}{[A][B]} \tag{3.12}$$

3.5 The relationship between retroactivity and fuel consumption for an insulating push-pull motif

Having argued about the costs of suppressing retroactivity through the design of $\mathcal{U}$, we now turn to the alternative approach of using insulating motifs as shown in Fig. 3.1E.ii. We reiterate the question first raised by [93]: is increased consumption of fuel necessary for suppression of retroactivity in a push-pull network? To approach this question, we will consider an insulating push-pull motif for an upstream network $\mathcal{U}$ with a fixed total amount of Z. This was the simplest $\mathcal{U}$ considered in Section 3.3. This choice is motivated not only by simplicity, but also because it is the most challenging context for an insulator (underlying retroactive effects are strongest) and because the alternative choices of $\mathcal{U}$ are associated with their own additional resource costs. Our system is therefore defined by the set of reactions below, following Ref. [93]:

$$\begin{aligned} Z + X &\underset{\beta_2}{\overset{\beta_1}{\rightleftharpoons}} C_1 \xrightarrow{k_1} X^* + Z, \\ Y + X^* &\underset{\alpha_2}{\overset{\alpha_1}{\rightleftharpoons}} C_2 \xrightarrow{k_2} X + Y, \\ X^* + p &\underset{k_{\text{off}}}{\overset{k_{\text{on}}}{\rightleftharpoons}} C. \end{aligned} \tag{3.13}$$

Here we have assigned mass-action rate constants to each step, and assumed that the role of the molecular fuel molecules ATP, ADP and P_i can be implicitly absorbed into rate constants, as is common. We have also assumed that the free energy of ATP breakdown, ΔG_{ATP}, is sufficiently large that dephosphorylation via Z and phosphorylation via Y are never observed.

An explicit representation of the intermediate catalyst-substrate complexes C_1 and C_2 allows for a quantification of sequestration of Z by the push-pull insulator. Again, we consider the fractional reduction in free Z due to the introduction of the downstream system:

$$\mathcal{R} = \left| 1 - \frac{[Z^{\mathrm{ss}}]}{[Z^{\mathrm{ss}}_{\mathcal{D},\mathcal{I}\to\emptyset}]} \right| = \frac{[C_1^{\mathrm{ss}}]}{[Z_{\mathrm{tot}}]}. \tag{3.14}$$

The system illustrated in Eq. 3.13 turns over a single molecule of ATP per phosphorylation/dephosphorylation cycle, in which a molecule of X is first activated by Z then deactivated by Y. Thus the fuel consumption rate per unit volume is given by the net flux Ψ of X molecules around this cycle. The overall power per unit volume is given by $w = \Psi \Delta G_{\mathrm{ATP}}$, where ΔG_{ATP} is the free energy released by the breakdown of a single ATP molecule. From inspection,

$$\Psi = k_1[C_1^{\mathrm{ss}}], \quad w = \Psi \Delta G_{\mathrm{ATP}} = k_1[C_1^{\mathrm{ss}}]\Delta G_{\mathrm{ATP}}. \tag{3.15}$$

Strictly speaking, ΔG_{ATP} should be infinite since we have approximated the catalytic reactions as microscopically irreversible, as in [93, 151, 173, 174]. In practice, ΔG_{ATP} is assumed to have a large, fixed, negative value. The power consumption will then be essentially determined by the flux Ψ through the cycle. As a consequence, we see that both retroactivity and fuel consumption grow proportionally to $[C_1^{\mathrm{ss}}]$. In intuitive terms both fuel turnover through phosphorylation, and the retroactivity, increase if Z is frequently bound to X to form the enzyme-substrate complex C_1. This observation raises questions about the conclusion that increased energy consumption is necessary to *suppress* retroactivity from Ref. [93]. To explore this idea further, let us consider whether an insulating system can in general be tuned to reduce both retroactivity and fuel consumption whilst

maintaining its signal-transducing function. Specifically, let us ask whether combinations of system parameters can be changed so that the input-output relation $[C^{\rm ss}]([Z_{\rm tot}])$ is approximately preserved, but both w and $\mathcal{R}$ are reduced. In Refs. [151, 93, 154, 175, 176], optimisation of insulating circuits was only performed over the total concentrations of species. Natural evolution or deliberate design will, however, allow for moderation of at least some of the chemical rate constants, and we will focus on these parameters. If it were possible to take the catalytic rate constants $k_1, k_2 \to \infty$, we would obtain $[C_1^{\rm ss}] \to 0$ and $\mathcal{R} \to 0$, and retroactivity could be completely eliminated, regardless of the fuel consumption rate. However, these catalytic rate constants encode complex chemistry that is likely difficult to accelerate. We shall therefore take k_1, k_2 as fixed and instead focus on α_1, β_1, the rate constants of enzymatic binding. Whilst there is a diffusion-based limit to how far α_1, β_1 can be *increased* [177], we assume that it is always possible to reduce the speed with which a given catalyst binds to its substrate. Alternatively, we could have focussed on α_2, β_2 (the unbinding rate constants), assuming that it is always possible to increase them. In both cases, which yield similar results, we essentially assume that it is possible to reduce the catalyst-substrate binding affinity, either through design of synthetic systems or via evolution of natural systems. For example, this reduced affinity could arise from replacing a hydrophobic residue by a hydrophilic one in a protein-protein interaction or creating a mismatch in DNA-DNA interaction. We first consider the low

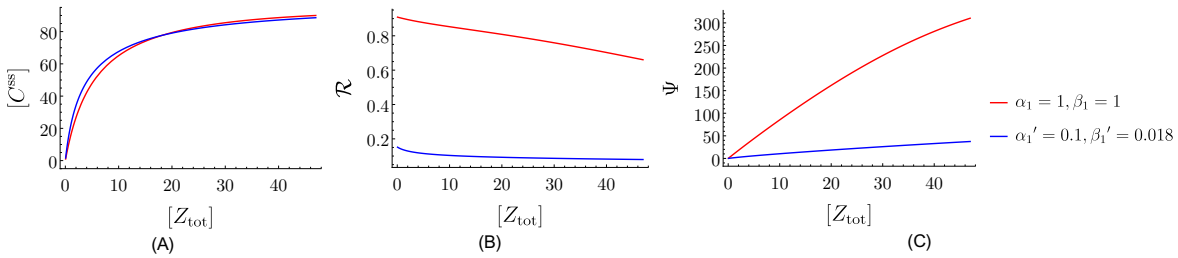


FIGURE 3.3: An example of reducing both fuel consumption and retroactivity whilst maintaining an approximately fixed input-output relation, even with high initial retroactivity. (A) Input-output relation for two systems with distinct α_1, β_1, but otherwise identical parameters. The second curve is obtained by setting $\beta_1' = 0.018\beta_1$, and adjusting α_1' to maximise the similarity between curves. (B) Retroactivity $\mathcal{R}$ for the system in (A), illustrating substantially lower retroactivity for $\alpha_1', \beta_1' < \alpha_1, \beta_1$. (C) Flux Ψ, which is much reduced for $\alpha_1', \beta_1' < \alpha_1, \beta_1$. Other parameters: $[X_{\rm tot}] = 200, [Y_{\rm tot}] = 100, [P_{\rm tot}] = 100, \alpha_2 = \beta_2 = k_1 = k_2 = k_{\rm on} = k_{\rm off} = 10$.

retroactivity limit, when $[C_1^{\text{ss}}]$ and $[C_2^{\text{ss}}]$ are both small. In this case,

$$[C^{\text{ss}}] \approx \frac{f(r) \pm \sqrt{f^2(r) - 4r^2 k_{\text{on}}^2 [X_{\text{tot}}][P_{\text{tot}}]}}{2rk_{\text{on}}} \tag{3.16}$$

where

$$r = \frac{\beta_1}{\alpha_1} \frac{k_1[Z_{\text{tot}}]}{k_2[Y_{\text{tot}}]} \frac{k_2 + \alpha_2}{\beta_2 + k_1}, \tag{3.17}$$

and

$$f(r) = k_{\text{off}} + r(k_{\text{on}}([P_{\text{tot}}] + [X_{\text{tot}}]) + k_{\text{off}}). \tag{3.18}$$

In this limit, $[C^{\text{ss}}]$ is a function of β_1/α_1, rather than α_1 and β_1 independently, if all other parameters are fixed. Moreover,

$$\mathcal{R}, \Psi \propto \beta_1 \text{ at fixed } \frac{\beta_1}{\alpha_1} \tag{3.19}$$

(see Section 7.2 in Appendix B for a derivation of these facts). If we then reduce β_1 and α_1 by the same factor ϕ whilst keeping all other parameters fixed, the input-output relation $[C^{\text{ss}}]([Z_{\text{tot}}])$ is unchanged whilst $\mathcal{R}$ and w are both reduced by ϕ. In principle, this simultaneous reduction of retroactivity and fuel consumption at fixed input-output relation can proceed arbitrarily far. The above observation forms the intuition behind the main claim of this chapter. Fundamentally, the push-pull network responds to a competition between Z and Y. We can therefore reduce the strength with which both Z and Y couple to X, whilst maintaining the same steady-state output. Reducing the coupling to X serves to minimise both retroactivity and energy consumption. If $[C_1^{\text{ss}}]$ and $[C_2^{\text{ss}}]$ are not both small, the input-output relationship is not a function of $\frac{\beta_1}{\alpha_1}$ only. It is therefore no longer possible to reproduce input-output relations exactly as outlined above. However, we can instead consider reducing $\beta_1 \to \beta_1'$, and identifying the corresponding change $\alpha_1 \to \alpha_1'$ that reproduces the original input-output curve as closely as possible. Specifically, we identify the new α_1' as the value that minimizes the following measure of

the difference between input-output relations

$$\int_l^u |[C^{\rm ss}]([Z_{\rm tot}], \alpha_1, \beta_1) - [C^{\rm ss}]([Z_{\rm tot}], \alpha_1', \beta_1') \mid \mathrm{d}[Z_{\rm tot}], \tag{3.20}$$

where l and u are such that $C^{\rm ss}(l, \alpha_1, \beta_1) \approx 0.01[P_{\rm tot}]$ and $C^{\rm ss}(u, \alpha_1, \beta_1) \approx 0.9[P_{\rm tot}]$.

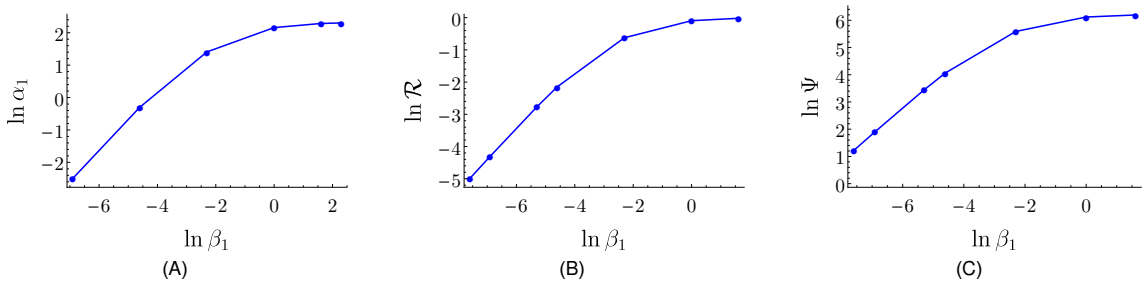


FIGURE 3.4: In the limit of low $[C_1^{\rm ss}]$ and $[C_2^{\rm ss}]$ or equivalently low α_1 and β_1, the steady-state concentration of the output $[C^{\rm ss}]$ depends on the ratio $\frac{\beta_1}{\alpha_1}$ and not α_1 and β_1 individually. (A) Scaling of α_1 and β_1 as they are simultaneously adjusted to retain a given input-output curve. (B) Scaling of retroactivity $\mathcal{R}$ with β_1 as this operation is performed. (C) Scaling of net flux Ψ with β_1 as this operation is performed. Both the flux and retroactivity decrease as β_1 is decreased, becoming proportional to β_1 at fixed ratio $\frac{\beta_1}{\alpha_1}$ in the low retroactivity limit. Constant parameters for the network: $[X_{\rm tot}] = 400, [Y_{\rm tot}] = 150, [Z_{\rm tot}] = 50, [P_{\rm tot}] = 100, \alpha_1 = 10, \alpha_2 = 15, \beta_1 = 10, \beta_2 = 15, k_1 = k_2 = k_{\rm on} = k_{\rm off} = 10$.

Even when retroactivity (and hence $[C_1^{\rm ss}]$) is high for the original parameters α_1, β_1, it is frequently possible to approximate the input-output relation with reduced α_1', β_1'. Consider, for example, the input-output curves in Fig.3.3, in which $\mathcal{R} \sim 0.8$. As expected, the reduced α_1', β_1' give substantially lower retroactivity and fuel consumption. Further examples are provided in Section 7.5 of Appendix B.

The above process can be iterated, producing even lower retroactivity and energy consumption by a continuing reduction in coupling strength between Z and the insulator. We illustrate the process in Fig. 3.4. Eventually, the limit of small $[C_1^{\rm ss}]$ and $[C_2^{\rm ss}]$ is reached and α_1 and β_1 decrease in proportion, with $\mathcal{R}$ and Ψ also scaling proportionally.

3.6 Incorporating microscopic reversibility for an insulating push-pull motif

In principle, all chemical reactions are microscopically reversible [178, 77]. This fact is often ignored when studying physiological ATP-driven systems, as it was by Barton and Sontag [93], since the high free energy of ATP hydrolysis [179] can render reverse reactions irrelevant to the eventual steady state. Such an approximation is typically reasonable for push-pull motifs driven by free energies substantially in excess of $4k_BT$ [162], as we confirm in our case in Section 7.3 of Appendix B. Nonetheless, a full understanding of the resource requirements of insulators requires explicit treatment of reverse reactions, since their contribution depends directly on the free energy consumed per cycle. Moreover, in developing synthetic systems, a chemical fuel with a free energy as high as physiological ATP may be unavailable or undesirable. We therefore explicitly incorporate microscopic reversibility into our discussion in this session. We introduce microscopically reversible reactions in the simplest possible way, still assuming a single long-lived catalyst/substrate complex. In particular, we set the rate constant corresponding to the reaction $X^* + Z \to C_1$ to be $\frac{\epsilon k_1 \beta_1}{\beta_2}$ and the rate constant corresponding to $X + Y \to C_2$ to be $\frac{\epsilon k_2 \alpha_1}{\alpha_2}$. The system is represented by the following

$$\begin{aligned} Z + X &\underset{\beta_2}{\overset{\beta_1}{\rightleftharpoons}} C_1 \underset{\frac{\epsilon k_1 \beta_1}{\beta_2}}{\overset{k_1}{\rightleftharpoons}} X^* + Z, \\ Y + X^* &\underset{\alpha_2}{\overset{\alpha_1}{\rightleftharpoons}} C_2 \underset{\frac{\epsilon k_2 \alpha_1}{\alpha_2}}{\overset{k_2}{\rightleftharpoons}} X + Y, \\ X^* + p &\underset{k_{\text{off}}}{\overset{k_{\text{on}}}{\rightleftharpoons}} C. \end{aligned} \tag{3.21}$$

Here, ϵ is the parameter that modulates the distance of the system from equilibrium; $\epsilon = 0$ corresponds to a completely irreversible network, with infinite driving, while $\epsilon = 1$ corresponds to an equilibrium network. To illustrate this, consider the reactions corre-

sponding to the phosphorylation-dephosphorylation cycle

$$Z + X \underset{\beta_2}{\overset{\beta_1}{\rightleftharpoons}} C_1 \underset{\frac{\epsilon k_1 \beta_1}{\beta_2}}{\overset{k_1}{\rightleftharpoons}} X^* + Z,$$
$$Y + X^* \underset{\alpha_2}{\overset{\alpha_1}{\rightleftharpoons}} C_2 \underset{\frac{\epsilon k_2 \alpha_1}{\alpha_2}}{\overset{k_2}{\rightleftharpoons}} X + Y, \tag{3.22}$$

Using Equation 3.12, the free energy of the molecular fuel consumed in a single cycle is given by $\Delta G_{\text{ATP}} = k_B T(\ln \frac{\beta_2 [C_1]}{\beta_1 [Z][X]} + \ln \frac{[Z][X^*]\frac{\epsilon k_1 \beta_1}{\beta_2}}{k_1 [C_1]} + \ln \frac{\alpha_2 [C_2]}{\alpha_1 [Y][X^*]} + \ln \frac{[Y][X]\frac{\epsilon k_2 \alpha_1}{\alpha_2}}{k_2 [C_2]}) = 2k_B T \ln \epsilon$. Therefore, $\epsilon = 0$ corresponds to $\Delta G_{\text{ATP}} \sim -\infty$ and $\epsilon = 1$ corresponds to an equilibrium network with $\Delta G_{\text{ATP}} = 0$

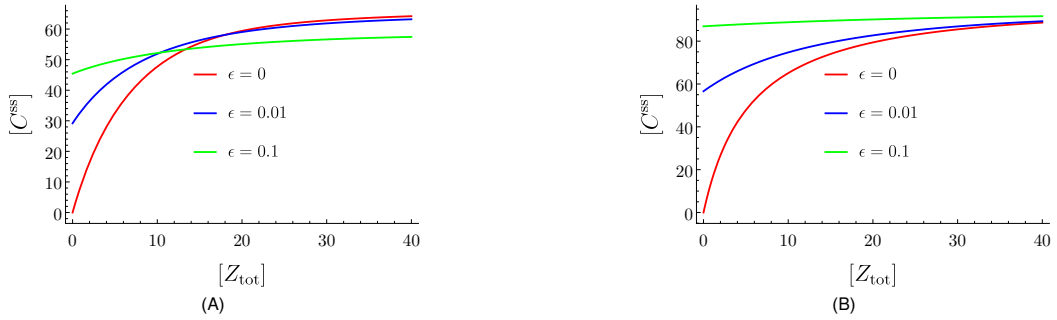


FIGURE 3.5: Increasing microscopic reversibility limits the ability of the network to produce a wide range of output. Parameters used for the network: A) $[X_{\text{tot}}] = 100, [Y_{\text{tot}}] = 100, [P_{\text{tot}}] = 100, \alpha_1 = \beta_1 = \alpha_2 = \beta_2 = 0.1, k_1 = k_2 = 1, k_{\text{on}} = k_{\text{off}} = 10$. B) $[X_{\text{tot}}] = 200, [Y_{\text{tot}}] = 100, [P_{\text{tot}}] = 100, \alpha_1 = \beta_1 = 0.1, \alpha_2 = \beta_2 = k_1 = k_2 = 1, k_{\text{on}} = k_{\text{off}} = 10$.

Fundamentally, imposing a finite free energy per fuel molecule through microscopic reversibility limits the overall range of the input-output function $[C^{\text{ss}}]([Z_{\text{tot}}])$. Intuitively, if catalysts function in both directions, $[X] > 0$ even if $[Z_{\text{tot}}]/[Y_{\text{tot}}] \to \infty$. Similarly, $[X^*] > 0$ even if $[Z_{\text{tot}}]/[Y_{\text{tot}}] \to 0$. Indeed, as we show in Section 7.4 of the Appendix, the following relation holds true:

$$\epsilon \leq \frac{[X^*]}{[X]} \leq \frac{1}{\epsilon}, \tag{3.23}$$

implying a reduced dynamic range of the insulator and hence a weaker propagation of the signal from Z to the output. We illustrate this intuition for a specific system in Figure 3.5. The overall range of the input-output function drops as ϵ increases from 0 towards 1. Although potentially problematic for signal propagation, the reduction in

range of the input-output function $[C^{\rm ss}]([Z_{\rm tot}])$ is not an inherently retroactive effect. It is not a direct consequence, nor a cause, of sequestration of Z by the downstream subsystem. We now explicitly consider the effect of microscopic reversibility on retroactivity. We show that decreasing the free energy consumed per fuel molecule can increase retroactivity due to rebinding of products to catalysts. However, the essential arguments of Section 3.5 remain valid; it is still possible to simultaneously reduce free-energy consumption and retroactivity through weak coupling of Z and Y to X. For the push-pull motif that explicitly incorporates the microscopically reversible reactions,

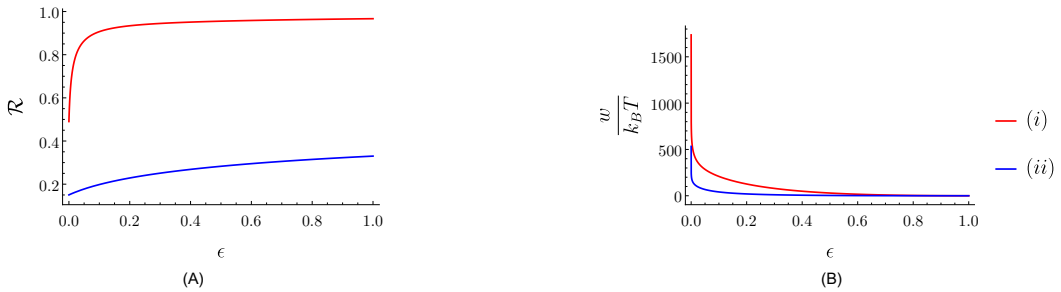


FIGURE 3.6: A) Inclusion of microscopically reversible reactions can increase the retroactivity of the motif due to rebinding of products to catalysts; $\mathcal{R}$ increases with ϵ for a specific system while all other parameters are fixed. B) The rate of energy consumption decreases with increase in ϵ. However, this is not primarily due to a decrease in flux, but rather due to the decrease in free energy of ATP molecules. Parameters of the network: *i)* $[X_{\rm tot}] = 300, [Y_{\rm tot}] = 50, = [Z_{\rm tot}] = 100, [P_{\rm tot}] = 100, \alpha_1 = \alpha_2 = \beta_1 = \beta_2 = k_1 = k_2 = 1, k_{\rm on} = k_{\rm off} = 10$.*(ii)* $[X_{\rm tot}] = [Y_{\rm tot}] = [Z_{\rm tot}] = [P_{\rm tot}] = 100, \alpha_1 = \alpha_2 = \beta_1 = \beta_2 = 0.1, k_1 = k_2 = 1, k_{\rm on} = k_{\rm off} = 10$.

$$\Psi = k_1[C_1^{\rm ss}] - \frac{\epsilon k_1 \beta_1}{\beta_2}[X^{*\rm ss}][Z^{\rm ss}], \tag{3.24}$$

and

$$w = 2k_B T \ln \epsilon \left(k_1[C_1^{\rm ss}] - \frac{\epsilon k_1 \beta_1}{\beta_2}[X^{*\rm ss}][Z^{\rm ss}] \right), \tag{3.25}$$

whilst the retroactivity remains $\mathcal{R} = [C_1^{\rm ss}]/[Z_{\rm tot}]$. Away from the microscopically irreversible limit of $\epsilon \to 0$, retroactivity and energy consumption are less directly related than in Section 3.5. Indeed, if we simply keep all other parameters fixed whilst varying ϵ, it is possible to simultaneously increase retroactivity and decrease overall power consumption (or vice-versa). In particular, both Ψ and w tend to decrease as $\epsilon \to 1$, but $\mathcal{R}$ can be

enhanced as both X and X^* can bind to Z to produce C_1. A specific example is given in Fig. 3.6.

The above observation, however, does not imply that reduction of retroactivity necessarily requires high rates of free energy consumption. In particular, for a push-pull motif of fixed ϵ, we can play essentially the same trick as before: reduce the strength of coupling to the push-pull by decreasing both α_1 and β_1 in such a way that approximately maintains the input-output relation $[C^{ss}]([Z_{tot}])$. In fact, just as in the completely irreversible case, we show that in the limit of low $[C_1^{ss}]$ and $[C_2^{ss}]$, the steady-state output is a function of $\frac{\beta_1}{\alpha_1}$ but not α_1 and β_1 separately. Moreover,

$$\mathcal{R}, w \propto \beta_1 \text{ at fixed } \frac{\beta_1}{\alpha_1} \tag{3.26}$$

still holds (see Section 7.2 of Appendix B). In the limit of low retroactivity, one can

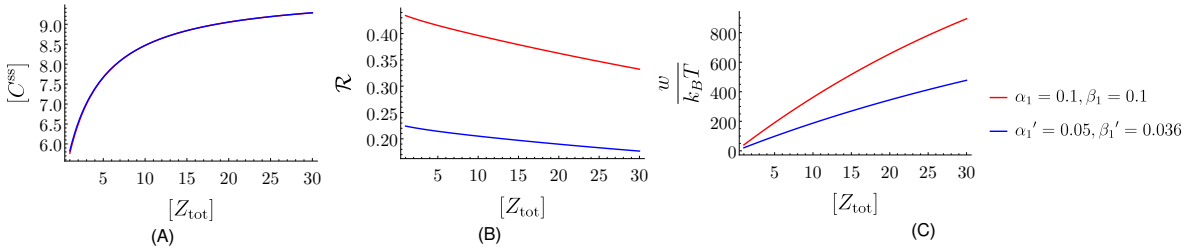


FIGURE 3.7: Simultaneous reduction in retroactivity and energy consumption whilst approximately maintaining the input-output relation at a fixed and finite free energy stored per fuel molecule. (A) Two different sets of binding rates α_1, β_1 and α_1', β_1' that give a similar input-output relation with all other parameters fixed (α_1' is chosen by minimising Eq. 3.20 for given $\alpha_1, \beta_1, \beta_1'$). (B) and (C) show retroactivity $\mathcal{R}$ and power w for the two cases. Other Parameters used $[X_{tot}] = [Y_{tot}] = 100, [P_{tot}] = 10, \alpha_1 = \beta_1 = 0.1, \alpha_2 = \beta_2 = 1, k_1 = k_2 = k_{on} = k_{off} = 10, \epsilon = 0.01$.

therefore decrease both α_1 and β_1 in proportion to give the same input/output curve at reduced retroactivity and energy consumption, as before. For higher $[C_1^{ss}]$ and $[C_2^{ss}]$, just as in Section 3.5, it is not possible to obtain the same input-output curve by varying β_1 at fixed $\frac{\beta_1}{\alpha_1}$. However, we again observe that in most cases a good fit to the input-output relation can be obtained by reducing β_1 and α_1 in such a way as to minimize Eq. 3.20, even when $\mathcal{R}$ is appreciable, in the process reducing both $\mathcal{R}$ and w. We demonstrate this behaviour for a specific system in Fig. 3.7; other examples are given in Section 7.5 of Appendix B. Yet another way to confirm this effect is to see how retroactivity behaves with

the rate of energy consumption as we vary the coupling to the push-pull motif (by changing α_1, β_1), whilst maintaining the same steady-state output. As we expect, retroactivity decreases with decrease in the rate of energy consumption. Figure 3.8 confirms this fact.

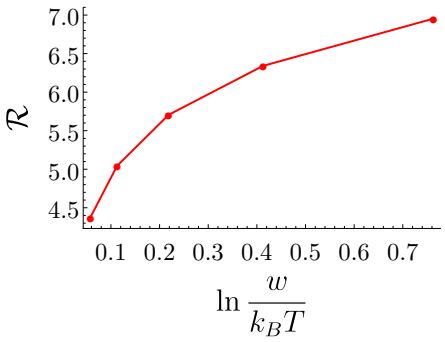


FIGURE 3.8: Decreasing the coupling to the push-pull motif (by decreasing α_1, β_1), whilst maintaining the same input-output relationship decreases both retroactivity and the rate of energy consumption. Parameters used: $[X_{\text{tot}}] = [Y_{\text{tot}}] = 100, [Z_{\text{tot}}] = 15, [P_{\text{tot}}] = 10, \alpha_2 = \beta_2 = 1, k_1 = k_2 = k_{\text{on}} = k_{\text{off}} = 10, \epsilon = 0.01$.

3.7 Arbitrarily weak coupling to an insulator causes vulnerability to cross-talk

Biochemical signalling pathways do not exist in isolation; in both natural and complex synthetic systems multiple information transmission pathways based on similar reactions must co-exist [180, 56, 181, 182]. Transferring information to only the desired downstream recipients is a challenge in specificity; the possibility of unintended interference would compromise information transduction. In this section we demonstrate how cross-talk limits the degree to which weak insulator-coupling allows effective signalling with low fuel consumption and low retroactivity. To do so we consider the system in Eq. 3.27, with an additional upstream molecule Z' that couples to X through an accidental leak

reaction. As a result, we get the following network:

$$\begin{aligned} Z' + X &\underset{\gamma_2}{\overset{\gamma_1}{\rightleftharpoons}} C_4 \xrightarrow{k_3} X^* + Z', \\ Z + X &\underset{\beta_2}{\overset{\beta_1}{\rightleftharpoons}} C_1 \xrightarrow{k_1} X^* + Z, \\ Y + X^* &\underset{\alpha_2}{\overset{\alpha_1}{\rightleftharpoons}} C_2 \xrightarrow{k_2} X + Y, \\ X^* + p &\underset{k_{\text{off}}}{\overset{k_{\text{on}}}{\rightleftharpoons}} C. \end{aligned} \tag{3.27}$$

Conceptually, Z' represents the combined effect of many alternative catalysts that could cause accidental activation of X via a leak reaction. It will therefore be challenging to reduce γ_1 arbitrarily far, either by evolution or design, whilst retaining the functions of these alternative catalysts within their intended pathways. In Fig. 3.9, we repeat the protocol of progressively weakening α_1 and β_1 while attempting to preserve the input-output curve $[C^{\text{ss}}]([Z_{\text{tot}}])$ as outlined in Section 3.5, comparing a system with $\gamma_1 = 0$ to a system with a fixed and finite γ_1. Crucially, we now observe that as the coupling between Z and X becomes weaker, Z' starts to dominate the insulator's behaviour. When the coupling between Z' and X exceeds the coupling between Z and X, we are no longer able to reduce α_1 and β_1 to give a close match to the original curve, since $[C^{\text{ss}}]$ responds primarily to Z' rather than Z. The range of the input-output function $[C^{\text{ss}}]([Z_{\text{tot}}])$ is consequently compromised, and signal propagation becomes ineffective.

The strength of leak reactions or cross-talk thus determines the degree to which effective signalling can be maintained despite weak coupling between $\mathcal{U}$ and the insulator $\mathcal{I}$. The system $\mathcal{U}$ must couple more strongly than cross-talk reactions, and consequently fuel turnover and retroactivity cannot be suppressed arbitrarily far whilst retaining a functioning network. A similar consideration shows that the rates of spontaneous phosphorylation and dephosphorylation also limit the degree to which Z can couple weakly to the insulator and remain effective. It is important to note, however, the logical distinction between the observation that some degree of fuel turnover and retroactivity are required for effective signal propagation, and the suggestion that increased fuel consumption is

FIGURE 3.9: Evidence that accidental leak reactions limit the degree to which coupling between $\mathcal{U}$ and $\mathcal{I}$ can be reduced whilst maintaining a functioning network. (A) In the absence of a leak reaction, one can decrease α_1 and β_1 successively using exactly the procedure outlined in Section 3.5, maintaining an approximately constant input-output curve. (B) In the presence of a leak reaction caused by a signal molecule Z', one is able to reduce α_1 and β_1 to match a given input-output curve whilst the coupling of Z' remains relatively weak. However, eventually Z' becomes dominant and one cannot find suitable parameters α_1 and β_1 to match a given input-output curve well. Parameters used for both networks: $[X_{\text{tot}}] = 150, [Y_{\text{tot}}] = 100, [Z'_{\text{tot}}] = 100, [P_{\text{tot}}] = 100, \alpha_1 = \beta_1 = 1, \alpha_2 = \beta_2 = 10, \gamma_1 = 0.01, \gamma_2 = k_1 = k_2 = k_{\text{on}} = k_{\text{off}} = 10, k_3 = 1$.

required to suppress retroactivity.

3.8 Conclusions

We have considered the suppression of retroactivity in molecular signal transduction systems by both the design of the upstream subsystem $\mathcal{U}$, and by incorporating an insulator $\mathcal{I}$ between the $\mathcal{U}$ and the downstream subsystem $\mathcal{D}$. Using the fractional reduction in the concentration of the output of $\mathcal{U}$ due to the presence of $\mathcal{D}/\mathcal{I}$ as a metric for retroactivity [151, 154], we find that retroactivity is strongly dependent on the design of $\mathcal{U}$, and that insulators can suppress retroactivity at low levels of fuel consumption.

In particular, if $\mathcal{U}$ consists of a single species Z undergoing production and decay on a fast time-scale relative to signal switching, retroactivity can be eliminated in the steady state for certain downstream systems $\mathcal{D}$ (as previously noted in [156] for a specific case). More generally, birth/death dynamics serves to buffer the concentration of Z against the influence of $\mathcal{D}$, reducing retroactivity. However, such a buffering would incur substantial resource costs, requiring a high turnover of molecules or the establishment and maintenance of a very large buffer population.

We then consider the behaviour of certain catalytic circuits called push-pull motifs that can act as insulators $\mathcal{I}$, to explore whether they can reduce retroactivity at low cost. These insulators do not require a high production rate of signalling molecules, nor the establishment and maintenance of a large population of said molecules. Instead, the insulators consume fuel, typically by converting ATP into ADP and inorganic phosphate.

We argue that coupling $\mathcal{U}$ to $\mathcal{I}$ weakly reduces both the retroactivity and fuel consumption. Moreover, in the steady-state signalling limit, it is often possible to simultaneously reduce both retroactivity and fuel consumption to arbitrary low levels, whilst maintaining an approximately fixed signal propagation from $\mathcal{U}$ to the output of $\mathcal{D}$. Note that we do not claim that one can *always* match an input-output curve with weaker coupling. In particular, motifs based on zero-order ultra-sensitivity [174, 183, 184, 185] actually leverage retroactive effects. However, in such contexts retroactivity is a key ingredient of the system, rather than a nuisance to be eliminated. Additionally, in these cases it would be incorrect to say that suppressing retroactivity requires more fuel consumption – instead, suppressing retroactivity and fuel consumption simultaneously comes at the expense of signal alteration.

Therefore it is in general possible to suppress retroactivity at low cost through insulation, and an engineer could design a signalling network with low energy consumption and low retroactivity. This observation still holds when the finite free energy associated with the breakdown of each ATP is explicitly modelled through microscopically reversible reactions. We note that unlike increasing the concentration of insulator molecules [151, 93, 154, 175, 176], which strongly influences the interactions of both $\mathcal{U}$ and $\mathcal{D}$ with $\mathcal{I}$, there is no reason why changing the coupling of $\mathcal{I}$ to $\mathcal{U}$ at their mutual interface should make $\mathcal{I}$ more subject to "retroactivity to the output" at its interface with downstream subsystems [151]. This fact supports the approach of considering only the retroactivity on $\mathcal{U}$ in our analysis. However, the presence of unintended leak reactions limits the degree to which the coupling to the insulator can be weakened before signal transduction is

compromised.

In this work we have assumed that the signalling network reaches steady state. We have thus not considered its ability to respond to fast variation of the parameters of the upstream subsystem, as in some previous studies [151, 93]. Tracking rapid variation in $\mathcal{U}$ is impossible if $\mathcal{U}$ only couples weakly to downstream subsystems since insulator molecules must undergo catalytic cycling on a time-scale comparable to the variation in $\mathcal{U}$ in order to propagate the time-varying signal. We therefore expect that, like robustness to leak reactions, the need to respond to time-varying signals will set a limit on how weak the coupling between $\mathcal{U}$ and $\mathcal{I}$ can be whilst retaining functional signalling.

Note, however, that neither the constraints that arise from leak reactions nor those from signal-tracking imply that a high level of fuel consumption is necessary to suppress retroactivity. Rather, an alternative trade-off is suggested: retroactivity can be reduced at low free-energy cost, but at the expense of reduced response speed and robustness of the signalling pathway. Exploring this putative trade-off in more depth, and with more detailed models of chemical reactions, will be the subject of the next chapter.

From the perspective of understanding and engineering actual biochemical systems in an experimental context, relevant questions are: how weak can the coupling be in practice before signalling is disrupted, and is the principle of relatively weak coupling applicable in natural systems? In particular, if weaker coupling is used in biology to minimize fuel turnover, one would expect different circuits to find different optimal trade-offs. Circuits with many possible leak reactions, or which need to vary on a rapid time scale, will exhibit stronger coupling (faster fuel turnover) than others. Furthermore, our analysis may explain why an activation reaction that is known to be vital for cellular function nonetheless has a slow rate.

When designing a synthetic signalling network, either from proteins or nucleic-acid based analogs, researchers could consider varying coupling strength to optimize performance,

and indeed might consider different coupling strengths for different tasks. Importantly, making strong binding weaker by mutating a binding interface is relatively simple – at least when compared to making an already-strong interface stronger. We note that it is important to make interactions with both the activating and deactivating catalysts weaker.

From a fundamental biophysics perspective, our results emphasize an important and often mis-understood point. Catalytic circuits must be dissipative (consume fuel) in order to function. But given an inherently dissipative structure it does not follow that an increased dissipation rate leads to better performance. Fundamentally, a fuel-consuming network structure is needed to ensure that catalysts overwhelmingly activate rather than deactivate their substrates (or vice versa), which is a question of relative reaction rates. The rate of dissipation, however, depends not only on these relative rates but also absolute rates, which may not help to improve circuit functionality [186].

One can analyse the push-pull motif from a cost/speed/reliability perspective as in Chapter 2. As mentioned earlier, the inactive and active states of the substrate would correspond to the different states of a bit. The notion of switching time can be likened to the amount of time spent in the intermediate catalyst-substrate complex. The notion of reliability can be mapped to the absence or inefficacy of leak-reactions in the network. The premise of this chapter seems to suggest that switching a biological bit rapidly and reliably requires low cost. This is not actually true – the point is that the push-pull motif can be an effective insulator without needing to actually switch between its states very often (by coupling the kinase and phosphatase weakly to the push-pull motif). It is known that switching a bit with arbitrary accuracy in an autonomous system costs substantial amount of energy [80]. We have not delved into the precise biophysical details to talk about trade-offs with reliability or speed of switching. We will take some steps in that direction in the next chapter.

Chapter 4

Optimizing enzymatic catalysts for rapid turnover of substrates with low sequestration

Abstract

We analyse the mechanism of enzyme-substrate catalysis from the perspective of alleviating retroactivity, whilst maintaining a minimum rate of the reaction. In particular, we ask: How should we choose binding energies of the intermediates so that the motif minimizes retroactivity, while proceeding at a certain minimal rate? Under reasonable assumptions on the rate constants, we find that there is a trade-off to choosing the optimal binding energies. Choosing high binding energies results in the motif having low retroactivity, but hampers the effective progress of the reaction. Selecting low binding energies results in high sequestration of the enzyme into complexes, giving high retroactivity. We find that the difference between the optimal binding energies of the intermediate complexes is related to the free energy difference between the products and reactants. In addition, we show that many of these observations carry over qualitatively upon relaxing the assumptions on the rate constants.

4.1 Introduction

This piece of work can be viewed as a sequel to Chapter 3. Instead of focussing on the entire push-pull motif, we consider a single enzyme-substrate reaction and analyse its dynamics. The idea is to build a slightly more complex model for enzyme-substrate catalysis, to observe the inherent constraints that arise, but are not obvious from simpler models considered in the previous chapter. Our primary emphasis will be on the states

in the middle of the switching transition (i.e. conversion of substrate from its inactive (active) form to its active (inactive) form by the catalyst).

Enzymatic catalysts form a ubiquitous part of biological pathways and analysing their mechanism is essential to understanding the exquisite behaviour exhibited by them. Enzymes are proteins that alter the rate of certain reactions and/or aid in conversion of a substrate from one form to another [50]. Even though the naive idea of catalysis gives the impression that the catalyst alters the rate of the reaction without being consumed, there is a non-zero time for which it is bound to the substrate. This sequestration of enzymatic catalysts results from the formation of catalyst-substrate intermediates and gives rise to effects like retroactivity. Even before the introduction of the term "retroactivity", ten Wolde et al. [187] discuss how back-interactions in biological systems break modularity assumptions. Though their primary interest lies in the consequences of these interactions for noise propagation, they make the above (general) point quite nicely. It is important to observe that sequestration of catalysts implies that one is using up a valuable resource – the enzyme's "time" that could be spent processing other substrates.

Natural systems seem to evolve towards minimizing sequestration [188](unless they are actively being used for things like zero-order ultrasensitivity [174, 163, 184]). But equally, in this age of synthetic biology and DNA nanotechnology, we might want to design our own systems that minimize sequestration [189]. We analyse simple models to understand how the underlying features of the enzymatic process might be tuned to achieve this goal.

The chapter is structured as follows: In section 4.2, we present the precise model of the enzyme-substrate network that we analyse, with appropriate assumptions. Initially, we assume that catalyst-substrate binding reactions are diffusion controlled, so that their rate constants are taken to be fixed. We formulate the problem of building optimal circuits (from the view-point of retroactivity) that proceed with a minimum flux at steady-state, as an optimization problem (Problem 4.2.1). More precisely, we ask: How do we choose binding energies (refer to Section 3.4 for a definition of binding energy) of the

intermediate complexes so that the circuit minimizes retroactivity, whilst proceeding at a minimum flux? In Section 4.3, we analyse this optimization problem to find that there is an energy trade-off to designing optimal circuits. Choosing very high binding energies for the intermediate complexes reduces retroactivity, but also reduces the flux through the circuit - which results in failing the minimal flux requirement. On the other hand, choosing very low binding energies implies that the system spends a large proportion of time in the intermediate states, increasing the retroactivity of the system. Hence, there is a trade-off to choosing the optimal binding energies. We show that not only are the optimal binding energies moderate, but they are strongly related to each other. In particular, the difference of optimal binding energies is connected to the intrinsic free energy difference between the products and reactants. In addition, we also show that the optimal circuit saturates the bound on the flux requirement. In Section 4.4, we relax the assumption of diffusion control for the binding rates and analyse the same optimization problem. We find that many of our observations carry over to this new regime and the optimal binding energies typically still try to remain tuned with each other.

4.2 Model and methods

In what follows, we consider discrete-state models of biochemical reactions. In particular, we model the enzyme-substrate mechanism traditionally represented by the following equation

$$Z + X \rightleftharpoons XZ \rightarrow X^* + Z \tag{4.1}$$

Here Z is the catalytic enzyme, X the inactive form of substrate, XZ the catalyst-substrate intermediate and X^* the activated form of the substrate. The two classical approaches for analysing the mechanism of enzyme-substrate catalysis are the Michaelis-Menten [190, 191] and Briggs-Haldane [192] approximations. One issue with these classic

models is that they conflate the notions of the underlying biochemical catalysis and the release of products. To get around this, we decouple the notion of catalysis and binding/unbinding using two separate reactions [193] by explicitly incorporating a second intermediate X^*Z via the reactions $XZ \to X^*Z$ and $X^*Z \to X^* + Z$. In addition, we incorporate the notion of reversibility [194, 195] by making all our reactions microscopically reversible. The model corresponding to these assumptions can be represented by the following:

$$Z + X \underset{k_{-0}}{\overset{k_{+0}}{\rightleftharpoons}} XZ \underset{k_{-\mathrm{cat}}}{\overset{k_{+\mathrm{cat}}}{\rightleftharpoons}} X^*Z \underset{k_{-1}}{\overset{k_{+1}}{\rightleftharpoons}} X^* + Z \tag{4.2}$$

The essentially catalytic parts are governed by the reactions $XZ \rightleftharpoons X^*Z$ and the binding/unbinding parts by the reactions $Z + X \rightleftharpoons XZ$ and $X^*Z \rightleftharpoons X^* + Z$. The model represented by Equation 4.2 forms the basis of our analysis in this chapter – it is the minimal model that meaningfully allows us to ask how optimal enzymes should be designed. It is important to stress that such unpacking of the model is necessary to analyse the structure of the optimal enzyme-substrate network. Handling the catalytic and binding/unbinding parts of the reaction separately allows us to make reasonable assumptions on the respective rate constants. Indeed as remarked earlier, we find that the two optimal binding energies are linked to each other by the free energy difference of the reaction.

We provide a rough sketch of the analysis of this system from the viewpoint of retroactivity - the back signal experienced by the upstream system when connected to the downstream system [92, 150, 152, 90, 95].

As is common in cellular biology [196], we assume that the reservoirs of both the inactive and active substrates X and X^* are constantly topped up throughout the course of the reaction. This means that their concentrations remain buffered to a fixed value. As a consequence, we have pseudo first-order ODEs that describe the rate of change of concentrations of enzymatic states (i.e. $k_{+0}[X]$ and $k_{-1}[X^*]$ become simple "pseudo-first order rate constants"). Consequently, each molecule of the enzyme Z can be treated

independently, and we can analyse the system using stochastic dynamics of a single enzyme [197] evolving according to a master equation. Let us denote by p_Z, p_{XZ}, p_{X^*Z} the probabilities of being in states Z, XZ and X^*Z respectively. Then we have the following master equation describing the time evolution of the probabilities corresponding to the different states.

$$\begin{pmatrix} \dot{p}_Z \\ \dot{p}_{XZ} \\ \dot{p}_{X^*Z} \end{pmatrix} = \begin{pmatrix} -(k_{-1}[X^*] + k_{+0}[X]) & k_{-0} & k_{+1} \\ k_{+0}[X] & -(k_{-0} + k_{+\text{cat}}) & k_{-\text{cat}} \\ k_{-1}[X^*] & k_{+\text{cat}} & -(k_{+1} + k_{-\text{cat}}) \end{pmatrix} \begin{pmatrix} p_Z \\ p_{XZ} \\ p_{X^*Z} \end{pmatrix} \tag{4.3}$$

Our system can also be represented by a continuous time Markov chain as shown in Figure 4.1.

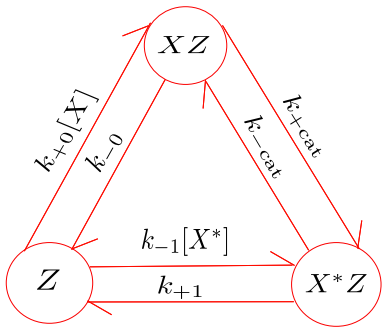


FIGURE 4.1: Markov chain corresponding to the enzyme-substrate catalysis given by Equation 4.1

Maintaining such infrastructure requires us to drive the system out of equilibrium, requiring energy consumption in the form of fuel molecules. In our case, the buffers of X and X^* act as fuel molecules [77, 159, 198, 199, 200]. In what follows, we analyse various properties of this out-of-equilibrium system in steady-state (often referred to as "non-equilibrium steady-state"). The notion of non-equilibrium steady-states has also been shown to be relevant in other contexts like kinetic proofreading [201, 202, 50] and more recently by Gunawardena et.al. [203] in the context of overcoming the "Hopfield

barrier" to sharpness of gene regulation [204].

Let us define the intrinsic free energy difference between X^* and X as $\Delta\mu$. A large magnitude of $\Delta\mu$ pushes the reaction forwards. The net free energy difference of the reaction becomes $\Delta G = \ln \frac{[X^*]}{[X]} - \Delta\mu$. It is also possible to explicitly incorporate ancillary fuel molecules (e.g ATP/ADP) under certain approximations. A possible mechanism for phosphorylation($XZ \rightleftharpoons X^*Z$) in the presence of fuel molecules is given by the following

$$XZ + ATP \rightleftharpoons XZ\text{--}ATP \rightleftharpoons X^*Z\text{--}ADP \rightleftharpoons X^*Z + ADP \tag{4.4}$$

If the binding/unbinding reactions ($XZ + ATP \rightleftharpoons XZ\text{--}ATP$ and $X^*Z\text{--}ADP \rightleftharpoons X^*Z + ADP$) are fast enough to not be treated as explicit reactions in modelling, then the mechanism looks like

$$XZ + ATP \rightleftharpoons X^*Z + ADP \tag{4.5}$$

Let us denote the intrinsic free energy difference between ATP and ADP as $\Delta\mu_{ATP,ADP}$. Therefore the net free energy difference for the reaction becomes $\Delta G = \ln \frac{[X^*]}{[X]} - \Delta\mu + \ln \frac{[ADP]}{[ATP]} - \Delta\mu_{ATP,ADP} = \ln \frac{[X^*]}{[X]} - \Delta\mu'$, where $\Delta\mu' = \Delta\mu - \ln \frac{[ADP]}{[ATP]} + \Delta\mu_{ATP,ADP}$. Thus the fuel molecules do alter the free energy, but their effect can be incorporated into an effective intrinsic free energy difference.

Let E_{XZ} and E_{X^*Z} denote the binding energies corresponding to the complexes XZ and X^*Z. These binding energies fix the ratio of the transition rates. Letting $k_BT = 1$, we get $\frac{k_{-0}}{k_{+0}} = e^{E_{XZ}}$ and $\frac{k_{+1}}{k_{-1}} = e^{E_{X^*Z}}$ and $\frac{k_{+\text{cat}}}{k_{-\text{cat}}} = e^{-E_{X^*Z}+E_{XZ}+\Delta\mu}$. The binding reactions $Z + X \rightleftharpoons XZ$ and $X^*Z \rightleftharpoons X^* + Z$ require association of diffusing molecules. We assume that they are diffusion-controlled [205, 206] (i.e. the on-rate is fixed). In particular, we set $k_{+0} = k_0$ and $k_{-1} = k_1$. In the language of [207], fixing the on-rates corresponds to the reactions being "backward" and "forward" labile, respectively. We could in principle make the catalytic part of our network $XZ \rightleftharpoons X^*Z$ either forward or backward labile; however as we show in the Section 4.3 that this gives rise to pathological

results for the question we ask. One way to get around this problem is to upper bound both the forward and backward catalytic rate constants whilst still maintaining the ratio $\frac{k_{+\text{cat}}}{k_{-\text{cat}}} = e^{-E_{X^*Z}+E_{XZ}+\Delta\mu}$. Therefore, we choose $k_{+\text{cat}} = k_{\text{cat}} \min(1, e^{-E_{X^*Z}+E_{XZ}+\Delta\mu+\Delta G_c})$ and $k_{-\text{cat}} = k'_{\text{cat}} \min(1, e^{-E_{XZ}+E_{X^*Z}-\Delta\mu-\Delta G_c})$, where $k'_{\text{cat}} = k_{\text{cat}} e^{\Delta G_c}$ for some $\Delta G_C \in \mathbb{R}$. ΔG_C represents an offset between the maximum forward and backward catalytic rate constants. Figure 4.2 gives a graphical representation of the rates as a function of their respective binding energies. Our system can therefore be represented as

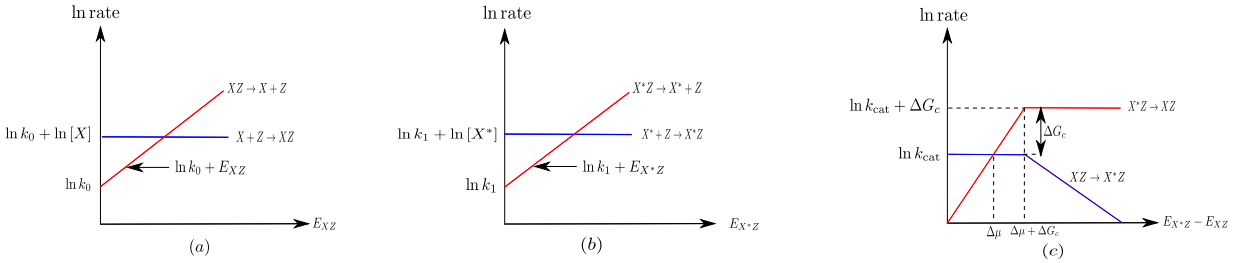


FIGURE 4.2: Graphical illustration of transition rates of the Markov chain as a function of respective binding energies. Binding rate constants are assumed to be fixed due to diffusion. Catalytic rate constants are bounded, whilst maintaining $\frac{k_{+\text{cat}}}{k_{-\text{cat}}} = e^{-E_{X^*Z}+E_{XZ}+\Delta\mu}$.

$$Z + X \underset{k_0 e^{E_{XZ}}}{\overset{k_0}{\rightleftharpoons}} XZ \underset{k_{\text{cat}} e^{\Delta G_c} \min(1, e^{-E_{XZ}+E_{X^*Z}-\Delta\mu-\Delta G_c})}{\overset{k_{\text{cat}} \min(1, e^{-E_{X^*Z}+E_{XZ}+\Delta\mu+\Delta G_c})}{\rightleftharpoons}} X^*Z \underset{k_1}{\overset{k_1 e^{E_{X^*Z}}}{\rightleftharpoons}} X^* + Z \tag{4.6}$$

Recalling Equation 3.3 from chapter 3, the retroactivity metric is $\mathcal{R} = \left|1 - \frac{[Z^{\text{ss}}]}{[Z^{\text{ss}}_{\mathcal{D}\to\emptyset}]}\right|$, where $\mathcal{D}$ denotes the downstream system. Starting with an initial concentration of Z, say $[Z_{\text{tot}}]$, the retroactivity metric translates to

$$\mathcal{R} = 1 - \frac{[Z^{\text{ss}}]}{[Z_{\text{tot}}]} \tag{4.7}$$

Then, the flux (or net flow across any reaction) in steady-state corresponding to Equation 4.2 is given by

$$\begin{aligned}
\Psi &= k_{+0}[Z^{\text{ss}}][X^{\text{ss}}] - k_{-0}[XZ^{\text{ss}}]e^{E_{XZ}} \\
&= k_{\text{cat}} \min(1, e^{-E_{X^*Z}+E_{XZ}+\Delta\mu+\Delta G_c})[XZ^{\text{ss}}] - k_{\text{cat}} e^{\Delta G_c} \min(1, e^{-E_{XZ}+E_{X^*Z}-\Delta\mu-\Delta G_c})[X^*Z^{\text{ss}}] \\
&= k_{+1}[X^*Z^{ss}]e^{E_{X^*Z}} - k_{-1}[X^{*\text{ss}}][Z^{\text{ss}}]
\end{aligned} \tag{4.8}$$

Our primary goal in this chapter is to design efficient enzyme-substrate circuits that

proceed at a minimum flux whilst consuming as little of the enzyme as possible (i.e. minimize retroactivity). More specifically,

Our main question is: How should binding energies E_{XZ} and E_{X^*Z} be chosen to minimize retroactivity $\mathcal{R}$, whilst proceeding with a minimal flux Ψ_0 through the circuit?

Let us call the optimal binding energies as E_{XZ}^{opt} and $E_{X^*Z}^{\text{opt}}$, the optimal flux Ψ_{opt} and the required minimum flux to be $\Psi_0 \in \mathbb{R}_{>0}$. Then, we can state the problem in the following form

Problem 4.2.1.

$$(E_{XZ}^{\text{opt}}, E_{X^*Z}^{\text{opt}}) = \arg \min_{E_{XZ}, E_{X^*Z}} \mathcal{R}$$
$$s.t.\ \Psi_{\text{opt}} \geq \Psi_0$$

Sivak and Brown [207] have also considered optimizing flux albeit in a different context; there the authors decide on splitting the free energy available to maximise the flux in the circuit. They find that maximizing the flux results in having fewer intermediate states and that uneven splitting of free energy is optimal for a fixed amount of transition states. Our work differs from theirs in the sense that they are exclusively looking to maximize flux, while we are concerned with maximizing flux whilst minimizing the time spent in the intermediate states.

4.3 Optimal binding energies and flux

It is not hard to see that E_{XZ}^{opt} is finite - increasing E_{XZ}^{opt} increases the unbinding rate ($XZ \to X + Z$), which decreases the flux exponentially. Since we have a minimum flux requirement, we cannot make E_{XZ}^{opt} arbitrarily large. Similarly, $E_{X^*Z}^{\text{opt}}$ cannot be made arbitrarily large since it essentially prevents the reaction from progressing further. But it is not a priori clear how E_{XZ}^{opt} and $E_{X^*Z}^{\text{opt}}$ should be related. As we shall show in this

chapter, $E^{\text{opt}}_{X^*Z} = E^{\text{opt}}_{XZ} + \Delta\mu + \Delta G_c$. In addition, we also claim that $\Psi_{\text{opt}} = \Psi_0$. We first present a rough biophysical argument to justify both these claims and then present more formal proofs using the theory of continuous time Markov chains.

4.3.1 Biophysical argument

Consider the free energy profile corresponding to our problem, as depicted in Figure 4.3. Naively, one might expect to get a nice "ladder" of (roughly) evenly-spaced decreasing intrinsic free energies. But in general, we do not see this effect; the intrinsic free energies of the intermediate states get pushed as high as possible to avoid sequestration whilst maintaining the desired flux. With respect to this figure, the retroactivity metric is proportional to the fraction of time spent in the intermediate metastable states XZ and X^*Z (or is inversely proportional to the fraction of time spent in the free Z state). The flux is directly proportional to both the fraction of time spent in the free Z state and the difference in probabilities of going from the start state to the end state and back (in this case, the start and end states are the free Z states).

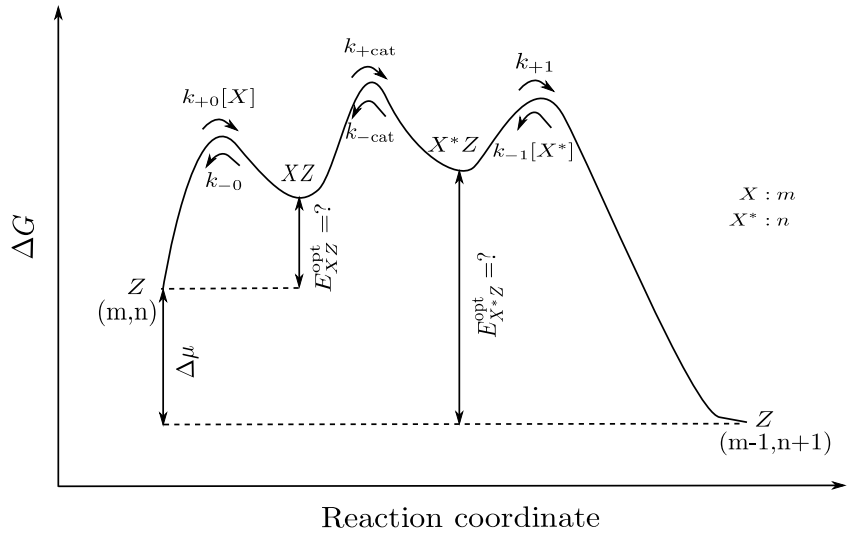


FIGURE 4.3: Free energy profile of the reaction. The metastable states denote the intermediates XZ and X^*Z. Our primary question is: How to choose binding energies E_{XZ} and E_{X^*Z} such that it minimizes retroactivity, whilst maintaining a minimum flux.

Let us assume for contradiction that $E_{X^*Z} \neq E_{XZ} + \Delta\mu + \Delta G_c$. Then, we have two cases

1. $E_{X^*Z} < E_{XZ} + \Delta\mu + \Delta G_c$: Equation 4.6 becomes

$$Z + X \underset{k_0 e^{E_{XZ}}}{\overset{k_0}{\rightleftharpoons}} XZ \underset{k_{\text{cat}} e^{-E_{XZ}+E_{X^*Z}-\Delta\mu}}{\overset{k_{\text{cat}}}{\rightleftharpoons}} X^*Z \underset{k_1}{\overset{k_1 e^{E_{X^*Z}}}{\rightleftharpoons}} X^* + Z \tag{4.9}$$

Comparing Equations 4.9 and 4.2, we get that increasing E_{X^*Z} increases both k_{+1} and $k_{-\text{cat}}$ proportionately. Keeping in mind Figure 4.3, this amounts to making the well corresponding to state X^*Z shallower. As a consequence, one is more likely to pop out of the X^*Z state quickly; but the relative probabilities of hopping to its neighbouring states remain unaffected. Therefore, one has to spend a smaller proportion of time in the X^*Z state and a large proportion of time in the free Z state. By Equation 3.3, since retroactivity is inversely proportional to the amount of time spent in the free Z state, we get that the retroactivity decreases with increase in E_{X^*Z}. In addition, since the relative probabilities of hopping between states remain unchanged, the flux increases with increase in E_{X^*Z}. So, given any configuration of binding energies obeying $E_{X^*Z} < E_{XZ} + \Delta\mu + \Delta G_c$, one can always increase E_{X^*Z} to reach a more optimal state.

If we don't bound the backward catalytic rate constant (i.e. if the system is 100% backwards labile on the catalytic transition) and let our system evolve according to

$$Z + X \underset{k_0 e^{E_{XZ}}}{\overset{k_0}{\rightleftharpoons}} XZ \underset{k_{\text{cat}} e^{-E_{XZ}+E_{X^*Z}-\Delta\mu}}{\overset{k_{\text{cat}}}{\rightleftharpoons}} X^*Z \underset{k_1}{\overset{k_1 e^{E_{X^*Z}}}{\rightleftharpoons}} X^* + Z \tag{4.10}$$

we can keep on increasing E_{X^*Z} (by the argument above) to get systems with lower flux and higher retroactivity. Thus, $E^{\text{opt}}_{X^*Z}$ is arbitrarily high for this system, which is clearly unphysical.

2. $E_{X^*Z} > E_{XZ} + \Delta\mu + \Delta G_c$, then Equation 4.6 becomes

$$Z + X \underset{k_0 e^{E_{XZ}}}{\overset{k_0}{\rightleftharpoons}} XZ \underset{k_{\text{cat}}}{\overset{k_{\text{cat}} e^{E_{XZ}-E_{X^*Z}+\Delta\mu}}{\rightleftharpoons}} X^*Z \underset{k_1}{\overset{k_1 e^{E_{X^*Z}}}{\rightleftharpoons}} X^* + Z \tag{4.11}$$

Arguing as above, one can conclude in this case that increasing E_{XZ} decreases retroactivity and increases the flux of the motif, and hence cannot be optimal.

Similar to the first case, if we do not bound the forward catalytic rate constant (i.e. if system is 100% forwards labile on the catalytic transition) and let our system evolve according to

$$Z + X \underset{k_0 e^{E_{XZ}}}{\overset{k_0}{\rightleftharpoons}} XZ \underset{k_{\text{cat}}}{\overset{k_{\text{cat}} e^{E_{XZ} - E_{X^*Z} + \Delta\mu}}{\rightleftharpoons}} X^*Z \underset{k_1}{\overset{k_1 e^{E_{X^*Z}}}{\rightleftharpoons}} X^* + Z \tag{4.12}$$

we can keep on increasing E_{XZ} to get systems with lower flux and higher retroactivity. Thus, E_{XZ}^{opt} is arbitrarily high for this system, which is pathological. This justifies the energetic-dependence of our catalytic rates.

As a result, if $E_{X^*Z} \neq E_{XZ} + \Delta\mu + \Delta G_c$, one can increase either E_{XZ} or E_{X^*Z} to reach a more optimal state. Therefore, we get $E_{X^*Z}^{\text{opt}} = E_{XZ}^{\text{opt}} + \Delta\mu + \Delta G_c$ and we have $k_{+\text{cat}} = k_{\text{cat}}$ and $k_{-\text{cat}} = k_{\text{cat}} e^{\Delta G_c}$. It is interesting to observe what happens if we increase both E_{XZ} and E_{X^*Z} by the same amount when $E_{X^*Z} = E_{XZ} + \Delta\mu + \Delta G_c$. With respect to Figure 4.3, increasing E_{XZ} and E_{X^*Z} makes it easier to pop out of both wells into the free Z state. Since the catalytic rate constants $k_{+\text{cat}} = k_{\text{cat}}$ and $k_{-\text{cat}} = k_{\text{cat}} e^{\Delta G_c}$ are fixed, it also increases the probability of going from XZ and X^*Z to the free Z state. This reduces the retroactivity of the motif. As a consequence, if the optimal flux did not saturate the bound on the constraint i.e. if $\Psi_{\text{opt}} > \Psi_0$, one could increase both the binding energies by a sufficiently small amount so that it still satisfies the constraint on the flux with lower retroactivity. Repeatedly performing this operation of increasing both binding energies by the same amount will eventually suppress flux and the the optimal flux will always saturate the bound on the flux constraint i.e. $\Psi_{\text{opt}} = \Psi_0$.

4.3.2 Formal proof using Markov chains

To make these arguments more formal, we require machinery from the theory of Markov chains. The continuous time Markov chain corresponding to our network is shown in Figure 4.4. For convenience, we will use the following notation. Let us assume that

we start from the state Z and let the Markov chain evolve until it returns back to Z. Let us denote by p_i^{ss}, the steady-state probability of being state i; $\langle \tau_i \rangle$, the expected lifetime of state i; $\langle \tau_Z^r \rangle$, the expected return time of Z; $\langle \tau_{i \to j} \rangle$, the expected time to reach state j given that you started in state i and by $p_{i,j}$: the probability of transition from state i to j. By the memoryless property of continuous time Markov chains, the lifetime of a state is exponentially distributed with parameter equal to the total rate of outward transition from that state, implying that $\langle \tau_i \rangle = \frac{1}{\lambda_i}$, where λ_i is the rate of outward transition from state i. The steady-state probabilities of the three states in the Markov chain are $p_Z^{\text{ss}} = \frac{[Z^{\text{ss}}]}{[Z_{\text{tot}}]}, p_{XZ}^{\text{ss}} = \frac{[XZ^{\text{ss}}]}{[Z_{\text{tot}}]}$ and $p_{X^*Z}^{\text{ss}} = \frac{[X^*Z^{\text{ss}}]}{[Z_{\text{tot}}]}$. Since we have an irreducible Markov chain, using [208, Theorem 3.8.1] and Equation 4.7, we get $\mathcal{R} = 1 - \frac{[Z^{\text{ss}}]}{[Z_{\text{tot}}]} = 1 - p_Z^{\text{ss}} \xrightarrow{a.s.}{}^{2} 1 - \frac{\langle \tau_Z \rangle}{\langle \tau_Z^r \rangle}$. Using $\langle \tau_Z^r \rangle = \langle \tau_Z \rangle + p_{Z,XZ} \langle \tau_{XZ \to Z} \rangle + p_{Z,X^*Z} \langle \tau_{X^*Z \to Z} \rangle$, we get $\mathcal{R} \xrightarrow{a.s.} 1 - \frac{\langle \tau_Z \rangle}{\langle \tau_Z \rangle + p_{Z,XZ} \langle \tau_{XZ \to Z} \rangle + p_{Z,X^*Z} \langle \tau_{X^*Z \to Z} \rangle}$. Noting that $p_{Z,XZ} = k_{+0} \langle \tau_Z \rangle$ and $p_{Z,X^*Z} = k_{-1} \langle \tau_Z \rangle$, we get

$$\mathcal{R} \xrightarrow{a.s.} 1 - \frac{1}{1 + k_{+0} \langle \tau_{XZ \to Z} \rangle + k_{-1} \langle \tau_{X^*Z \to Z} \rangle} \tag{4.13}$$

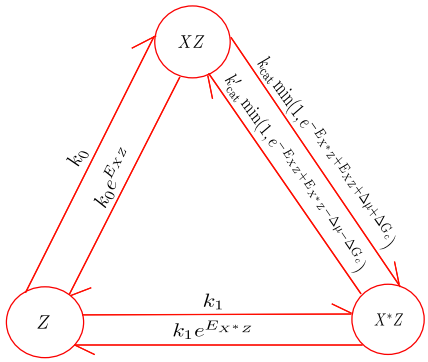


FIGURE 4.4

Observing that $\langle \tau_{XZ \to Z} \rangle = \langle \tau_{XZ} \rangle + p_{XZ,X^*Z} \langle \tau_{X^*Z \to Z} \rangle$ and $\langle \tau_{X^*Z \to Z} \rangle = \langle \tau_{X^*Z} \rangle + p_{X^*Z,XZ} \langle \tau_{XZ \to Z} \rangle$, we get that $\langle \tau_{XZ \to Z} \rangle = \frac{\langle \tau_{X^*Z} \rangle + p_{X^*Z,XZ} \langle \tau_{XZ} \rangle}{1 - p_{X^*Z,XZ} p_{XZ,X^*Z}}$ and $\langle \tau_{X^*Z \to Z} \rangle = \frac{\langle \tau_{XZ} \rangle + p_{XZ,X^*Z} \langle \tau_{X^*Z} \rangle}{1 - p_{X^*Z,XZ} p_{XZ,X^*Z}}$. Equation 4.13 then becomes

$$\mathcal{R} = 1 - \frac{1}{1 + k_{+0} \left(\frac{\langle \tau_{X^*Z} \rangle + p_{X^*Z,XZ} \langle \tau_{XZ} \rangle}{1 - p_{X^*Z,XZ} p_{XZ,X^*Z}} \right) + k_{-1} \left(\frac{\langle \tau_{XZ} \rangle + p_{XZ,X^*Z} \langle \tau_{X^*Z} \rangle}{1 - p_{X^*Z,XZ} p_{XZ,X^*Z}} \right)} \tag{4.14}$$

[2] almost surely

We will now formally show that $E^{\text{opt}}_{X^*Z} = E^{\text{opt}}_{XZ} + \Delta\mu + \Delta G_c$. For contradiction, assume not. Then, we have two cases:

1. $E^{\text{opt}}_{X^*Z} < E^{\text{opt}}_{XZ} + \Delta\mu + \Delta G_c$. This corresponds precisely to the situation in Figure 4.5. Eliminating $[XZ^{\text{ss}}]$ and $[X^*Z^{\text{ss}}]$ from system 4.8 with $k_{+\text{cat}} = k_{\text{cat}}, k_{-\text{cat}} = k_{\text{cat}} e^{E^{\text{opt}}_{X^*Z} - E^{\text{opt}}_{XZ} - \Delta\mu}$ and noting that $[Z^{\text{ss}}] = [Z_{\text{tot}}](1 - \mathcal{R})$, we get that the flux in steady-state is given by

$$\Psi = \frac{(1-\mathcal{R})k_0 k_1 k_{\text{cat}}[Z_{\text{tot}}]([X]e^{\Delta\mu} - [X^*])}{k_0 k_1 e^{\Delta\mu + E^{\text{opt}}_{XZ}} + k_0 k_{\text{cat}} + k_1 k_{\text{cat}} e^{\Delta\mu}} \tag{4.15}$$

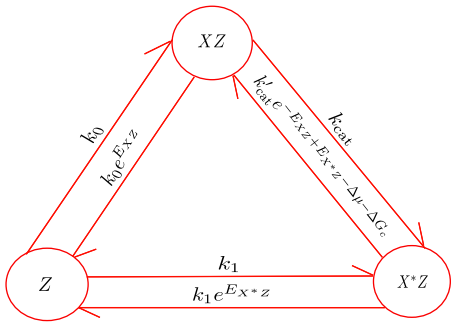


FIGURE 4.5

Note that the transition probabilities $p_{XZ,X^*Z} = \frac{k_{\text{cat}}}{k_0 e^{E^{\text{opt}}_{XZ}} + k_{\text{cat}}}$ and $p_{X^*Z,XZ} = \frac{k_{\text{cat}} e^{-E^{\text{opt}}_{XZ} - \Delta\mu}}{k_1 + k_{\text{cat}} e^{-E^{\text{opt}}_{XZ} - \Delta\mu}}$ remain unchanged as $E^{\text{opt}}_{X^*Z}$ increases. As one increases $E^{\text{opt}}_{X^*Z}$, the expected lifetime of $X^*Z, \langle\tau_{X^*Z}\rangle = \frac{1}{e^{E^{\text{opt}}_{X^*Z}}(k_1 + k_{\text{cat}} e^{-E^{\text{opt}}_{XZ} - \Delta\mu})}$ decreases, while $\langle\tau_Z\rangle = \frac{1}{k_0 + k_1}$ and $\langle\tau_{XZ}\rangle = \frac{1}{k_{\text{cat}} + k_0 e^{E^{\text{opt}}_{XZ}}}$ remain unchanged. By Equation 4.14, we get that $\mathcal{R}$ decreases as $E^{\text{opt}}_{X^*Z}$ increases. Using Equation 4.15, we get that flux Ψ increases as $E^{\text{opt}}_{X^*Z}$ increases. Therefore, increasing $E^{\text{opt}}_{X^*Z}$ by an amount $\Delta E > 0$ gives a tuple $(E^{\text{opt}}_{XZ}, E^{\text{opt}}_{X^*Z} + \Delta E)$ that has flux greater than Ψ_0 but lower retroactivity, contradicting the optimality of the tuple $(E^{\text{opt}}_{XZ}, E^{\text{opt}}_{X^*Z})$.

2. $E^{\text{opt}}_{X^*Z} > E^{\text{opt}}_{XZ} + \Delta\mu + \Delta G_c$. This corresponds exactly to the situation in Figure 4.6. Setting $k_{+\text{cat}} = k_{\text{cat}} e^{-E^{\text{opt}}_{X^*Z} + E^{\text{opt}}_{XZ} + \Delta\mu}, k_{-\text{cat}} = k_{\text{cat}}$ and eliminating $[XZ^{\text{ss}}]$ and $[X^*Z^{\text{ss}}]$ from system 4.8 with $[Z^{\text{ss}}] = [Z_{\text{tot}}](1-\mathcal{R})$, we find that the steady-state flux is given

by

$$\Psi = \frac{(1-\mathcal{R})k_0 k_1 k_{\text{cat}}[Z_{\text{tot}}]([X]e^{\Delta\mu} - [X^*])}{k_0 k_1 e^{E^{\text{opt}}_{X^*Z}} + k_0 k_{\text{cat}} + k_1 k_{\text{cat}} e^{\Delta\mu}} \tag{4.16}$$

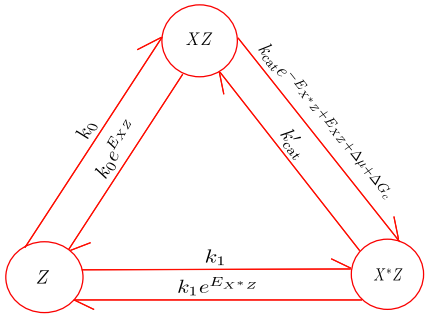


FIGURE 4.6

Since $\langle\tau_{XZ}\rangle = \frac{1}{e^{E^{\text{opt}}_{XZ}}(k_0 + k_{\text{cat}}e^{-E^{\text{opt}}_{X^*Z}+\Delta\mu+\Delta G_c})}$, we get that $\langle\tau_{XZ}\rangle$ decreases with increase in E^{opt}_{XZ}, while the transition probabilities $p_{XZ,X^*Z} = \frac{k_{\text{cat}}e^{-E^{\text{opt}}_{X^*Z}+\Delta\mu+\Delta G_c}}{k_0 + k_{\text{cat}}e^{-E^{\text{opt}}_{X^*Z}+\Delta\mu+\Delta G_c}}, p_{X^*Z,XZ} = \frac{k'_{\text{cat}}}{k'_{\text{cat}}+k_1 e^{E^{\text{opt}}_{X^*Z}}}$ and the lifetimes of other states remain unchanged. Using Equation 4.14, we get that $\mathcal{R}$ decreases as E^{opt}_{XZ} increases. Using Equation 4.16, we get that flux Ψ increases as E^{opt}_{XZ} increases. Increasing E^{opt}_{XZ} by an amount $\Delta E > 0$ gives a tuple $(E^{\text{opt}}_{XZ} + \Delta E, E^{\text{opt}}_{X^*Z})$ that has flux greater than Ψ_0 but lower retroactivity, contradicting the optimality of the tuple $(E^{\text{opt}}_{XZ}, E^{\text{opt}}_{X^*Z})$.

Our observations are summarised in Figure 4.7.

Next, we show that $\Psi_{\text{opt}} = \Psi_0$. For contradiction, assume that $\Psi_{\text{opt}} > \Psi_0$. We first prove that $\mathcal{R}$ decreases if you increase both E^{opt}_{XZ} and $E^{\text{opt}}_{X^*Z}$ by the same amount. Since the optimal binding energies satisfy $E^{\text{opt}}_{X^*Z} = E^{\text{opt}}_{XZ} + \Delta\mu + \Delta G_c$, our system can be represented by the Markov chain in Figure 4.8. Note that increasing both E^{opt}_{XZ} and $E^{\text{opt}}_{X^*Z}$ by the same amount decreases both $p_{XZ\to X^*Z} = \frac{k_{\text{cat}}}{k_{\text{cat}}+k_0 e^{E^{\text{opt}}_{XZ}}}$ and $p_{X^*Z\to XZ} = \frac{k_{\text{cat}}}{k'_{\text{cat}}+k_0 e^{E^{\text{opt}}_{X^*Z}}}$; further the expected lifetimes $\langle\tau_{XZ}\rangle = \frac{1}{k_0 e^{E^{\text{opt}}_{XZ}}+k_{\text{cat}}}, \langle\tau_{X^*Z})\rangle = \frac{1}{k_1 e^{E^{\text{opt}}_{X^*Z}}+k_{\text{cat}}}$ decrease, while $\langle\tau_Z\rangle = \frac{1}{k_0+k_1}$ remains unaffected. Therefore, using Equation 4.14 we get that $\mathcal{R}$ decreases as we increase both E^{opt}_{XZ} and $E^{\text{opt}}_{X^*Z}$ by the same amount. As a consequence, one can increase both the binding energies E^{opt}_{XZ} and $E^{\text{opt}}_{X^*Z}$ by a sufficiently small amount say

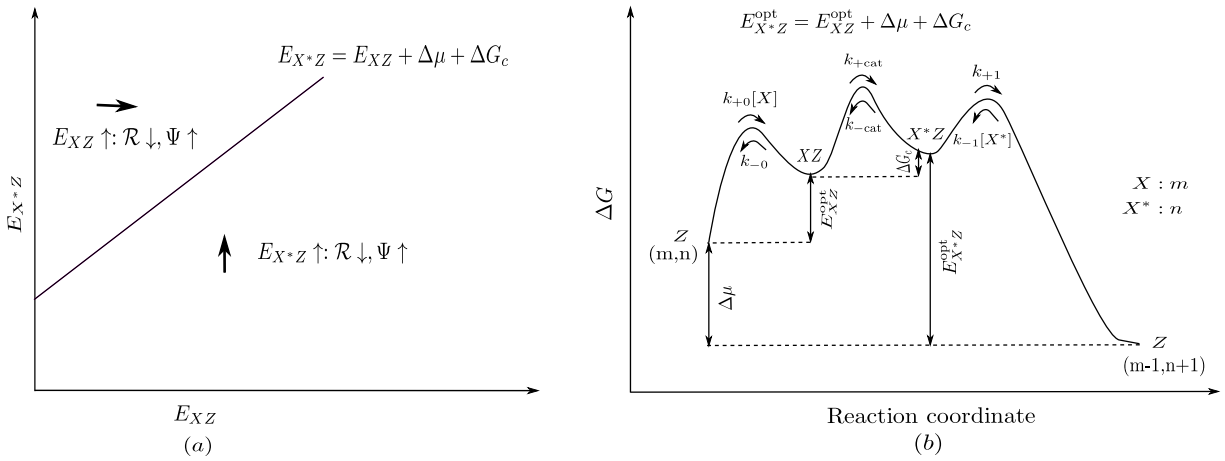


FIGURE 4.7: (a) In the regime $E_{X^*Z} < E_{XZ} + \Delta\mu + \Delta G_c$, one can increase E_{X^*Z} to get a system having lower retroactivity and higher flux. In the regime $E_{X^*Z} > E_{XZ} + \Delta\mu + \Delta G_c$, one can increase E_{XZ} to get a system having lower retroactivity and higher flux. As a consequence, optimal binding energies are constrained to lie on $E^{\text{opt}}_{X^*Z} = E^{\text{opt}}_{XZ} + \Delta\mu + \Delta G_c$. (b) Free energy landscape corresponding to the optimal reaction.

$E_0 > 0$, to get a tuple $(E^{\text{opt}}(XZ) + E_0, E^{\text{opt}}(X^*Z) + E_0)$ that has flux greater than Ψ_0, but has lower retroactivity than the tuple $(E^{\text{opt}}_{XZ}, E^{\text{opt}}_{X^*Z})$, contradicting the optimality of $(E^{\text{opt}}_{XZ}, E^{\text{opt}}_{X^*Z})$. If we keep on increasing both the binding energies, eventually the flux will start decreasing, ultimately dropping below Ψ_0. Therefore, we have $\Psi_{\text{opt}} = \Psi_0$.

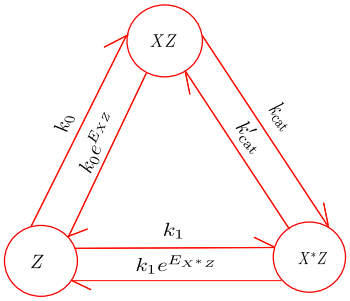


FIGURE 4.8

We have solved Problem 4.2.1 for the network corresponding to Figure 4.4 using numerical optimization to verify our observations. Figure 4.9 gives a contour plot of flux and retroactivity as a function of the binding energies E_{XZ} and E_{X^*Z}. Embedded on them are the optimal binding energies obtained as a result of numeric optimization for a range of target fluxes. Increasing E_{XZ} increases flux up to a certain point; then the flux starts to decrease with further increase in E_{XZ}. A similar behaviour is observed for E_{X^*Z}. The optimal binding energies for different values of target flux lie on a straight

line, consonant with our observations. The fact that the optimal binding energies obey precisely $E^{\text{opt}}_{X^*Z} = E^{\text{opt}}_{XZ} + \Delta\mu + \Delta G_c$ is verified numerically in Fig 4.9. In addition, the optimal flux saturates the constraint on the target flux i.e. $\Psi_{\text{opt}} = \Psi_0$.

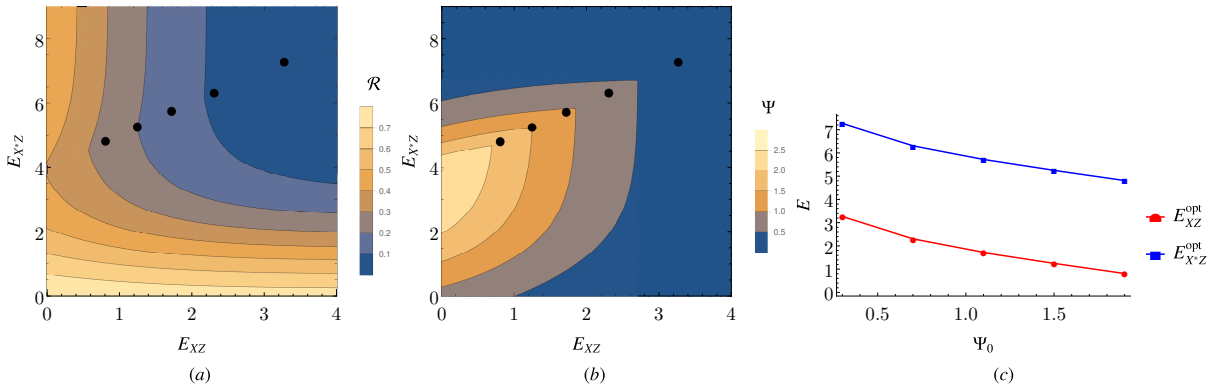


FIGURE 4.9: (a) Retroactivity contour plot as a function of binding energies. (b) Flux contour as a function of binding energies. (c) Optimal binding energies as a function of target flux. The black dots indicate optimal binding energies corresponding to certain target fluxes. Parameters used for optimization: $Z_{\text{tot}} = 10, a = 1, b = 3, k_0 = 1, k_1 = 1, k_{\text{cat}} = 1, \Delta\mu = 3, \Delta G_c = 1$. As is evident from part (c), the optimal binding energies obey $E^{\text{opt}}_{X^*Z} = E^{\text{opt}}_{XZ} + \Delta\mu + \Delta G_c$.

The observations we made above suggests that there is an energy trade-off to designing optimal (minimally retroactive) biological circuits. Ideally, one would want binding energies to be arbitrarily high, so that the system spends very little time in the intermediate states, thereby reducing the retroactivity of the motif. However, one cannot increase the binding energies arbitrarily, since they start affecting the flux of the circuit. We wish to design biological circuits that possess reactions happening at a certain rate - which constrains the binding energies to certain values. Interestingly, we find that the difference between the optimal binding energies is a constant - $(\Delta\mu + \Delta G_c)$, where $\Delta\mu$ is the free energy difference between X^* and X and ΔG_c is the offset. Further, this difference does not depend upon the concentrations of $[X]/[X^*]$, which is slightly counter-intuitive. One might have expected that a high amount of $[X^*]$ would coerce us to increase E_{X^*Z} relative to E_{XZ} to avoid spending a lot of time in the X^*Z state; however our results show that the difference $E_{X^*Z} - E_{XZ}$ is independent of $[X]/[X^*]$.

4.4 Effect of non-diffusion controlled binding rates

Until now, we assumed that the binding rate constants k_{+0}, k_{-1} are diffusion-controlled and cannot be changed arbitrarily, but placed no constraint on the dissociation rate constants. We now relax this assumption by limiting the maximum rate of dissociation, and assume that beyond a certain point, association is reduced instead. This assumption is incorporated in the model represented by Equation 4.1 with $k_{+\text{cat}} = k_{\text{cat}} \min(1, e^{-E_{X^*Z}+E_{XZ}+\Delta\mu+\Delta G_c})$, $k_{-\text{cat}} = k'_{\text{cat}} \min(1, e^{-E_{XZ}+E_{X^*Z}-\Delta\mu-\Delta G_c})$, $k_{+0} = k_0 \min(1, e^{-(E_{XZ}-E_1)})$, $k_{-0} = k'_0 \min(1, e^{E_{XZ}-E_1})$, $k_{+1} = k'_1 \min(1, e^{E_{X^*Z}-E_2})$, $k_{-1} = k_1 \min(1, e^{-(E_{X^*Z}-E_2)})$, where $k'_{\text{cat}} = k_{\text{cat}} e^{\Delta G_c}$, $k'_0 = k_0 e^{E_1}$ and $k'_1 = k_1 e^{E_2}$. A schematic illustrating the distinct regions of this model is shown in Figure 4.10a. This phase diagram demarcates various regions in the space of binding energies, along with their respective relationships with retroactivity and flux. These relationships help us determine the structure and geometry of the optimal binding energies, which would have been hard to follow otherwise.

We have already analysed regions I and II in the previous section. By similar approaches, we deduce which directions on the E_{XZ} and E_{X^*Z} plane are guaranteed to decrease/increase flux and retroactivity in the different regions shown in Figure 4.10a. Our results are plotted in Figure 4.10b. The continuous time Markov chains corresponding to each of these regions is depicted in Figure 4.11.

- Region VI: Employing the same argument that we used for Region I, one can show that increasing E_{X^*Z} decreases retroactivity but increases flux in this region.
- Region III: Arguing as we did for Region II, one can show that increasing E_{XZ} decreases retroactivity but increases flux in this region.
- Region V: Note that the transition probabilities $p_{XZ,X^*Z} = \frac{k_{\text{cat}}}{k'_0+k_{\text{cat}}}$, $p_{X^*Z,XZ} = \frac{k'_{\text{cat}} e^{-E_{XZ}+E_{X^*Z}-\Delta\mu-\Delta G_c}}{k'_1+k'_{\text{cat}} e^{-E_{XZ}+E_{X^*Z}-\Delta\mu-\Delta G_c}}$ and the expected lifetimes X^*Z, $\langle \tau^l_{X^*Z} \rangle = \frac{1}{k'_1+k'_{\text{cat}} e^{-E_{XZ}+E_{X^*Z}-\Delta\mu-\Delta G_c}}$,

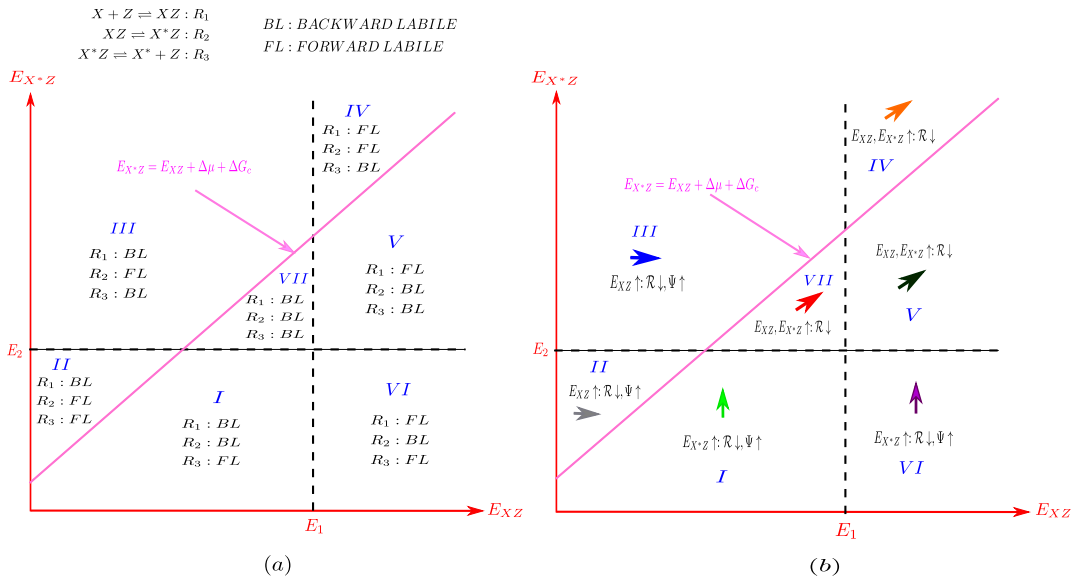


FIGURE 4.10: Behaviour of binding energies in the setting where binding reactions may or may not be diffusion controlled. (a) The space of binding energies can be divided into multiple regions in which different reactions are forward and backward labile. (b) In regions II and III, one can always increase E_{XZ} to get to a state with higher flux and lower retroactivity. In regions I and VI, one can always increase E_{X^*Z} to get to a state with higher flux and lower retroactivity. Consequently, optimal binding energies lie either on $E^{\text{opt}}_{X^*Z} = E^{\text{opt}}_{XZ} + \Delta\mu + \Delta G_c$ or in the regions IV, V, VII. In regions IV, V, VII, one can increase both E_{XZ} and E_{X^*Z} by the same amount to reach a state with lower retroactivity.

$\langle \tau^l_{XZ} \rangle = \frac{1}{k'_0 + k_{\text{cat}}}$ remain unchanged, while k_{+0} and k_{-1} decrease as we increase both E_{XZ} and E_{X^*Z} by the same amount. By Equation 4.14, this decreases the retroactivity $\mathcal{R}$ of the system.

- Region IV: Arguing as we did for Region V, one can show that $\mathcal{R}$ decreases as we increase both E_{XZ} and E_{X^*Z} by the same amount in this region.

- Region VII: Note that the transition probability $p_{XZ,X^*Z} = \frac{k_{\text{cat}}}{k'_0 e^{E_{XZ} - E_1} + k_{\text{cat}}}$ and the expected lifetime $\langle \tau^l_{XZ} \rangle = \frac{1}{k_{\text{cat}} + k'_0 e^{E_{XZ} - E_1}}$ decreases, while $p_{X^*Z,XZ} = \frac{k'_{\text{cat}} e^{-E_{XZ} + E_{X^*Z} - \Delta\mu - \Delta G_c}}{k'_1 + k'_{\text{cat}} e^{-E_{XZ} + E_{X^*Z} - \Delta\mu - \Delta G_c}}$ and $\langle \tau^l_{X^*Z} \rangle = \frac{1}{k'_1 + k'_{\text{cat}} e^{-E_{XZ} + E_{X^*Z} - \Delta\mu - \Delta G_c}}$ remain unchanged if we increase both E_{XZ} and E_{X^*Z} by the same amount. In addition k_{-1} decreases with increase in E_{X^*Z}. Therefore, by Equation 4.14, retroactivity decreases if we increase both E_{XZ} and E_{X^*Z} by the same amount.

As a consequence, the optimal binding energies either satisfy $E^{\text{opt}}_{X^*Z} = E^{\text{opt}}_{XZ} + \Delta\mu + \Delta G_c$ or lie in regions IV, V, VII. We have verified this numerically for some sample target

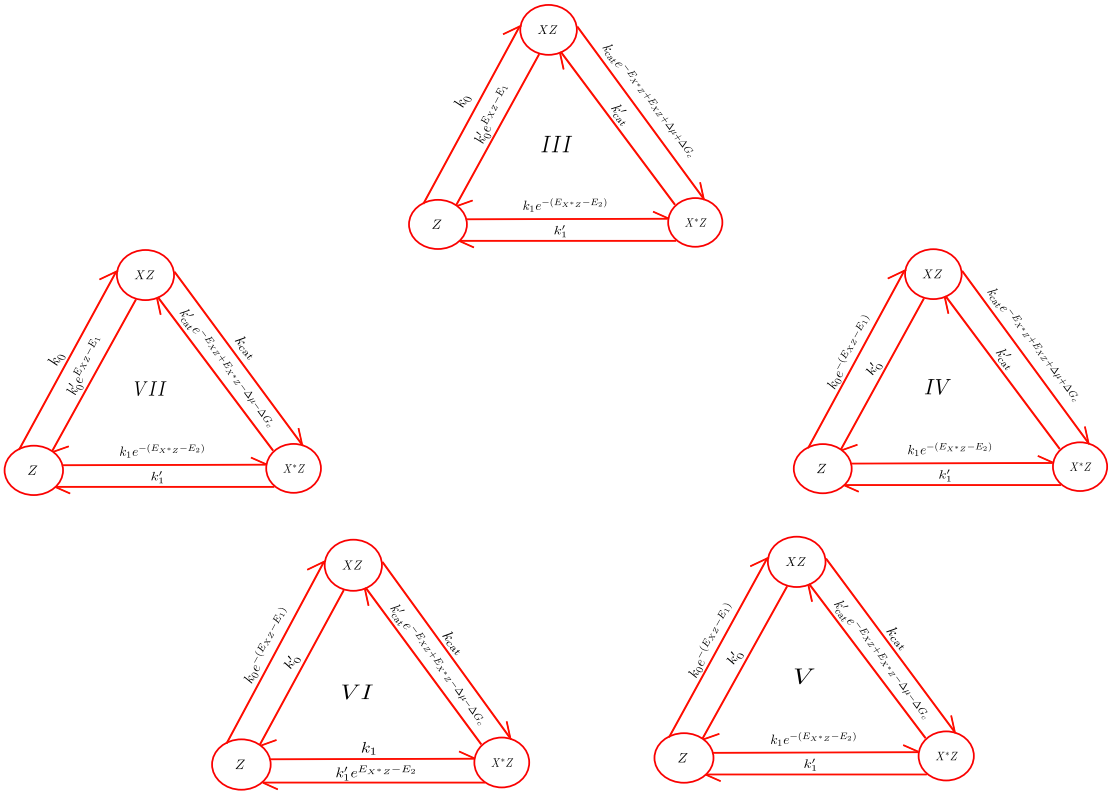


FIGURE 4.11: Continuous time markov chains corresponding to regions III, IV, V, VI, VII in Figure 4.10.a.

fluxes. Figure 4.12 illustrates this result. We can also show that $\Psi_{\text{opt}} = \Psi_0$ for optimal binding energies lying in regions IV, V, VII using the same argument we used for those lying on $E^{\text{opt}}_{X^*Z} = E^{\text{opt}}_{XZ} + \Delta\mu + \Delta G_c$. We repeat the argument here. For contradiction, assume that $\Psi_{\text{opt}} > \Psi_0$. Now increase both the binding energies E^{opt}_{XZ} and $E^{\text{opt}}_{X^*Z}$ by a sufficiently small amount ΔE such that the constraint on flux is still satisfied. As observed above, this reduces the retroactivity of the motif. Thus, we get a new tuple $(E^{\text{opt}}_{XZ} + \Delta E, E^{\text{opt}}_{X^*Z} + \Delta E)$ that has lower retroactivity; whilst still satisfying the constraint on flux. Morever, increasing both the binding energies eventually suppresses the flux, ensuring that it falls below Ψ_0. Therefore, the optimal binding energies satisfy $\Psi_{\text{opt}} = \Psi_0$.

It is interesting to observe that in regions I, II, III, VI, we can always increase one of the binding energies to go to a state with lower retroactivity and higher flux. We could not do the same in regions IV, V, VII. However, it is true that if we increase the appropriate binding energies (E_{XZ} in region V and E_{X^*Z} in region IV), then eventually, retroactivity

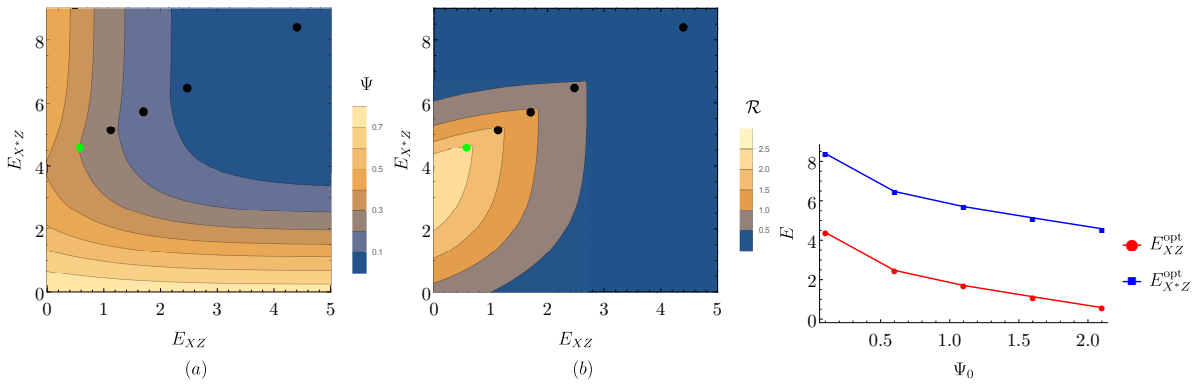


FIGURE 4.12: (a) Retroactivity contour plot as a function of binding energies. (b) Flux contour plot as a function of binding energies. The point coloured green in plots a and b lies somewhere on the intersection of regions I and II. Black dots indicate optimal binding energies outside regions I and II. (c) Optimal binding energies as a function of target flux. Parameters used for optimization: $Z_{\text{tot}} = 10, a = 1, b = 3, k_0 = 1, k_1 = 1, k_{\text{cat}} = 1, \Delta\mu = 3, \Delta G_c = 1, E_1 = 4, E_2 = 5$. The difference between the optimal binding energies $E_{X^*Z} - E_{XZ}$ is constant.

will converge on a finite value, but flux will be suppressed exponentially to 0. In contrast, if we increase both the binding energies simultaneously by the same amount, we will suppress retroactivity and, eventually the flux. Consequently, the optimal solutions tend to balance out the energies. As one varies the requirement on the target flux, one obtains a locus of optimal binding energies. In some cases, this locus intersects regions I and II. When it does, we observe that the difference between E_{X^*Z} and E_{XZ} is a constant given by $\Delta\mu + \Delta G_c$. Figure 4.12 illustrates this fact.

When the locus lies entirely outside regions I and II, we see that the difference between E_{X^*Z} and E_{XZ} might get (slightly) larger as you increase the target flux. This effect is depicted in Figure 4.13. One simple criterion to decide when the locus of optimal binding energies lies outside regions I and II would be the following: Consider all binding energies that lie on the line $E_{X^*Z} = E_{XZ} + \Delta\mu + \Delta G_c$. Amongst these binding energies, if the binding energies that give the maximum value of flux lie outside regions I and II, then the claim is that the optimal binding energies cannot lie in regions I and II. This is simply because, retroactivity decreases as we move along $E_{X^*Z} = E_{XZ} + \Delta\mu + \Delta G_c$ by increasing both binding energies. If the optimal binding energies lie within regions I and II, then the binding energies satisfying $E_{X^*Z} = E_{XZ} + \Delta\mu + \Delta G_c$ that correspond to the

maximum flux have higher flux and lower retroactivity than the optimal binding energies, a contradiction.

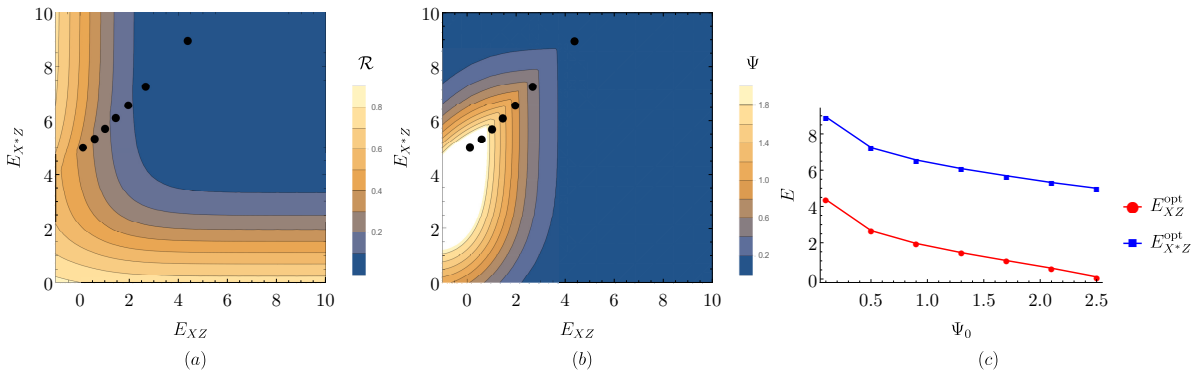


FIGURE 4.13: (a) Retroactivity contour plot as a function of binding energies. (b) Flux contour plot as a function of binding energies. (c) Optimal binding energies as a function of target flux. Black dots indicate optimal binding energies outside regions I and II. Parameters used for optimization: $Z_{\text{tot}} = 10, a = 1, b = 3, k_0 = 1, k_1 = 1, k_{\text{cat}} = 1, \Delta\mu = 3, \Delta G_c = 2, E_1 = 4, E_2 = 5$. The difference between the optimal binding energies $E_{X^*Z} - E_{XZ}$ gets larger as you increase the target flux.

We do not yet have a rigorous explanation for explaining these phenomenon. Exploring these effects is a subject of future work.

4.5 Conclusion and future work

We have analysed a model of enzyme-substrate catalysis. Our model has two intermediate states (or enzyme-substrate complexes), namely XZ and X^*Z, which decouple the catalytic reactions from the binding/unbinding ones. The notion of retroactivity corresponds to the time spent in these intermediate states. Our primary question was: how to choose binding energies (E_{XZ}, E_{X^*Z}) corresponding to these intermediate states so that it minimizes retroactivity, whilst maintaining a certain minimum flux for the reaction? We make reasonably modest assumptions about the reaction network like fixing the rate constant for the binding reactions and upper-bounding the catalytic rate constants. Given these constraints on the intrinsic dynamics, we find that there is an energy-tradeoff for designing optimal biological networks. Choosing binding energies that are too low implies that the intermediate states act as sinks, with the system spending a substantial

proportion of time in them. This increases the retroactivity of the system. Choosing very high binding energies implies that the system spends meagre amount of time in the intermediate states, thereby reducing the retroactivity of the system. However, high binding energies jeopardises the progress of the reaction, failing to meet the minimal flux requirement.

Interestingly, we find that in the system we consider, the optimal binding energies are tuned with each other. In particular, their difference $E_{X^*Z} - E_{XZ}$ is a constant and is related to the intrinsic free energy difference between the products and reactants. Further, this difference is independent of the concentrations of $[X]/[X^*]$. However, more complex models could break this tuning. It would be worth seeing how many of the results that we have for this simple model carry over to more complex networks (e.g. wherein one might have more intermediates states). We expect the idea that binding energies could be as high as possible without disrupting the flow of the reaction to still hold in these cases. In addition, if we don't allow binding reactions to be limited by diffusion, we find that most of our results carry over qualitatively with some minor differences; in particular, the difference between the optimal binding energies may not be perfectly constant.

The present work gives a macro-level insight into the analysis of enzyme-substrate catalysis. It is possible to explore a micro-level approach wherein one can design such biochemical systems using *in-vitro* DNA nanotechnology. Our work also opens up possible avenues for the analysis of push-pull motifs [95, 160, 161] since our system can be thought of as half of a push-pull motif.

This chapter flows out as a natural extension to the previous chapter. In chapter 3, one could reduce both the retroactivity and energy consumption by coupling weakly to the push-pull motif; however we could not do this arbitrarily due to cross-talk caused by leak reactions in the circuit. In this chapter, we effectively analyse how retroactivity might be reduced at fixed energy consumption and coupling strength i.e. what can be achieved at the level of single interactions, if the network-level approach fails.

One question to explore is whether an increased free energy difference between X^* and X (given by $\Delta\mu$) leads to circuits having lower retroactivity (as per the optimization Problem 4.2.1). This essentially amounts to asking whether stronger driving by ancillary fuel molecules produces better circuits. In addition, we also need to explain the behaviour of optimal binding energies when the binding rate constants are not diffusion-controlled.

Chapter 5

Conclusion

We have tried to present different facets of computation in the context of physical and biological systems. In particular, our focus has been on the notion of information processing in these systems. The idea of bits and their computations provides a link between the actual information processing tasks performed by these systems and its mathematical abstraction. We summarise the important contributions and highlight certain similarities and differences between physical and biological computations.

Our primary operations of interest are switching and erasing a bit. The former corresponds to flipping the state of the bit (i.e. taking it from 0 to 1 or from 1 to 0). The latter corresponds to setting a random bit to a fixed value (i.e. taking a bit that is equally likely to be in 0 or 1 to a state 0). Though we discuss erasure in the context of physical bits, and switching in the context of biological bits, we can effectively refer to them as "switching" without affecting the results of either systems. In the physical context, we represent a bit by a particle in a double well obeying Langevin dynamics. In the biological context, the bit is represented by a substrate, on which the enzyme acts. The notion of intermediate states involved in switching a bit in both these contexts is the coalescing feature of the thesis. There are slight differences though - our models of physical bits possess a continuum of intermediate states, while we work with discrete intermediates in

the biological context. Strictly speaking, this is an approximation – it just turns out that the biological bits we think about can be well approximated by the discrete state limit. This assumption implies that things like friction are not relevant and there is no notion of "inertia" in the biological context. The presence of these intermediate states gives rise to the notion of erasing time in the physical context and retroactivity in the biological setting.

In the case of physical bits, the notion of reliability corresponds to the amount of time spent by a particle in a single well. In the biological context, bits can switch spontaneously with some time-scale which sets the reliability time. Cross-talk, which is present in biological networks but perhaps not in physical models of bits, effectively increases the rate at which uncontrolled switching happens. In some sense, reliability time in biological systems depends on a slightly broader context than in standard computation; the same bit can be less reliable in a network with more potential cross-talking kinases. This is reflective of both the strength and the weakness of biological computation, which is that all molecules can diffuse around and interact with each other without any computational cost. The notion of erasing/switching time in a physical context corresponds to the time taken for a particle to start from a certain well and settle down in a different well. In a biological context, one can define switching time as the inverse of the rate of switching in a bit. This is (generally) not equal to the time spent in the intermediates - i.e.., it's finite even in a perfect two-state model. The common thread connecting the Langevin (physical) system and the biological bit is the interaction period (attachment of enzyme to substrate in biology and application of control in physical bits) which slows down physical computers and causes retroactivity in biological systems. Though, the exact details of implementations are not comparable, the properties of intermediate states play a vital role in both these systems. The notion of work in the the context of physical bits corresponds to the energy spent in raising the particle to a certain potential so that it is ready to settle down in a different well. Cost in the biological context corresponds to the energy required by fuel molecules to maintain the system out of equilibrium. It is also

interesting to note that in the context of physical bits, both reliability and erasing time-scales are equilibration processes; while expending work to lift the particle amounts to putting it in a non-equilibrium state with respect to the potential. On the biological side, enzymatic circuits consume fuel to maintain the system in a non-equilibrium steady-state. Our study of both physical and biological systems emphasize the physical constraints under which these operate. In biology, it is not practical to implement an arbitrarily fast catalytic reaction. This gives rise to a trade-off for choosing the appropriate binding energies to implement an optimal catalyst. In the physical setting, the constraints arise from choosing to work within a specific family of control protocols for erasing the bit. As mentioned in Chapter 2, allowing the use of arbitrary control strategies does not lead to effective trade-offs.

Yet another parallel between physical and biological computation is provided by the nature of problems involved. Some of the effects observed in Chapter 2 and 4 arise as a manifestation of an optimization problem. In the physical context, we seek to design bits that meet the required reliability and erasing time requirements with minimal work cost. For the enzyme-substrate system in Chapter 4, we wish to choose binding energies of the intermediates such that it minimizes the retroactivity of the motif, whilst proceeding at a certain minimum rate. In the physical case, the optimal bits saturate the bound on the erasing time requirement, but can exceed the bound on the reliability time requirement. For the enzyme-substrate system, the optimal binding energies saturate the bound on the flux requirement. In addition, the proofs involved in proving the notions of saturation/unsaturation are quite similar in both systems.

The idea of trade-offs is common to both physical and biological computation. Friction plays a non-trivial role in the analysis of physical bits. Both erasing and reliability time-scales are non-monotonic in the underlying friction. High reliability is favoured by very low or very high friction, whereas low erasing time is favoured by moderate friction leading to a trade-off. In fact, we find that the optimal friction is critically damped or

close to critical damping. In a similar vein, there is a trade-off to choosing the optimal binding energies for enzyme-substrate catalysis. Choosing binding energies that are too low implies that the system spends a large proportion of time in the intermediate states, which increases the retroactivity of the motif. Selecting binding energies that are too high decreases retroactivity, but also simultaneously decreases the flux through the circuit. Thus, making binding energies extremely high effectively hampers the progress of the reaction.

The overall take home message is that physical and biological computation are not that disjoint. Fundamentally, a lot of assumptions/observations apply generally to both settings; albeit being manifested in different forms. Clearly, it requires substantial energy consumption to perform operations accurately. Thinking about concrete (bio)physical settings with constrained possibilities is important in understanding engineering limitations. We only saw an energy requirement when trying to optimise reliability and erasing time requirements because we limited our control strategies. Similarly, one can only see the energy trade-off in the previous chapter if one realistically constrains certain reaction rates. High energy consumption was needed in Chapter 3 only when we considered realistic constraints on cross-talk.

It seems that there are natural biological analogues of computations performed in a physical context, as we saw in the case of switching a bit. But more complex computations in biology might not fall so nicely into in-silico paradigms. One example taking forward the ideas of computation in a biophysical framework is to build artificially engineered motifs (e.g. enzyme-substrate systems or push-pull motifs) and try to see if the behaviour of the system matches our theoretical observations.

Bibliography

[1] J. Maxwell. Letter to P.G Tait. *Life and Scientific Work of Peter Guthrie Tait, C. Knott (ed.)*, 1:213–215, 1911.

[2] H. Leff and A. Rex. *Maxwell's Demon: Entropy, Information, Computing*. Princeton University Press, 2014.

[3] K. Maruyama, F. Nori, and V. Vedral. Colloquium: The physics of Maxwell's demon and information. *Rev. Mod. Phys.*, 81:1–23, 2009.

[4] M. Smoluchowski. Experimentell nachweisbare, der üblichen thermodynamik widersprechende molekularphänomene. *Phys. Z*, 13:1069–1080, 1912.

[5] M. Smoluchowski. Vortgage uber die kinetische theorie der materie und der elektrizitat. *M. Planck et al.,(Teubner und Leipzig, Berlin 1914)*, pages 89–121, 1914.

[6] L. Szilárd. On the Decrease of Entropy in a Thermodynamic System by the Intervention of Intelligent Beings. *Z. Physik*, 53:840–856, 1929.

[7] L. Brillouin. Maxwell's Demon Cannot Operate: Information and Entropy. I. *J. Appl. Phys.*, 22(3):334–337, 1951.

[8] D. Gabor. IV light and information. In *Progress in Optics*, volume 1, pages 109–153. 1961.

[9] R. Landauer. Irreversibility and Heat Generation in the Computing Process. *IBM J. Res. Dev.*, 5(3):183–191, 1961.

[10] C. Bennett. The Thermodynamics of Computation – a Review. *Int. J. Theor. Phys.*, 21(12):905–940, 1982.

[11] Y. Lecerf. Machines de turing reversibles-reursive insolubilite en n n de l'equation u= $\theta\hat{}$ nu, ou θ est un isomorphisme de codes. *Comptes Rendus hebdomadaires des seances de l'academie des sciences*, 257:2597–2600, 1963.

[12] C. Bennett. Logical Reversibility of Computation. *IBM J. Res. Dev.*, 17(6):525–532, 1973.

[13] E. Fredkin and T. Toffoli. Conservative Logic. In *Collision-based computing*, pages 47–81. 2002.

[14] T. Toffoli. Computation and Construction Universality of Reversible Cellular Automata. *J. Comput. Syst. Sci.*, 15(2):213–231, 1977.

[15] J. Parrondo, J. Horowitz, and T. Sagawa. Thermodynamics of information. *Nat. Phys.*, 11(2):131, 2015.

[16] T. Sagawa and M. Ueda. Information Thermodynamics: Maxwell's Demon in Nonequilibrium Dynamics. *Nonequilibrium Statistical Physics of Small Systems: Fluctuation Relations and Beyond*, pages 181–211, 2013.

[17] J. von Neumann, A. Burks, et al. Theory of Self-Reproducing Automata. *IEEE Trans. Neural Netw.*, 5(1):3–14, 1966.

[18] J. Swanson. Physical versus Logical Coupling in Memory Systems. *IBM J. Res. Dev.*, 4(3):305–310, 1960.

[19] M. Ondrechen, B. Andresen, and R. Berry. Thermodynamics in finite time: Processes with temperature-dependent chemical reactions. *J. Chem. Phys.*, 73(11):5838–5843, 1980.

[20] F. Curzon and B. Ahlborn. Efficiency of a Carnot engine at maximum power output. *Am. J. Phys.*, 43(1):22–24, 1975.

[21] Y. Band, O. Kafri, and P. Salamon. Finite time thermodynamics: Optimal expansion of a heated working fluid. *J. Appl. Phys.*, 53(1):8–28, 1982.

[22] B. Andresen, P. Salamon, and R. Berry. Thermodynamics in finite time. *Phys. Today*, 37(9):62–70, 1984.

[23] B. Andresen, P. Salamon, and S. Berry. Thermodynamics in finite time: extremals for imperfect heat engines. *J. Chem. Phys.*, 66(4):1571–1577, 1977.

[24] B. Andresen, S. Berry, A. Nitzan, and P. Salamon. Thermodynamics in finite time. *i*. the step-*c*arnot cycle. *Phys. Rev. A*, 15(5):2086, 1977.

[25] K. Sekimoto. Kinetic Characterization of Heat Bath and the Energetics of Thermal Ratchet Models. *J. Phys. Soc. Jpn.*, 66(5):1234–1237, 1997.

[26] C. Jarzynski. Nonequilibrium Equality for Free Energy Differences. *Phys. Rev. Lett.*, 78(14):2690, 1997.

[27] G. Crooks. Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences. *Phys. Rev. E*, 60(3):2721, 1999.

[28] U. Seifert. Stochastic thermodynamics, fluctuation theorems and molecular machines. *Rep. Prog. Phys.*, 75(12):126001, 2012.

[29] K. Sekimoto and S. Sasa. Complementarity relation for irreversible process derived from stochastic energetics. *J. Phys. Soc. Jpn.*, 66(11):3326–3328, 1997.

[30] T. Schmiedl and U. Seifert. Optimal Finite-Time Processes In Stochastic Thermodynamics. *Phys. Rev. Lett.*, 98:108301, 2007.

[31] D. Sivak and G. Crooks. Thermodynamic metrics and optimal paths. *Phys. Rev. Lett.*, 108(19):190602, 2012.

[32] P. Zulkowski, D. Sivak, and M. DeWeese. Optimal control of transitions between nonequilibrium steady states. *PloS one*, 8(12):e82754, 2013.

[33] E. Aurell, K. Gawdzki, C. Mejía-Monasterio, R. Mohayaee, and P. Muratore-Ginanneschi. Refined Second Law of Thermodynamics for Fast Random Processes. *J. Stat. Phys.*, 147(3):487–505, 2012.

[34] P. Zulkowski and M. DeWeese. Optimal finite-time erasure of a classical bit. *Phys. Rev. E*, 89:052140, 2014.

[35] G. Diana, G. Bagci, and M. Esposito. Finite-time erasing of information stored in fermionic bits. *Phys. Rev. E*, 87:012111, 2013.

[36] I. Lehman, M. Bessman, E. Simms, and A. Kornberg. Enzymatic synthesis of deoxyribonucleic acid I. preparation of substrates and partial purification of an enzyme from Escherichia coli. *J. Biol. Chem.*, 233(1):163–170, 1958.

[37] R. Kornberg. Chromatin structure: a repeating unit of histones and DNA. *Science*, 184(4139):868–871, 1974.

[38] J. Matthaei and M. Nirenberg. Characteristics and stabilization of DNAase-sensitive protein synthesis in e. coli extracts. *Proc. Natl. Acad. Sci. U.S.A.*, 47(10):1580–1588, 1961.

[39] S. Levine. Enzyme Amplifier Kinetics. *Science*, 152(3722):651–653, 1966.

[40] R. Macfarlane. An Enzyme Cascade in the Blood Clotting Mechanism, and its Function as a Biochemical Amplifier. *Nature*, 202(4931):498, 1964.

[41] G. Wald. Visual Excitation and Blood Clotting. *Science*, 150(3699):1028–1030, 1965.

[42] J. Ferrell. Tripping the switch fantastic: how a protein kinase cascade can convert graded inputs into switch-like outputs. *Trends Biochem. Sci.*, 21(12):460 – 466, 1996.

[43] G. Brown, J. Hoek, and B. Kholodenko. Why do protein kinase cascades have more than one level? *Trends Biochem. Sci.*, 22(8):288, 1997.

[44] Yuhai T. Quantitative Modeling of Bacterial Chemotaxis: Signal Amplification and Accurate Adaptation. *Annu. Rev. Biophys*, 42(1):337–359, 2013.

[45] A. Bren and M. Eisenbach. How Signals Are Heard during Bacterial Chemotaxis: Protein-Protein Interactions in Sensory Signal Propagation. *J. Bacteriol.*, 182(24):6865–6873, 2000.

[46] M. Eisenbach. *Bacterial Chemotaxis*. John Wiley and Sons, Ltd, 2001.

[47] D. Bray, M. Levin, and C. Morton-Firth. Receptor clustering as a cellular mechanism to control sensitivity. *Nature*, 393(6680):85, 1998.

[48] J E Segall, S M Block, and H C Berg. Temporal comparisons in bacterial chemotaxis. *Proc. Natl. Acad. Sci. U.S.A.*, 83(23):8987–8991, 1986.

[49] V. Sourjik and H. Berg. Functional interactions between receptors in bacterial chemotaxis. *Nature*, 428(6981):437, 2004.

[50] U. Alon. *An Introduction to Systems Biology: Design Principles of Biological Circuits*. Chapman and Hall/CRC, 2006.

[51] L. Adleman. Molecular Computation of Solutions to Combinatorial Problems. *Science*, 266(11):1021–1024, 1994.

[52] N. Kallenbach, R. Ma, and N. Seeman. An immobile nucleic acid junction constructed from oligonucleotides. *Nature*, 305(5937):829, 1983.

[53] P. Rothemund. Folding DNA to create nanoscale shapes and patterns. *Nature*, 440(7082):297, 2006.

[54] G. Seelig, D. Soloveichik, D. Zhang, and E. Winfree. Enzyme-Free Nucleic Acid Logic Circuits. *Science*, 314(5805):1585–1588, 2006.

[55] D. Zhang and E. Winfree. Control of DNA Strand Displacement Kinetics Using Toehold Exchange. *J. Am. Chem. Soc.*, 131(47):17303–17314, 2009.

[56] L. Qian and E. Winfree. Scaling Up Digital Circuit Computation with DNA Strand Displacement Cascades. *Science*, 332(6034):1196–1201, 2011.

[57] F. Horn and R. Jackson. General Mass Action Kinetics. *Arch. Ration. Mech. Anal.*, 47(2):81–116, 1972.

[58] F. Horn. The dynamics of open reaction systems, Mathematical aspects of chemical and biochemical problems and quantum chemistry (new york). *Proc. SIAM-AMS Sympos. Appl. Math*, 8, 1974.

[59] M. Feinberg and F. Horn. Dynamics of open chemical systems and the algebraic structure of the underlying reaction network. *Chem. Eng. Sci.*, 29(3):775–787, 1974.

[60] M. Feinberg. Chemical reaction network structure and the stability of complex isothermal reactors I. The deficiency zero and deficiency one theorems. *Chem. Eng. Sci.*, 42(10):2229–2268, 1987.

[61] M. Feinberg. Chemical reaction network structure and the stability of complex isothermal reactors II. Multiple steady states for networks of deficiency one. *Chem. Eng. Sci.*, 43(1):1–25, 1988.

[62] M. Feinberg. Mathematical aspects of mass action kinetics. *Chemical Reactor Theory: A Review*, pages 1–78, 1977.

[63] M. Feinberg. Chemical Oscillations, Multiple Equilibria, and Reaction Network Structure. In *Dynamics and modelling of reactive systems*, pages 59–130. 1980.

[64] E. Sontag. Structure and stability of certain chemical networks and applications to the kinetic proofreading model of T-cell receptor signal transduction. *IEEE transactions on automatic control*, 46(7):1028–1047, 2001.

[65] D. Angeli and E. Sontag. Translation-invariant monotone systems, and a global convergence result for enzymatic futile cycles. *Nonlinear Anal. Real World Appl.*, 9(1):128–140, 2008.

[66] M. Banaji and G. Craciun. Graph-theoretic criteria for injectivity and unique equilibria in general chemical reaction systems. *Adv. Appl. Math.*, 44(2):168–184, 2010.

[67] B. Kholodenko. Negative feedback and ultrasensitivity can bring about oscillations in the mitogen-activated protein kinase cascades. *The FEBS Journal*, 267(6):1583–1588, 2000.

[68] M. Banaji. Inheritance of oscillation in chemical reaction networks. *Appl. Math. Comput.*, 325:191–209, 2018.

[69] M. Banaji. Cycle structure in SR and DSR graphs: implications for multiple equilibria and stable oscillation in chemical reaction networks. In *Transactions on Petri nets and other models of concurrency V*, pages 1–21. 2012.

[70] D. Soloveichik, G. Seelig, and E. Winfree. DNA as a universal substrate for chemical kinetics. *Proc. Natl. Acad. Sci. U.S.A*, 107(12):5393–5398, 2010.

[71] H. Chen, D. Doty, and D. Soloveichik. Deterministic function computation with chemical reaction networks. *Natural computing*, 13(4):517–534, 2014.

[72] D. Soloveichik, M. Cook, E. Winfree, and J. Bruck. Computation with finite stochastic chemical reaction networks. *Nat. Comput.*, 7(4):615–633, 2008.

[73] L. Cardelli, M. Kwiatkowska, and L. Laurenti. Programming discrete distributions with chemical reaction networks. *Nat. Comput.*, 2017.

[74] L. Cardelli and A. Csikász-Nagy. The Cell Cycle Switch Computes Approximate Majority. *Sci. Rep.*, 2:656, 2012.

[75] M. Gopalkrishnan. A Scheme for Molecular Computation of Maximum Likelihood Estimators for Log-Linear Models. In Yannick Rondelez and Damien Woods, editors, *DNA Computing and Molecular Programming*. Springer, 2016.

[76] Muppirala V. Virinchi, A. Behera, and M. Gopalkrishnan. A Stochastic Molecular Scheme for an Artificial Cell to Infer Its Environment from Partial Observations. In Robert Brijder and Lulu Qian, editors, *DNA Computing and Molecular Programming*. Springer, 2017.

[77] T. E. Ouldridge. The importance of thermodynamics for molecular systems, and the importance of molecular systems for thermodynamics. *Nat. Comput.*, pages 1–27, 2017.

[78] R. Cummings, D. Doty, and D. Soloveichik. Probability 1 computation with chemical reaction networks. *Nat. Comput.*, 15(2):245–261, 2016.

[79] D. Chu. Performance limits and trade-offs in entropy-driven biochemical computers. *J. Theor. Biol.*, 443:1–9, 2018.

[80] T. E. Ouldridge, C. Govern, and P. ten Wolde. Thermodynamics of Computational Copying in Biochemical Systems. *Phys. Rev. X*, 7:021004, 2017.

[81] D. Oyarzún and G. Stan. Synthetic gene circuits for metabolic control: design trade-offs and constraints. *J. Royal Soc. Interface*, 10(78), 2013.

[82] J. Lugagne, D. Oyarzún, and G. Stan. Stochastic simulation of enzymatic reactions under transcriptional feedback regulation. In *ECC*, pages 3646–3651, 2013.

[83] D. Oyarzun, J. Lugagne, and G. Stan. Noise Propagation in Synthetic Gene Circuits for Metabolic Control. *ACS synthetic biology*, 4(2):116–125, 2014.

[84] N. Kashtan and U. Alon. Spontaneous evolution of modularity and network motifs. *Proc. Natl. Acad. Sci. U.S.A.*, 102(39):13773–13778, 2005.

[85] E. Raff and R. Raff. Dissociability, modularity, evolvability. *Evol. Dev*, 2(5), 2000.

[86] J. Clune, J. Mouret, and H. Lipson. The evolutionary origins of modularity. *Proc. R. Soc. Lond. B Biol. Sci.*, 280(1755), 2013.

[87] G. Wagner, M. Pavlicev, and J. Cheverud. The road to modularity. *Nat. Rev. Genet.*, 8(12):921, 2007.

[88] T. Tran and Y. Kwon. The relationship between modularity and robustness in signalling networks. *J. Royal Soc. Interface*, 10(88), 2013.

[89] M. Rorick and G. Wagner. Protein Structural Modularity and Robustness are Associated with Evolvability. *Genome Biol. Evol.*, 3:456–475, 2011.

[90] L. Pantoja-Hernández and J. Martínez-García. Retroactivity in the Context of Modularly Structured Biomolecular Systems. *Front. Bioeng. Biotechnol.*, 3:85, 2015.

[91] J. Saez-Rodriguez, A. Kremling, and E. Gilles. Dissecting the puzzle of life: modularization of signal transduction networks. *Comput. Chem. Eng.*, 29(3):619–629, 2005.

[92] D. Del Vecchio, A. Ninfa, and E. Sontag. Modular cell biology: retroactivity and insulation. *Mol. Syst. Biol.*, 4(1):161, 2008.

[93] J. Barton and E. Sontag. The Energy Costs of Insulators in Biochemical Networks. *Biophys. J.*, 104(6):1380–1390, 2013.

[94] A. Deshpande, M. Gopalkrishnan, T. Ouldridge, and N. Jones. Designing the optimal bit: balancing energetic cost, speed and reliability. *Proc. R. Soc. A*, 473(2204):20170117, 2017.

[95] A. Deshpande and T.E Ouldridge. High rates of fuel consumption are not required by insulating motifs to suppress retroactivity in biochemical circuits. *Eng. Biol.*, 1(2), 2017.

[96] C. Bennett. Notes on the history of reversible computation. *IBM J. Res. Dev.*, 32(1):16–23, 1988.

[97] C. Bennett. Notes on Landauer's principle, Reversible Computation, and Maxwell's Demon. *Stud. Hist. Philos. M P*, 34:501–510, 2003.

[98] M. Frank. The physical limits of computing. *Computing in Science and Engineering*, 4(3):16–26, 2002.

[99] E. Pop. Energy dissipation and transport in nanoscale devices. *Nano. Res.*, 3(3):147–169, 2010.

[100] T. Ouldridge, C. Govern, and P. Wolde. The thermodynamics of Computational Copying in Biochemical Systems. *Phys. Rev. X*, 7:021004, 2017.

[101] R. Sarpeshkar. *Ultra Low Power Bioelectronics: Fundamentals, Biomedical Applications, and Bio-Inspired Systems*. 2010.

[102] R. Sarpeshkar. Analog versus Digital: Extrapolating from Electronics to Neurobiology. *Neural Comput.*, 10(7):1601–1638, 1998.

[103] B. Rapoport, J. Kedzierski, and R. Sarpeshkar. A Glucose Fuel Cell for Implantable BrainMachine Interfaces. *PloS one*, 7(6):e38436, 2012.

[104] Y. Tu. The nonequilibrium mechanism for ultrasensitivity in a biological switch: Sensing by Maxwell's demons. *Proc. Natl. Acad. Sci. U.S.A.*, 105(33):11737–11741, 2008.

[105] G. Lan, P. Sartori, S. Neumann, V. Sourjik, and Y. Tu. The energy-speed-accuracy trade-off in sensory adaptation. *Nat. Phys.*, 8(5):422–428, 2012.

[106] C. Govern and P. ten Wolde. Optimal resource allocation in cellular sensing systems. *Proc. Natl. Acad. Sci. U.S.A.*, 111(49):17486–17491, 2014.

[107] S. Bo, M. Giudice, and A. Celani. Thermodynamic limits to information harvesting by sensory systems. *J. Stat. Mech. Theory Exp.*, 2015(1):P01014, 2015.

[108] A. Barato, D. Hartich, and U. Seifert. Efficiency of cellular information processing. *New J. Phys.*, 16(10):103024, 2014.

[109] W. Anacker. Josephson Computer Technology: An IBM Research Project. *IBM J. Res. Dev.*, 24(2):107–112, 1980.

[110] M. Büttiker, E. Harris, and R. Landauer. Thermal activation in extremely underdamped Josephson-junction circuits. *Phys. Rev. B*, 28:1268–1275, 1983.

[111] M. Klein and A. Mukherjee. Thermal noise induced switching of Josephson logic devices. *Appl. Phys. Lett.*, 40(8):744–747, 1982.

[112] K. Likharev. Classical and quantum limitations on energy consumption in computation. *Int. J. Theor. Phys*, 21(3):311–326, 1982.

[113] P. Kloeden and E. Platen. Higher-order implicit strong numerical schemes for stochastic differential equations. *J. Stat. Phys.*, 66(1):283–314, 1992.

[114] G. Pavliotis. *Stochastic processes and applications*, volume 60. Springer, 2014.

[115] J. Mattingly and A. Stuart. Geometric Ergodicity of Some Hypo-Elliptic Diffusions for Particle Motions. *Markov Process relat.*, 8(2):199–214, 2002.

[116] M. Gopalkrishnan. The Hot Bit I: The Szilard-Landauer Correspondence. *arXiv:1311.3533*, [cs.IT], 2013.

[117] A. Bérut, A. Arakelyan, A. Petrosyan, S. Ciliberto, R. Dillenschneider, and E. Lutz. Experimental verification of Landauer's principle linking information and thermodynamics. *Nature*, 483(7388):187–189, 2012.

[118] Y. Jun, M. Gavrilov, and J. Bechhoefer. High-precision test of landauer's principle in a feedback trap. *Phys. Rev. Lett.*, 113:190601, 2014.

[119] T. Sagawa and M. Ueda. Nonequilibrium thermodynamics of feedback control. *Phys. Rev. E*, 85:021104, 2012.

[120] T. Sagawa and M. Ueda. Generalized Jarzynski Equality under Nonequilibrium Feedback Control. *Phys. Rev. Lett.*, 104:090602, 2010.

[121] M. Ponmurugan. Generalized detailed fluctuation theorem under nonequilibrium feedback control. *Phys. Rev. E*, 82:031129, 2010.

[122] D. Abreu and U. Seifert. Thermodynamics of Genuine Nonequilibrium States under Feedback Control. *Phys. Rev. Lett.*, 108:030601, 2012.

[123] J. Horowitz and S. Vaikuntanathan. Nonequilibrium detailed fluctuation theorem for repeated discrete feedback. *Phys. Rev. E.*, 82:061120, 2010.

[124] L. Sourabh, R. Shubhashis, and A.M. Jayannavar. Fluctuation theorems in the presence of information gain and feedback. *J. Phys. A*, 45(6):065002, 2012.

[125] J. Vega, R. Guantes, and S. Miret-Artes. Mean first passage time and the Kramers turnover theory in activated atom-surface diffusion. *Phys. Chem. Chem. Phys.*, 4:4985–4991, 2002.

[126] L. Ruslan, H. Richard, and M. Tretyakov. Langevin thermostat for rigid body dynamics. *J. Chem. Phys.*, 130(23):234101, 2009.

[127] K. Sekimoto. Langevin Equation and Thermodynamics. *Prog. Theor. Phys. Supp.*, 130:17–27, 1998.

[128] H. Kramers. Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica*, 7(4):284–304, 1940.

[129] V. Melnikov and S. Meshkov. Theory of activated rate processes: Exact solution of the Kramers problem. *J. Chem. Phys.*, 85(2):1018–1027, 1986.

[130] E. Pollak, H. Grabert, and P. Hänggi. Theory of activated rate processes for arbitrary frequency dependent friction: Solution of the turnover problem. *J. Chem. Phys.*, 91(7):4073–4087, 1989.

[131] P. Hänggi, P. Talkner, and M. Borkovec. Reaction-rate theory: fifty years after Kramers. *Rev. Mod. Phys.*, 62:251–341, Apr 1990.

[132] *Über die Dissociationswärme und den Einfluss der Temperatur auf den Dissociationsgrad der Elektrolyte.*

[133] S. Wright and J. Nocedal. Numerical optimization. *Springer Science*, 35:67–68, 1999.

[134] T. Rapcsák. Smooth nonlinear optimization in R^n. *Springer*, 19:376, 1997.

[135] T. Schmiedl and U. Seifert. Optimal Finite-Time Processes In Stochastic Thermodynamics. *Phys. Rev. Lett.*, 98:108301, 2007.

[136] H. Then and A. Engel. Computing the optimal protocol for finite-time processes in stochastic thermodynamics. *Phys. Rev. E*, 77:041105, 2008.

[137] E. Aurell, C. Mejía-Monasterio, and P. Muratore-Ginanneschi. Optimal Protocols and Optimal Transport in Stochastic Thermodynamics. *Phys. Rev. Lett.*, 106:250601, 2011.

[138] T. Gingrich, G. Rotskoff, G. Crooks, and P. Geissler. Near-optimal protocols in complex nonequilibrium transformations. *Proc. Natl. Acad. Sci. U.S.A.*, 113(37):10263–10268, 2016.

[139] M. Gopalkrishnan. A Cost/Speed/Reliability Tradeoff to Erasing. *Entropy*, 18(5), 2016.

[140] W. Gallagher, E. Harris, and M. Ketchen. Superconductivity at *ibma* centennial review: *p*art *is*uperconducting computer and device applications. In *Proc. IEEE/CSC ESAS Eur. Supercond. News Forum*, number 21, pages 1–34, 2012.

[141] A. Robinson. IBM drops superconducting computer project. *Science*, 222(4623):492–494, 1983.

[142] U. Alon. Network motifs: theory and experimental approaches. *Nat. Rev. Genet.*, 8(6):450–461, 2007.

[143] A. Hintze and C. Adami. Evolution of Complex Modular Biological Networks. *PLoS. Comput. Biol.*, 4(2):e23, 2008.

[144] J. Kim and K. Cho. Robustness Analysis of Network Modularity. *IEEE Trans. Control Netw. Syst*, 3(4):348–357, 2016.

[145] L. Hartwell, J. Hopfield, S. Leibler, and A. Murray. From molecular to modular cell biology. *Nature*, 402:C47–C52, 1999.

[146] D. Lauffenburger. Cell signaling pathways as control modules: Complexity for simplicity? *Proc. Natl. Acad. Sci. U.S.A.*, 97(10):5031–5033, 2000.

[147] P Purnick and R Weiss. The second wave of synthetic biology: from modules to systems. *Nat. Rev. Mol. Cell. Biol.*, 10(6):410–422, 2009.

[148] E. Andrianantoandro, S. Basu, D. Karig, and R. Weiss. Synthetic biology: new engineering rules for an emerging discipline. *Mol. Syst. Biol.*, 2(1), 2006.

[149] J. Rodriguez, A. Kremling, and E. Gilles. Dissecting the puzzle of life: modularization of signal transduction networks. *Comput. Chem. Eng.*, 29(3):619 – 629, 2005.

[150] A. Ventura, J. Sepulchre, and S. Merajver. A Hidden Feedback in Signaling Cascades Is Revealed. *PLoS. Comput. Biol.*, 4(3):e1000041, 2008.

[151] D. Del Vecchio, A. Ninfa, and E. Sontag. Modular cell biology: retroactivity and insulation. *Mol. Syst. Biol.*, 4(1):161, 2008.

[152] D. Del Vecchio and R. Murray. *Biomolecular Feedback Systems*. Princeton University Press, 2015.

[153] H. Kim and H. Sauro. Fan-out in gene regulatory networks. *J. Biol. Eng.*, 4(1):16, 2010.

[154] R. Shah and D. Del Vecchio. Signaling Architectures that Transmit Unidirectional Information Despite Retroactivity. *bioRxiv*, page 111971, 2017.

[155] J. Barton and E. Sontag. Remarks on the energy costs of insulators in enzymatic cascades. *arXiv:1412.8065*.

[156] R. Ghaemi and D. Del Vecchio. Stochastic analysis of retroactivity in transcriptional networks through singular perturbation. In *2012 American Control Conference (ACC)*, pages 2731–2736, 2012.

[157] M. Lemmon, K. Ferguson, and J. Schlessinger. PH domains: Diverse sequences with a Common Fold Recruit Signaling Molecules to the Cell Surface. *Cell*, 85:621–624, 1996.

[158] A. Green, P. Silver, J. Collins, and P. Yin. Toehold Switches: De-Novo-Designed Regulators of Gene Expression. *Cell*, 159:925–939, 2014.

[159] P. Mehta, A. Lang, and D. Schwab. Landauer in the Age of Synthetic Biology: Energy Consumption and Information Processing in Biochemical Networks. *J. Stat. Phys.*, 162:1153–1166, 2016.

[160] M. Robinson and M. Cobb. Mitogen-activated protein kinase pathways. *Curr. Opin. Cell. Biol.*, 9(2):180–186, 1997.

[161] M. Samoilov, S. Plyasunov, and A. Arkin. Stochastic amplification and signaling in enzymatic futile cycles through noise-induced bistability with oscillations. *Proc. Natl. Acad. Sci. U.S.A.*, 102(7):2310–2315, 2005.

[162] C. Govern and P. ten Wolde. Optimal resource allocation in cellular sensing systems. *Proc. Natl. Acad. Sci. U.S.A.*, 111(49):17486–17491, 2014.

[163] C. Huang and J. Ferrell. Ultrasensitivity in the mitogen-activated protein kinase cascade. *Proc. Natl. Acad. Sci. U.S.A.*, 93(19):10078–10083, 1996.

[164] E. Sontag. *Modularity, Retroactivity, and Structural Identification*, pages 183–200. Springer, 2011.

[165] S. Jayanthi, K. Nilgiriwala, and D. Del Vecchio. Retroactivity Controls the Temporal Dynamics of Gene Transcription. *ACS Synth. Biol.*, 2(8):431–441, 2013.

[166] H. Ossareh, A. Ventura, S. Merajver, and D. Del Vecchio. Long Signaling Cascades Tend to Attenuate Retroactivity. *Biophys. J.*, 100(7):1617–1626, 2011.

[167] A. Stock, V. Robinson, and P. Goudreau. Two-component signal transduction. *Annu. Rev. Biochem.*, 69(1):183–215, 2000.

[168] A. Novick and L. Szilard. Description of the Chemostat. *Science*, 112(2920):715–716, 1950.

[169] R. Aris. Prolegomena to the rational analysis of systems of chemical reactions. *Arch. Rational Mech. Anal.*, 19(2):81–99, 1965.

[170] M. Feinberg. Necessary and sufficient conditions for detailed balancing in mass action systems of arbitrary complexity. *Chem. Eng. Sci.*, 44(9):1819–1827, 1989.

[171] J. Gunawardena. Chemical reaction network theory for in-silico biologists. *Notes available for download at http://vcp. med. harvard. edu/papers/crnt. pdf*, 2003.

[172] D. Bender. *Introduction to Nutrition and Metabolism*. CRC Press, 2014.

[173] E. Stadtman and P. Chock. Superiority of interconvertible enzyme cascades in metabolic regulation: analysis of monocyclic systems. *Proc. Natl. Acad. Sci. U.S.A.*, 74(7):2761–2765, 1977.

[174] A. Goldbeter and D. Koshland. An amplified sensitivity arising from covalent modification in biological systems. *Proc. Natl. Acad. Sci. U.S.A.*, 78(11):6840–6844, 1981.

[175] S. Legewie, H. Herzel, H. Westerhoff, and N. Blüthgen. Recurrent design patterns in the feedback regulation of the mammalian signalling network. *Mol. Syst. Biol.*, 4(1):190, 2008.

[176] E. Dekel and U. Alon. Optimality and evolutionary tuning of the expression level of a protein. *Nature*, 436(7050):588, 2005.

[177] P. ten Wolde, N. Becker, T. Ouldridge, and A. Mugler. Fundamental Limits to Cellular Sensing. *J. Stat. Phys.*, 162:1395–1424, 2016.

[178] G. Lewis. A New Principle of Equilibrium. *Proc. Natl. Acad. Sci. U.S.A.*, 11(3):179–183, 1925.

[179] G. Zubay. *Biochemistry*. Brown Publishers, US, 1998.

[180] T. Friedlander, R. Prizak, C. Guet, N. Barton, and G. Tkačik. Intrinsic limits to gene regulation by global crosstalk. *Nat. Commun.*, 7, 2016.

[181] P. Swain and E. Siggia. The Role of Proofreading in Signal Transduction Specificity. *Biophys. J.*, 82:2928–2933, 2002.

[182] T. Ouldridge and P. ten Wolde. The Robustness of Proofreading to Crowding-Induced Pseudo-Processivity in the MAPK pathway. *Biophys. J.*, 107:2425 – 2435, 2014.

[183] O. Berg, J. Paulsson, and M. Ehrenberg. Fluctuations and Quality of Control in Biological Cells: Zero-Order Ultrasensitivity Reinvestigated. *Biophys. J.*, 79(3):1228–1236, 2000.

[184] J. Ferrell and S. Ha. Ultrasensitivity part I: Michaelian responses and zero-order ultrasensitivity. *Trends Biochem. Sci.*, 39(10):496–503, 2014.

[185] P. Jithinraj, U. Roy, and M. Gopalakrishnan. Zero-order ultrasensitivity: A study of criticality and fluctuations under the total quasi-steady state approximation in the linear noise regime. *J. Theor. Biol.*, 344:1–11, 2014.

[186] M. Baiesi and C. Maes. Life efficiency does not always increase with the dissipation rate. *arXiv:1707.09614*, 2017.

[187] S. Tănase-Nicola, P. Warren, and P. Ten Wolde. Signal detection, modularity, and the correlation between extrinsic and intrinsic noise in biochemical networks. *Phys. Rev. Lett.*, 97(6):068102, 2006.

[188] S. Jayanthi and D. Del Vecchio. Retroactivity Attenuation in Bio-molecular Systems based on Timescale Separation. *IEEE Trans. Autom. Control*, 56(4):748–761, 2011.

[189] L. Meijer, A. Joesaar, E. Steur, W. Engelen, R. Santen, M. Merkx, and T. Greef. Hierarchical control of enzymatic actuators using DNA-based switchable memories. *Nat. Commun.*, 8(1):1117, 2017.

[190] L. Menten and M. Michaelis. Die kinetik der invertinwirkung. *Biochem Z*, 49:333–369, 1913.

[191] J. Keener and J. Sneyd. *Mathematical Physiology I: Cellular Physiology*. Springer, 2010.

[192] G. Briggs and J. Haldane. A Note on the Kinetics of Enzyme Action. *Biochem. J.*, 19(2):338, 1925.

[193] C. Frenzen and P.K. Maini. Enzyme kinetics for a two-step enzymic reaction with comparable initial enzyme-substrate ratios. *J. Math. Biol.*, 26(6):689–703, 1988.

[194] M. Fisher and A. Kolomeisky. The force exerted by a molecular motor. *Proc. Natl. Acad. Sci. U.S.A.*, 96(12):6597–6602, 1999.

[195] R Astumian. Irrelevance of the Power Stroke for the Directionality, Stopping Force, and Optimal Efficiency of Chemically Driven Molecular Machines. *Biophys. J.*, 108(2):291–303, 2015.

[196] H. Ge and H. Qian. Dissipation, generalized free energy, and a self-consistent nonequilibrium thermodynamics of chemically driven open subsystems. *Phys. Rev. E*, 87(6):062125, 2013.

[197] A. Wachtel, R. Rao, and M. Esposito. Thermodynamically consistent coarse graining of biocatalysts beyond michaelismenten. *New J. Phys.*, 20(4):042002, 2018.

[198] P. Nelson and S. Doniach. Biological Physics: Energy, Information Life. *Phys. Today*, 57(11):63–64, 2004.

[199] H. Qian. Phosphorylation energy hypothesis: open chemical systems and their biological functions. *Annu. Rev. Phys. Chem.*, 58:113–142, 2007.

[200] D. Beard and H. Qian. *Chemical Biophysics: Quantitative Analysis of Cellular Systems*. Cambridge University Press, 2008.

[201] J. Hopfield. Kinetic Proofreading: A New Mechanism for Reducing Errors in Biosynthetic Processes Requiring High Specificity. *Proc. Natl. Acad. Sci. U.S.A.*, 71(10):4135–4139, 1974.

[202] J. Ninio. Kinetic amplification of enzyme discrimination. *Biochimie*, 57(5):587–595, 1975.

[203] J. Estrada, F. Wong, A. DePace, and J. Gunawardena. Information Integration and Energy Expenditure in Gene Regulation. *Cell*, 166(1):234–244, 2016.

[204] T Gregor, D. Tank, E. Wieschaus, and W. Bialek. Probing the Limits to Positional Information. *Cell*, 130(1):153–164, 2007.

[205] P. Atkins and J. De Paula. *Physical Chemistry for the Life Sciences*. Oxford University Press, USA, 2011.

[206] R. Alberty and G. Hammes. Application of the Theory of Diffusion-controlled Reactions to Enzyme Kinetics. *J. Phys. Chem. A*, 62(2):154–159, 1958.

[207] A. Brown and D. Sivak. Allocating dissipation across a molecular machine cycle to maximize flux. *Proc. Natl. Acad. Sci. U.S.A.*, 114(42):11057–11062, 2017.

[208] J. Norris. *Markov Chains*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 1997.

Chapter 6

Appendix A

6.1 Validating the timestep of the integrator

We validate the accuracy of our Langevin integrator by considering the dependence of thermodynamic expectations on the time step. We calculate the average potential and kinetic energies for a particle in a quadratic potential $W_{A,B} = A\left(\frac{x}{B} - 1\right)^2$, a quadratic proxy for a single well of the quartic resting-state potential. We plot the results in Figure 6.1 for a few representative values of the friction coefficient $\gamma = [0.1, 1.10, 100]$ and $A = 10$. Each result is based on an average from 10 simulations each of 5×10^8 time steps. As is evident from the figure, a time step of 0.001 gives good convergence to the equipartition limit of $k_B T/2$.

However, it is not sufficient to just compare the average kinetic and potential energies to the equipartition limit. We need to ensure that the observed kinetics are robust to our choice of time step. In particular, we need to test that a time step of 0.001 is sufficient for the highest values of our control parameter F, which presents the most severe challenge to integrating our Langevin equation (due to the behaviour near $x = 0$). Figure 6.2 confirms that a time step of 0.001 is appropriate for $F = 100$ and the full range of γ

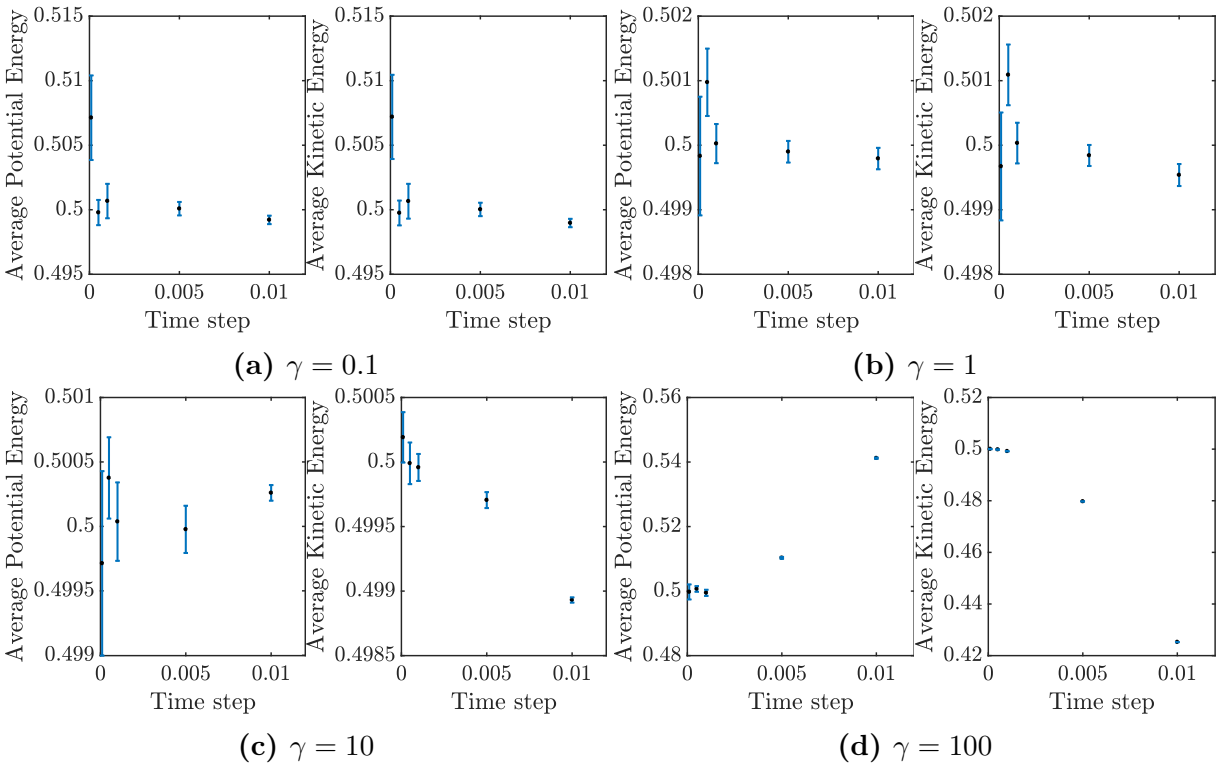


FIGURE 6.1: A time step of 0.001 is good enough to ensure that the average potential and kinetic energies approaches $\frac{k_B T}{2} = 0.5$ for a wide range of friction values.

tested. Each value in the figure is an average over 1000 initial conditions.

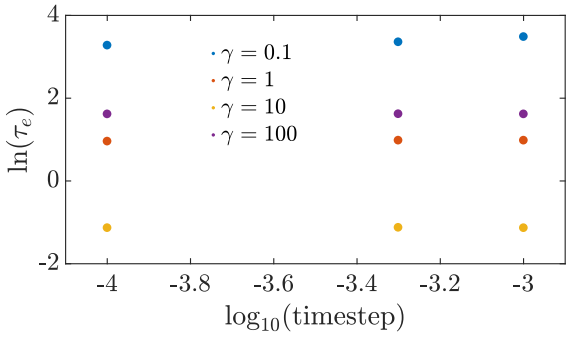


FIGURE 6.2: A time step of 0.001 gives reasonable values of erasing time for the largest value of the control parameter ($F = 100$) that we use in the simulations.

6.2 Erasure region

In this section, we demonstrate that our results are not limited to our specific definition of the erasure region by considering two alternative criteria for erasing and confirming that our earlier conclusions are supported.

6.2.1 Accuracy of erasure: Convergence of probability distribution

Recall that the erasing time is a sum of transport time (τ_t) and the mixing time (τ_m). Since the transport time is independent of the metric used to measure the mixing time, we analyse a proxy for the mixing process in isolation. Specifically, we consider the relaxation of a particle in a harmonic well, initially prepared in an arbitrary non-equilibrium distribution. As an alternative definition of mixing, we consider τ_m^ϵ as the first time when the probability distribution of the particle comes within a certain distance (in the appropriate norm, and relative to its initial distribution) of the Gibbs distribution corresponding to the well. More specifically,

$$\tau_m^\epsilon = \inf_{t\geq 0}\{||\text{law}((x(t),p(t)) - \pi_0(x,p)||_{L^2(\pi_0(x,p))} \leq \epsilon||\text{law}((x(0),p(0)) - \pi_0(x,p)||_{L^2(\pi_0(x,p))}\} \quad (6.1)$$

where $(x(t),p(t))$ is the solution to Equation 2.7 given appropriate initial conditions. Here, $\pi_0(x,p)$ is the stationary distribution of the harmonic well. As is usual, the weighted norm $L^2(\pi_0(x,p)) := \{f|\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}|f|^2\pi_0(x,p)dxdp < \infty\}$. Informally, τ_e^ϵ is the time required for the distribution to be a factor $\epsilon \ll 1$ "closer" to the equilibrium distribution than in the initial condition. We define $\eta = 1-\epsilon$ as the **accuracy** of erasure. Lesser the ϵ, the closer the distribution of the particle is to the Gibbs distribution of the harmonic

well and hence more accurate the erasure. Consider the modified Langevin equation

$$\begin{aligned} m\,dx &= p\,dt \\ dp &= -\gamma p\,dt - \partial_x N_{A,B}(x)\;dt + \sqrt{2m\gamma k_B T}\,dW \end{aligned} \tag{6.2}$$

Here $N_{A,B}(x) = \frac{1}{2}m\omega_0^2\,(x-B)^2$ where $\omega_0 = \sqrt{\frac{8A}{mB^2}}$ is the harmonic potential that approximates well "0". Equation6.2 has the generator [114, pp. 182] given by

$$\mathcal{L} = \frac{p}{m}\partial_x - (\partial_x N_{A,B}(x))\,\partial_p + \gamma\left(-p\partial_p + k_B T\partial_p^2\right) \tag{6.3}$$

It is common knowledge that the following equation is true [115].

$$||\text{law}((x(t),p(t)) - \pi_0(x,p)||_{L^2(\pi_0(x,p))} \leq e^{-\lambda t}||\text{law}((x(0),p(0)) - \pi_0(x,p)||_{L^2(\pi_0(x,p))} \tag{6.4}$$

where λ is the first non-zero eigenvalue of the generator $\mathcal{L}$ given by Equation 6.3. Setting $e^{-\lambda t} = \epsilon$, we get useful **upper bounds** on the mixing time. In particular, we get

$$\tau_m^\epsilon \leq \frac{1}{\lambda}\ln\frac{1}{\epsilon} \tag{6.5}$$

For the sake of rough scaling, we will use $\tau_m^\epsilon \approx \frac{1}{\lambda}\ln\frac{1}{\epsilon}$ as an approximate estimate of the mixing time. It is important to note that the generator $\mathcal{L}$ is not self-adjoint and may possess imaginary eigenvalues. The rate of convergence in such cases will be determined by the real part of the eigenvalue. In fact using [114, pp. 200], the first non-zero eigenvalue of the generator is

$$\lambda = \frac{\gamma}{2} - \frac{1}{2}\sqrt{\gamma^2 - 4\omega_0^2} \tag{6.6}$$

In the underdamped limit when $\gamma \ll 2\omega_0$, we have $Re(\lambda) = \frac{\gamma}{2}$. Therefore $\tau_m^\epsilon \approx \frac{2}{\gamma}\log(\frac{1}{\epsilon})$ in the low friction regime. When friction is very high i.e. $\gamma \gg \omega_0$, we have $\lambda \approx \frac{\omega_0^2}{\gamma}$. As a result we get $\tau_m^\epsilon \approx \frac{\gamma}{\omega_0^2}\log\frac{1}{\epsilon}$. Thus our proxy for the mixing process produces $\tau_m^\epsilon \propto \frac{1}{\gamma}$ in the low friction regime and $\tau_m^\epsilon \propto \gamma$ in the high friction regime, consonant with the scaling and non-monotonicity observed using the erasure region criterion. As a consequence, using a convergence criterion for erasure would not change the physics of the problem,

merely perturbing the erasing time-scale quantitatively.

6.2.2 $4k_BT$ criterion for erasure region

Within the framework of the original "erasure region" criterion discussed in Chapter 2, we now consider the robustness of results to changing the numerical value of the criterion. Specifically, we here define the erasure region as all phase space points with total energy atleast $4k_BT$ below the barrier height. More formally,

$$\tau_e^{4k_BT} = \mathbb{E}\left[\inf\{t \geq 0 \mid x(t) < 0 \text{ and } H(x(t), p(t)) \leq A - 4k_BT\}\right] \tag{6.7}$$

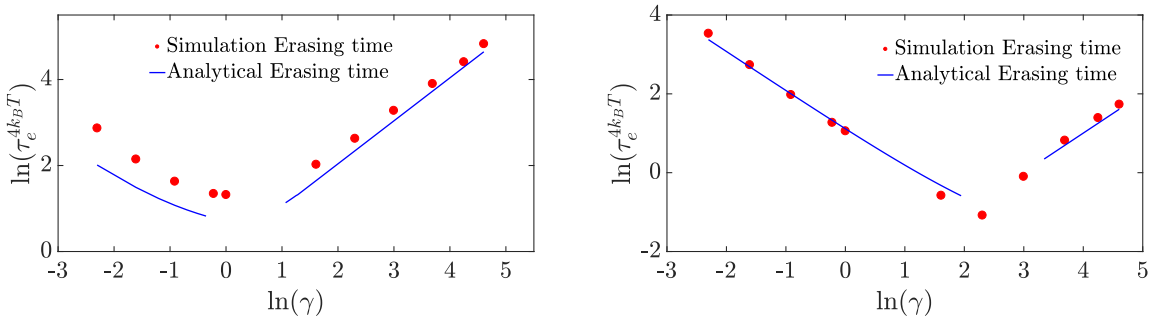
where $(x(t), p(t))$ is the solution to Equation 2.7 with the initial condition $(x(0), p(0)) \sim_{\text{law}} \pi_1(x, p)$. We now show that we get the same non-monotonicity and scaling of erasing time as a function of friction-coefficient that we got using the $3k_BT$ criterion. In particular, the erasing time scales as $\frac{1}{\gamma}$ in the low friction regime and scales as γ at high friction. Figure 6.3 illustrates this fact. Fits are performed using analytical expressions equivalent to those discussed in Chapter 2, but adjusted for the new numerical value of the boundary of the erasure region.

1. Low friction regime:

$$\tau_e^{4k_BT} \approx \sqrt{\frac{2mB}{F}} + \frac{1}{\gamma} \ln \frac{A + F \cdot B}{A - 4k_BT}. \tag{6.8}$$

2. High friction regime:

$$\tau_e^{4k_BT} \approx \frac{mB\gamma}{F} + \frac{2mB^2\gamma}{5A} \tag{6.9}$$



(a) $A = 10, F = 1$ **(b)** $A = 10, F = 100$

FIGURE 6.3: Evidence from simulation that the use of $4k_BT$ to define the erasure reason does not change the fundamental physics of the problem.

6.3 Regression and Cross Validation

We use cubic regression to interpolate between simulation data for both the reliability and erasing time-scales. Let $F' = \log(F)$ and $\gamma' = \log(\gamma)$. Then we use the following polynomials to fit the time-scales.

1. **Erasing Polynomial**:

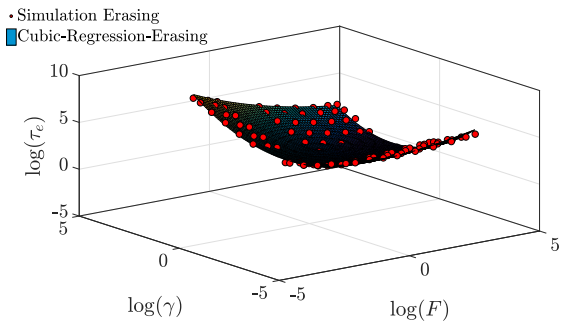
$$\begin{aligned}\log(\tau_e) = b_1 + b_2A^3 + b_3F'^3 + b_4\gamma'^3 + b_5A^2F' + b_6A'F'^2 + b_7F'^2\gamma' + b_8F'\gamma'^2 \\ + b_9A'^2\gamma' + b_{10}A\gamma'^2 + b_{11}A^2 + b_{12}F'^2 + b_{13}\gamma'^2 + b_{14}AF' + b_{15}F'\gamma' \\ + b_{16}A\gamma' + b_{17}A + b_{18}F' + b_{19}\gamma'\end{aligned} \tag{6.10}$$

2. **Reliability Polynomial**:

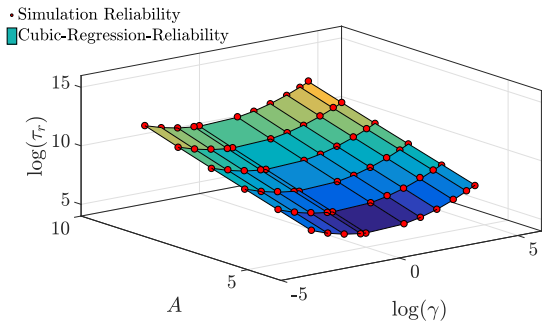
$$\log(\tau_r) = c_1 + c_2A^3 + c_3\gamma'^3 + c_4A^2\gamma' + c_5A\gamma'^2 + c_6A^2 + c_7\gamma'^2 + c_8A\gamma' + c_9A + c_{10}\gamma \tag{6.11}$$

, where the coefficients $b_1, b_2, \cdots b_{19}$ and $c_1, c_2, \cdots c_{10}$ are to be determined by regression.

Figure 6.4 gives a visual illustration of the fact that cubic fits offer a good approximation to the simulation results for both the erasing and reliability time-scales. In what follows,



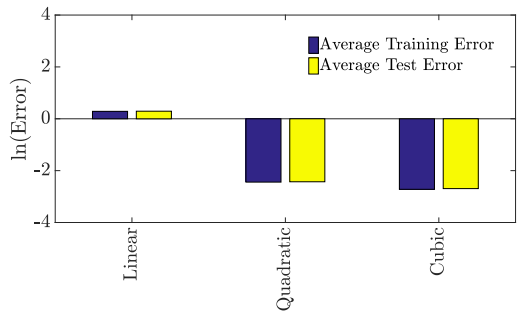
(a) Erasing time from simulation and cubic-regression for $A = 6$



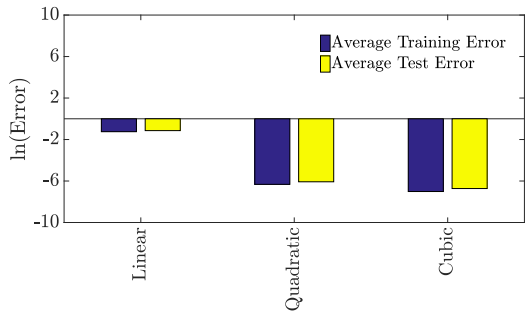
(b) Reliability time from simulation and cubic-regression

FIGURE 6.4

we present a more detailed and formal justification using cross-validation.



(a) Errors of regression fits for erasing time



(b) Errors of regression fits for reliability time

FIGURE 6.5

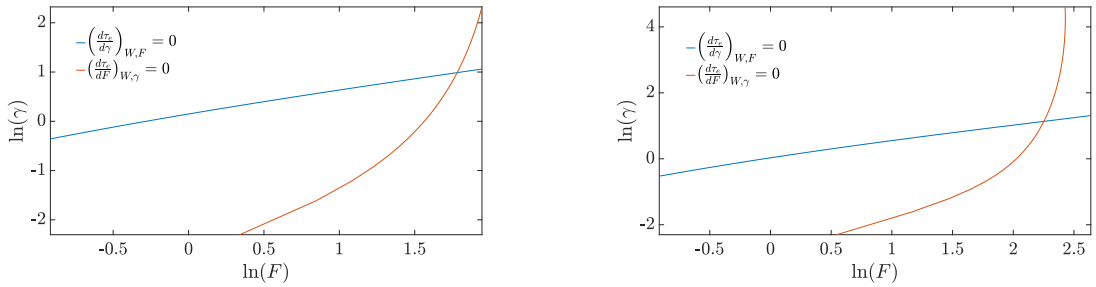
We perform "Leave-one-out" cross validation to justify the use of cubic regression. Figure 6.5 reports the mean square training and testing cross-validation errors corresponding to linear, quadratic and cubic fits. A lower value of the testing error indicates a good fit. Cubic regression has the lowest value of testing errors amongst the fits considered for both the reliability and erasing time-scales. Figure 6.5 confirms that the training and testing errors corresponding to cubic-regression for both the time-scales are roughly comparable(with the training error being slightly lower than the testing error). As a result, we can safely assume that the cubic polynomial does not over-fit the data and use it for modelling both the time-scales.

6.4 Locally trapped bits are uniquely trapped

In this section, we give typical plots for our family of controls that show no evidence of multiple local minima in erasing time within a level set of work. Towards this we let $F' = \log(F)$ and $\gamma' = \log(\gamma)$. Using the same form of regression polynomial as in Equation 6.10, but at constant work W, this translates to

$$\begin{aligned}\log(\tau_e) = {} & b_1 + b_2(W - e^{F'})^3 + b_3 F'^3 + b_4 \gamma'^3 + b_5(W - e^{F'})^2 F' + b_6(W - e^{F'})F'^2 \\ & + b_7 F'^2 \gamma' + b_8 F' \gamma'^2 + b_9(W - e^{F'})^2 \gamma' + b_{10}(W - e^{F'})\gamma'^2 + b_{11}(W - e^{F'})^2 \\ & + b_{12} F'^2 + b_{13}\gamma'^2 + b_{14}(W - e^{F'})F' + b_{15}F'\gamma' + b_{16}(W - e^{F'})\gamma' + b_{17}(W - e^{F'}) \\ & + b_{18}F' + b_{19}\gamma'\end{aligned} \tag{6.12}$$

Note that $\left(\frac{d\tau_e}{d\gamma'}\right)_{W,F'} = \gamma \left(\frac{d\tau_e}{d\gamma}\right)_{W,F}$ and $\left(\frac{d\tau_e}{dF'}\right)_{W,\gamma'} = F\left(\frac{d\tau_e}{dF}\right)_{W,\gamma}$. Therefore solving for $\left(\frac{d\tau_e}{d\gamma}\right)_{W,F} = \left(\frac{d\tau_e}{dF}\right)_{W,\gamma} = 0$ is equivalent to solving for $\left(\frac{d\tau_e}{d\gamma'}\right)_{W,F'} = \left(\frac{d\tau_e}{dF'}\right)_{W,\gamma'} = 0$. We solve $\left(\frac{d\tau_e}{d\gamma'}\right)_{W,F'} = \left(\frac{d\tau_e}{dF'}\right)_{W,\gamma'} = 0$ numerically and plot it in Figure 6.6. As illustrated by Figure 6.6, there is exactly one solution to the equations $\left(\frac{d\tau_e}{d\gamma}\right)_{W,F} = \left(\frac{d\tau_e}{dF}\right)_{W,\gamma} = 0$ within the broad range of parameters allowed, confirming our assumption that locally trapped bits are uniquely trapped.



(a) W=12 **(b)** W=19

FIGURE 6.6: Evidence that for our family of controls, locally trapped bits are uniquely trapped. The system $\left(\frac{d\tau_e}{d\gamma}\right)_{W,F} = \left(\frac{d\tau_e}{dF}\right)_{W,\gamma} = 0$ has exactly one solution within the broad range of parameters considered, which corresponds to a unique local minimum of erasing time in a level set of work. This is illustrated for work $W = 12$ and $W = 19$. The situation is representative for other values of work.

Chapter 7

Appendix B

In what follows, we will assume that we start with a fixed amount of transcription factor $[Z_{\text{tot}}]$, promoter $[P_{\text{tot}}]$, kinase $[Z_{\text{tot}}]$ and phosphatase $[Y_{\text{tot}}]$ unless specified otherwise.

7.1 Analytics for different Z dynamics

In this section we give analytical results for the steady-state dynamics corresponding to three different $\mathcal{U}$ subsystems discussed in Section 3.3 in Chapter 3.

1. **Fixed amount of Z:**

$$Z + P \underset{k_{\text{off}}}{\overset{k_{\text{on}}}{\rightleftharpoons}} C$$

In this case, we have the conservation laws $[Z^{\text{ss}}] + [C^{\text{ss}}] = [Z_{\text{tot}}]$ and $[P^{\text{ss}}] + [C^{\text{ss}}] = [P_{\text{tot}}]$. Solving for steady-state, we get $k_{\text{on}}[Z^{\text{ss}}]([P_{\text{tot}}] - [Z_{\text{tot}}] + [Z^{\text{ss}}]) = k_{\text{off}}([Z_{\text{tot}}] - [Z^{\text{ss}}])$ implying that

$$[Z^{\text{ss}}] = \frac{-\lambda \pm \sqrt{\lambda^2 + 4k_{\text{off}}k_{\text{on}}[Z_{\text{tot}}]}}{2k_{\text{on}}},$$

where $\lambda = k_{\text{off}} + k_{\text{on}}([P_{\text{tot}}] - [Z_{\text{tot}}])$. We choose the solution that makes physical sense for a given set of parameters i.e. the solution that satisfies $[Z^{\text{ss}}] \geq 0$ and $[Z^{\text{ss}}] \leq [Z_{\text{tot}}]$. The metric for retroactivity translates to

$$\mathcal{R} = \left| 1 - \frac{[Z^{\text{ss}}]}{[Z^{\text{ss}}_{\mathcal{D},\mathcal{I}\to\emptyset}]} \right| = 1 - \frac{[Z^{\text{ss}}]}{[Z_{\text{tot}}]} \tag{7.1}$$

2. **Constant birth/death dynamics:**

$$\phi \underset{\delta}{\overset{k}{\rightleftharpoons}} Z$$
$$Z + P \underset{k_{\text{off}}}{\overset{k_{\text{on}}}{\rightleftharpoons}} C$$

In this case, solving for steady-state we get $k - \delta[Z^{\text{ss}}] - k_{\text{on}}[Z^{\text{ss}}]([P_{\text{tot}}] - [C^{\text{ss}}]) + k_{\text{off}}[C^{\text{ss}}] = 0$ and $k_{\text{on}}[Z^{\text{ss}}]([P_{\text{tot}}] - [C^{\text{ss}}]) - k_{\text{off}}[C^{\text{ss}}] = 0$. Therefore $[Z^{\text{ss}}] = \frac{k}{\delta}$. In addition, note that $[Z^{\text{ss}}_{\mathcal{D},\mathcal{I}\to\emptyset}] = \frac{k}{\delta}$ implying that the retroactivity metric is

$$\mathcal{R} = \left| 1 - \frac{[Z^{\text{ss}}]}{[Z^{\text{ss}}_{\mathcal{D},\mathcal{I}\to\emptyset}]} \right| = 0. \tag{7.2}$$

3. **Active/Inactive forms of Z:**

$$Z_0 \underset{k_{\text{in}}}{\overset{k_{\text{ac}}}{\rightleftharpoons}} Z$$
$$Z + P \underset{k_{\text{off}}}{\overset{k_{\text{on}}}{\rightleftharpoons}} C$$

In this case, we start with a fixed amount of total Z, say $[Z'_{\text{tot}}]$. Solving for steady-state, we get

$$k_{\text{ac}}([Z'_{\text{tot}}] - [Z^{\text{ss}}] - [C^{\text{ss}}]) = k_{\text{in}}[Z^{\text{ss}}] \tag{7.3}$$

and $k_{\text{on}}[Z^{\text{ss}}]([p_{\text{tot}}] - [C^{\text{ss}}]) = k_{\text{off}}[C^{\text{ss}}]$ implying that

$$[Z^{\text{ss}}] = \frac{-\mu \pm \sqrt{\mu^2 + 4k_{\text{ac}}(k_{\text{ac}}+k_{\text{in}})k_{\text{off}}k_{\text{on}}[Z'_{\text{tot}}]}}{2(k_{\text{ac}}+k_{\text{in}})k_{\text{on}}}$$

where

$$\mu = k_{\text{in}}k_{\text{off}} + k_{\text{ac}}(k_{\text{off}} + k_{\text{on}}([P_{\text{tot}}] - [Z'_{\text{tot}}])) \tag{7.4}$$

As in case 1, we choose only those solutions that make physical sense i.e. those which satisfy $[Z^{\text{ss}}] \geq 0$ and $[Z^{\text{ss}}] \leq [Z'_{\text{tot}}]$. Note that $k_{\text{in}}[Z^{\text{ss}}_{\mathcal{D},\mathcal{I}\to\emptyset}] = k_{\text{ac}}([Z'_{\text{tot}}] - [Z^{\text{ss}}_{\mathcal{D},\mathcal{I}\to\emptyset}])$. Therefore,

$$[Z^{\text{ss}}_{\mathcal{D},\mathcal{I}\to\emptyset}] = \frac{k_{\text{ac}}}{k_{\text{ac}} + k_{\text{in}}}[Z'_{\text{tot}}] \tag{7.5}$$

Choosing $[Z^{\text{ss}}_{\mathcal{D},\mathcal{I}\to\emptyset}] = \frac{k}{\delta}$ allows the system to be compared sensibly to other designs of $\mathcal{U}$ with the same behaviour in this limit. In this case, the metric for retroactivity translates to

$$\mathcal{R} = \left| 1 - \frac{[Z^{\text{ss}}]}{[Z^{\text{ss}}_{\mathcal{D},\mathcal{I}\to\emptyset}]} \right| = 1 - \frac{\delta[Z^{\text{ss}}]}{k}. \tag{7.6}$$

Subsequently, Equation 7.5 becomes

$$\frac{k_{\text{ac}}}{k_{\text{ac}} + k_{\text{in}}}[Z'_{\text{tot}}] = \frac{k}{\delta} \tag{7.7}$$

In the limit when $k_{\text{in}} \gg k_{\text{ac}}$, Equation 7.7 reduces to

$$\frac{k_{\text{ac}}}{k_{\text{in}}}[Z'_{\text{tot}}] = \frac{k}{\delta} \tag{7.8}$$

When $k_{\text{in}} \gg k_{\text{ac}}$, one can safely assume that $[Z'_{\text{tot}}] - [Z^{\text{ss}}] - [C^{\text{ss}}] \sim [Z'_{\text{tot}}]$ since all of the Z is essentially present in the form of Z_0. Therefore, Equation 7.3 reduces to

$$\frac{k_{\text{ac}}}{k_{\text{in}}}[Z'_{\text{tot}}] = [Z^{\text{ss}}] \tag{7.9}$$

Comparing Equation 7.8 and Equation 7.9, we obtain $[Z^{\text{ss}}] = \frac{k}{\delta}$, thereby reproducing the steady-state concentration of constant birth-death process.

7.2 Effect of decreasing the coupling to the push-pull

Recall the microscopically reversible push-pull motif from Chapter 3:

$$\begin{aligned} Z + X &\underset{\beta_2}{\overset{\beta_1}{\rightleftharpoons}} C_1 \underset{\frac{\epsilon k_1 \beta_1}{\beta_2}}{\overset{k_1}{\rightleftharpoons}} X^* + Z \\ Y + X^* &\underset{\alpha_2}{\overset{\alpha_1}{\rightleftharpoons}} C_2 \underset{\frac{\epsilon k_2 \alpha_1}{\alpha_2}}{\overset{k_2}{\rightleftharpoons}} X + Y \\ X^* + p &\underset{k_{\text{off}}}{\overset{k_{\text{on}}}{\rightleftharpoons}} C. \end{aligned} \tag{7.10}$$

Here, $0 \leq \epsilon \leq 1$ is the parameter that defines the degree of microscopic reversibility. We analyse the effect of repeatedly reducing the coupling to the push-pull motif to match a given input/output curve. We show that the steady-state output of a push-pull motif is a function of the ratio $\frac{\beta_1}{\alpha_1}$ and not α_1 and β_1 individually, in the low retroactivity limit. Further the retroactivity and power is directly proportional to β_1 at fixed ratio $\frac{\beta_1}{\alpha_1}$. Our analysis is divided into two cases:

1. **Microscopically irreversible limit:** A push-pull motif coupled to fuel with an infinite free energy corresponds to case $\epsilon = 0$. Specifically, we have the following network:

$$\begin{aligned} Z + X &\underset{\beta_2}{\overset{\beta_1}{\rightleftharpoons}} C_1 \xrightarrow{k_1} X^* + Z \\ Y + X^* &\underset{\alpha_2}{\overset{\alpha_1}{\rightleftharpoons}} C_2 \xrightarrow{k_2} X + Y \\ X^* + p &\underset{k_{\text{off}}}{\overset{k_{\text{on}}}{\rightleftharpoons}} \end{aligned} \tag{7.11}$$

   As we reduce the coupling to the push-pull motif by making α_1 and β_1 sufficiently small, one can approximately ignore sequestration into complexes relative to $[C^{\text{ss}}]$,

$[X^{\text{ss}}]$ and $[X^{*\text{ss}}]$ and the network essentially boils down to the following:

$$X \underset{\frac{\alpha_1 k_2 [Y_{\text{tot}}]}{k_2+\alpha_2}}{\overset{\frac{k_1\beta_1[Z_{\text{tot}}]}{\beta_2+k_1}}{\rightleftharpoons}} X^*$$

$$X^* + P \underset{k_{\text{off}}}{\overset{k_{\text{on}}}{\rightleftharpoons}} C. \tag{7.12}$$

Solving for steady-state, we get

$$[X^{\text{ss}}]\tfrac{k_1\beta_1[Z_{\text{tot}}]}{\beta_2+k_1} = [X^{*\text{ss}}]\tfrac{\alpha_1 k_2[Y_{\text{tot}}]}{k_2+\alpha_2} \tag{7.13}$$

and

$$k_{\text{on}}[X^{*\text{ss}}]([P_{\text{tot}}] - [C^{\text{ss}}]) = k_{\text{off}}[C^{\text{ss}}] \tag{7.14}$$

with the conservation relation $[X^{\text{ss}}] + [X^{*\text{ss}}] + [C^{\text{ss}}] = [X_{\text{tot}}]$. Therefore, we have

$$[C^{\text{ss}}] = \tfrac{f(r) \pm \sqrt{f^2(r) - 4r^2 k_{\text{on}}^2 [X_{\text{tot}}][P_{\text{tot}}]}}{2rk_{\text{on}}} \tag{7.15}$$

where $r = \frac{\frac{k_1\beta_1[Z_{\text{tot}}]}{\beta_2+k_1}}{\frac{\alpha_1 k_2[Y_{\text{tot}}]}{k_2+\alpha_2}}$ and $f(r) = k_{\text{off}} + r(k_{\text{on}}([P_{\text{tot}}] + [X_{\text{tot}}]) + k_{\text{off}})$, justifying Equation 3.16 in Chapter 3. It follows that both $[X^{\text{ss}}]$ and $[X^{*\text{ss}}]$ are functions of α_1 and β_1 through the ratio $r = \frac{\beta_1}{\alpha_1}$. Solving for steady-state of $[C_1]$, we get $\beta_1[Z^{\text{ss}}][X^{*\text{ss}}] - (\beta_2 + k_1)[C_1^{\text{ss}}] = 0$, implying that $[C_1^{\text{ss}}] = \frac{\beta_1[Z^{\text{ss}}][X^{*\text{ss}}]}{(\beta_2+k_1)}$. As a consequence $C_1^{\text{ss}} \propto [Z^{\text{ss}}]\beta_1$ at fixed r. This implies that $\mathcal{R} = \frac{[C_1^{\text{ss}}]}{[Z_{\text{tot}}]} \propto \beta_1(1+\beta_1)^{-1} \approx \beta_1$ for sufficiently small β_1 and flux $\Psi = k_1[C_1^{\text{ss}}] \propto \beta_1$ at fixed r, justifying Equation 3.19 in Chapter 3.

2. **Finite free energy of fuel molecules:** Explicitly incorporating the presence of microscopically reversible reactions in the push-pull motif corresponds to the case $0 < \epsilon \leq 1$. As in the case of infinite free energy, making α_1 and β_1 sufficiently small amounts to neglecting sequestration into complexes relative to $[C^{\text{ss}}]$, $[X^{\text{ss}}]$

and $[X^{*\mathrm{ss}}]$, giving the following network:

$$X \underset{\frac{\alpha_1 k_2[Y_{\mathrm{tot}}]}{k_2+\alpha_2}+\epsilon\frac{k_1\beta_1[Z_{\mathrm{tot}}]}{\beta_2+k_1}}{\overset{\frac{k_1\beta_1[Z_{\mathrm{tot}}]}{\beta_2+k_1}+\epsilon\frac{\alpha_1 k_2[Y_{\mathrm{tot}}]}{k_2+\alpha_2}}{\rightleftharpoons}} X^*$$

$$X^* + p \underset{k_{\mathrm{off}}}{\overset{k_{\mathrm{on}}}{\rightleftharpoons}} C$$

Solving for steady-state, we get

$$X^{\mathrm{ss}}\left(\frac{k_1\beta_1[Z_{\mathrm{tot}}]}{\beta_2+k_1}+\epsilon\frac{\alpha_1 k_2[Y_{\mathrm{tot}}]}{k_2+\alpha_2}\right) = [X^{*\mathrm{ss}}]\left(\frac{\alpha_1 k_2[Y_{\mathrm{tot}}]}{k_2+\alpha_2}+\epsilon\frac{k_1\beta_1[Z_{\mathrm{tot}}]}{\beta_2+k_1}\right) \tag{7.16}$$

and

$$k_{\mathrm{on}}[X^{*\mathrm{ss}}]([P_{\mathrm{tot}}]-[C^{\mathrm{ss}}]) = k_{\mathrm{off}}[C^{\mathrm{ss}}] \tag{7.17}$$

with the conservation relation $[X^{\mathrm{ss}}]+[X^{*\mathrm{ss}}]+[C^{\mathrm{ss}}]=[X_{\mathrm{tot}}]$. Let $r=\frac{\frac{k_1\beta_1[Z_{\mathrm{tot}}]}{\beta_2+k_1}}{\frac{\alpha_1 k_2[Y_{\mathrm{tot}}]}{k_2+\alpha_2}}$ and $r'=\frac{r+\epsilon}{r\epsilon+1}$. Therefore, we have

$$[C^{\mathrm{ss}}] = \frac{f(r')\pm\sqrt{f^2(r')-4r'^2k_{\mathrm{on}}^2[X_{\mathrm{tot}}][P_{\mathrm{tot}}]}}{2r'k_{\mathrm{on}}} \tag{7.18}$$

where $f(r')=k_{\mathrm{off}}+r'(k_{\mathrm{on}}([P_{\mathrm{tot}}]+[X_{\mathrm{tot}}])+k_{\mathrm{off}})$. Solving for $[C_1^{\mathrm{ss}}]$, we get $\beta_1[Z^{\mathrm{ss}}][X^{*\mathrm{ss}}]-(\beta_2+k_1)[C_1^{\mathrm{ss}}]+\frac{\epsilon k_1\beta_1}{\beta_2}[X^{*\mathrm{ss}}]=0$, implying that

$$[C_1^{\mathrm{ss}}] = \frac{\beta_1[Z^{\mathrm{ss}}][X^{*\mathrm{ss}}]+\frac{\epsilon k_1\beta_1}{\beta_2}[X^{*\mathrm{ss}}]}{(\beta_2+k_1)}. \tag{7.19}$$

Since both $[X^{\mathrm{ss}}]$ and $[X^{*\mathrm{ss}}]$ depend only on the ratio r', we get that $\mathcal{R}=\frac{[C_1^{\mathrm{ss}}]}{[Z_{\mathrm{tot}}]}\propto\beta_1(1+\beta_1)^{-1}\approx\beta_1$ for sufficiently small β_1 at fixed r' and power $w=\Psi\Delta G_{\mathrm{ATP}}\propto\beta_1$ justifying equation 3.26 in Chapter 3.

7.3 Effectively irreversible push-pull motifs

We show that push-pull networks consuming free energy per cycle beyond a certain threshold are essentially equivalent to those without the microscopically reversible reactions, for the purposes of the steady-state concentrations. In our system, this threshold is $\sim 4k_BT$.

Figure 7.1 illustrates this fact for certain sets of parameters. Recall from Chapter 3 that

FIGURE 7.1: For a push-pull network consuming free energy in excess of $-2k_BT\ln\epsilon \sim 4k_BT$, the presence of microscopically reversible reactions has negligible effect on the steady-state output as demonstrated for two specific systems. Parameters used for the model: *a*) $[X_{\text{tot}}] = [Y_{\text{tot}}] = [Z_{\text{tot}}] = 50, [P_{\text{tot}}] = 100, \alpha_1 = \beta_1 = 0.1, \alpha_2 = \beta_2 = k_2 = k_{\text{on}} = k_{\text{off}} = 10, k_1 = 1$. *b*) $[X_{\text{tot}}] = 200, [Y_{\text{tot}}] = [Z_{\text{tot}}] = 100, [P_{\text{tot}}] = 100, \alpha_1 = \beta_1 = \alpha_2 = \beta_2 = k_1 = k_2 = 1, k_{\text{on}} = k_{\text{off}} = 10$.

the free energy of a push-pull motif having microscopically reversible reactions is given by $\Delta G_{\text{ATP}} = 2k_BT\ln\epsilon$.

7.4 Microscopic reversibility constrains overall range of the input-output function

As observed earlier in the main text, addition of microscopically reversible reactions restricts the input-output range. In this section, we derive quantitative bounds to illustrate this fact. In particular, we derive Equation 3.23 in the main text. Recall Equation 7.16. We rewrite it here for convenience.

$$[X^{\text{ss}}]\left(\frac{k_1\beta_1[Z_{\text{tot}}]}{\beta_2+k_1}+\epsilon\frac{\alpha_1k_2[Y_{\text{tot}}]}{k_2+\alpha_2}\right)=[X^{*\text{ss}}]\left(\frac{\alpha_1k_2[Y_{\text{tot}}]}{k_2+\alpha_2}+\epsilon\frac{k_1\beta_1[Z_{\text{tot}}]}{\beta_2+k_1}\right) \tag{7.20}$$

Dividing this Equation throughout by $[Y_{\text{tot}}]$ gives,

$$[X^{\text{ss}}]\left(\frac{k_1\beta_1[Z_{\text{tot}}]}{(\beta_2+k_1)[Y_{\text{tot}}]}+\epsilon\frac{\alpha_1k_2}{k_2+\alpha_2}\right)=[X^{*\text{ss}}]\left(\frac{\alpha_1k_2}{k_2+\alpha_2}+\epsilon\frac{k_1\beta_1[Z_{\text{tot}}]}{(\beta_2+k_1)[Y_{\text{tot}}]}\right) \tag{7.21}$$

Therefore, we get

$$\frac{[X^{*\text{ss}}]}{[X^{\text{ss}}]} = \frac{\left(\frac{k_1\beta_1[Z_{\text{tot}}]}{(\beta_2+k_1)[Y_{\text{tot}}]} + \epsilon\frac{\alpha_1 k_2}{k_2+\alpha_2}\right)}{\left(\frac{\alpha_1 k_2}{k_2+\alpha_2} + \epsilon\frac{k_1\beta_1[Z_{\text{tot}}]}{(\beta_2+k_1)[Y_{\text{tot}}]}\right)} \tag{7.22}$$

Since $\epsilon \leq 1$, we get

$$\epsilon \leq \frac{\left(\frac{k_1\beta_1[Z_{\text{tot}}]}{(\beta_2+k_1)[Y_{\text{tot}}]} + \epsilon\frac{\alpha_1 k_2}{k_2+\alpha_2}\right)}{\left(\frac{\alpha_1 k_2}{k_2+\alpha_2} + \epsilon\frac{k_1\beta_1[Z_{\text{tot}}]}{(\beta_2+k_1)[Y_{\text{tot}}]}\right)} \tag{7.23}$$

Hence $\epsilon \leq \frac{[X^{*\text{ss}}]}{[X^{\text{ss}}]}$. In the limit $[Y_{\text{tot}}] \gg [Z_{\text{tot}}]$, we get equality i.e. $\epsilon = \frac{[X^{*\text{ss}}]}{[X^{\text{ss}}]}$. Similarly, dividing Equation 7.20 by $[Z_{\text{tot}}]$ gives,

$$[X^{\text{ss}}]\left(\frac{k_1\beta_1}{\beta_2+k_1} + \epsilon\frac{\alpha_1 k_2[Y_{\text{tot}}]}{(k_2+\alpha_2)[Z_{\text{tot}}]}\right) = [X^{*\text{ss}}]\left(\frac{\alpha_1 k_2[Y_{\text{tot}}]}{(k_2+\alpha_2)[Z_{\text{tot}}]} + \epsilon\frac{k_1\beta_1}{\beta_2+k_1}\right) \tag{7.24}$$

Therefore, we get

$$\frac{[X^{*\text{ss}}]}{[X^{\text{ss}}]} = \frac{\left(\frac{k_1\beta_1}{\beta_2+k_1} + \epsilon\frac{\alpha_1 k_2[Y_{\text{tot}}]}{(k_2+\alpha_2)[Z_{\text{tot}}]}\right)}{\left(\frac{\alpha_1 k_2[Y_{\text{tot}}]}{(k_2+\alpha_2)[Z_{\text{tot}}]} + \epsilon\frac{k_1\beta_1}{\beta_2+k_1}\right)} \tag{7.25}$$

Using the fact that $\epsilon \leq 1$, we get

$$\frac{\left(\frac{k_1\beta_1}{\beta_2+k_1} + \epsilon\frac{\alpha_1 k_2[Y_{\text{tot}}]}{(k_2+\alpha_2)[Z_{\text{tot}}]}\right)}{\left(\frac{\alpha_1 k_2[Y_{\text{tot}}]}{(k_2+\alpha_2)[Z_{\text{tot}}]} + \epsilon\frac{k_1\beta_1}{\beta_2+k_1}\right)} \leq \frac{1}{\epsilon} \tag{7.26}$$

Hence $\frac{[X^{*\text{ss}}]}{[X^{\text{ss}}]} \leq \frac{1}{\epsilon}$. In the limit $[Z_{\text{tot}}] \gg [Y_{\text{tot}}]$, we get equality i.e. $\frac{[X^{*\text{ss}}]}{[X^{\text{ss}}]} = \frac{1}{\epsilon}$. In addition, it is easy to check that the fraction $\frac{[X^{*\text{ss}}]}{[X^{\text{ss}}]} = \frac{\left(\frac{k_1\beta_1}{\beta_2+k_1} + \epsilon\frac{\alpha_1 k_2[Y_{\text{tot}}]}{(k_2+\alpha_2)[Z_{\text{tot}}]}\right)}{\left(\frac{\alpha_1 k_2[Y_{\text{tot}}]}{(k_2+\alpha_2)[Z_{\text{tot}}]} + \epsilon\frac{k_1\beta_1}{\beta_2+k_1}\right)}$ is monotonic in $\frac{[Y_{\text{tot}}]}{[Z_{\text{tot}}]}$. This implies that the relation $\epsilon \leq \frac{[X^{*\text{ss}}]}{[X^{\text{ss}}]} \leq \frac{1}{\epsilon}$ holds over the entire range of $\frac{[Y_{\text{tot}}]}{[Z_{\text{tot}}]}$, justifying Equation 3.23 in the main text.

7.5 Retroactivity and rate of free-energy consumption for randomly parameterised push-pull motifs

In this section we show that the results presented in Fig. 3.7 of Chapter 3 – namely that it is possible to reproduce an input-output relation at weaker coupling, thereby reducing retroactivity and free energy consumption, is true for a good proportion of randomly generated systems. Push-pull motifs having infinite free energy are a limiting case of generic push-pull networks possessing microscopically reversible reactions. As noted earlier, having an infinite free energy corresponds to putting $\epsilon = 0$, where ϵ is the parameter that quantifies the amount of microscopic reversibility. It therefore suffices to consider push-pull motifs having finite free energy. We repeat the plots of Fig. 3.7 from Chapter 3, using randomly generated parameters. Specifically, we consider 10 systems randomly chosen from parameter distributions:$X_{\text{tot}}, Y_{\text{tot}} \sim \text{U}[1, 200], P_{\text{tot}} \sim \text{U}[1, X_{\text{tot}}], \alpha_1, \beta_1, k_{\text{on}} \sim 10^{\text{U}[-2,0]}, \alpha_2, \beta_2, k_1, k_2, k_{\text{off}} \sim 10^{\text{U}[-1,1]}, \epsilon \sim 10^{\text{U}[-3,-1]}$. Here, U indicates a uniform distribution and all samples are independent. The results are plotted in the following figures, demonstrating that often (though not always) it is possible to get a very close match to the input-output curve at weaker coupling. Moreover, even when the matching of the input-output curve is moderate, both retroactivity and fuel consumption still decrease.

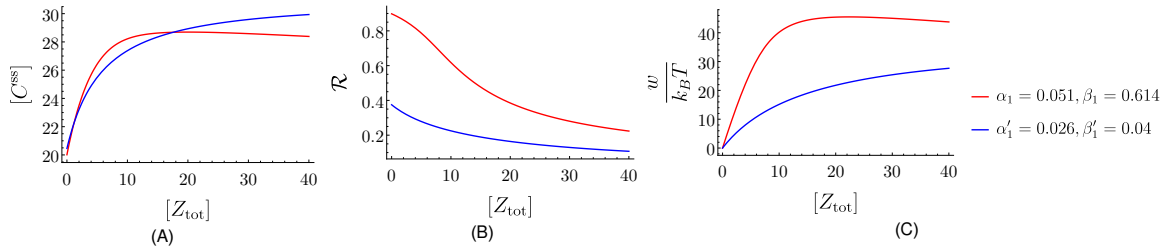


FIGURE 7.2: $[X_{\text{tot}}] = 41.198, [Y_{\text{tot}}] = 108.948, [P_{\text{tot}}] = 36.176, \alpha_1 = 0.051, \alpha_2 = 0.897, \beta_1 = 0.614, \beta_2 = 0.291, k_1 = 0.950, k_2 = 3.184, k_{\text{on}} = 0.517, k_{\text{off}} = 0.203, \epsilon = 0.026$.

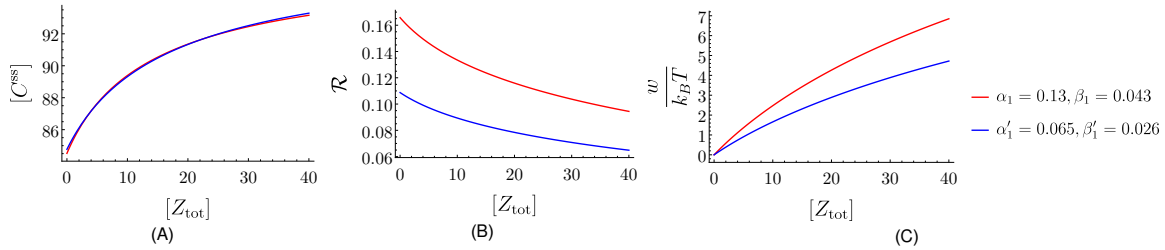


FIGURE 7.3: $[X_{\text{tot}}] = 124.747, [Y_{\text{tot}}] = 5.810, [P_{\text{tot}}] = 101.106, \alpha_1 = 0.126, \alpha_2 = 0.177, \beta_1 = 0.043, \beta_2 = 6.802, k_1 = 0.391, k_2 = 1.600, k_{\text{on}} = 0.366, k_{\text{off}} = 0.215, \epsilon = 0.065.$

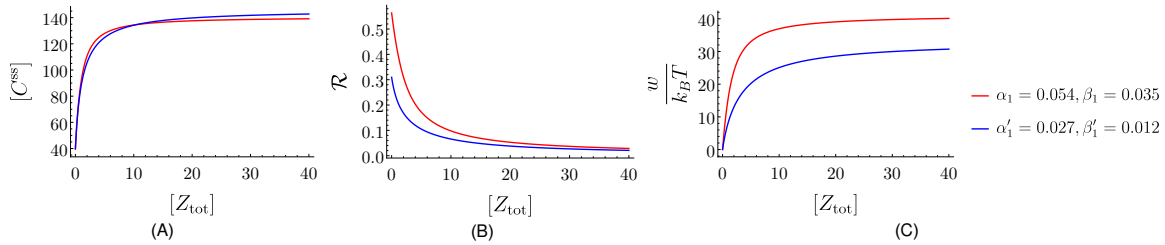


FIGURE 7.4: $[X_{\text{tot}}] = 182.861, [Y_{\text{tot}}] = 57.132, [P_{\text{tot}}] = 171.82, \alpha_1 = 0.054, \alpha_2 = 0.137, \beta_1 = 0.035, \beta_2 = 0.283, k_1 = 3.383, k_2 = 0.105, k_{\text{on}} = 0.432, k_{\text{off}} = 0.646, \epsilon = 0.027.$

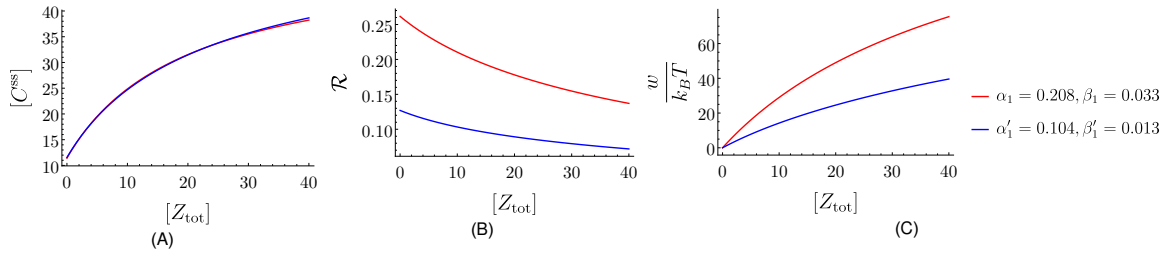


FIGURE 7.5: $[X_{\text{tot}}] = 95.631, [Y_{\text{tot}}] = 147.689, [P_{\text{tot}}] = 72.11, \alpha_1 = 0.208, \alpha_2 = 9.877, \beta_1 = 0.033, \beta_2 = 6.076, k_1 = 1.450, k_2 = 0.766, k_{\text{on}} = 0.418, k_{\text{off}} = 1.559, \epsilon = 0.104.$

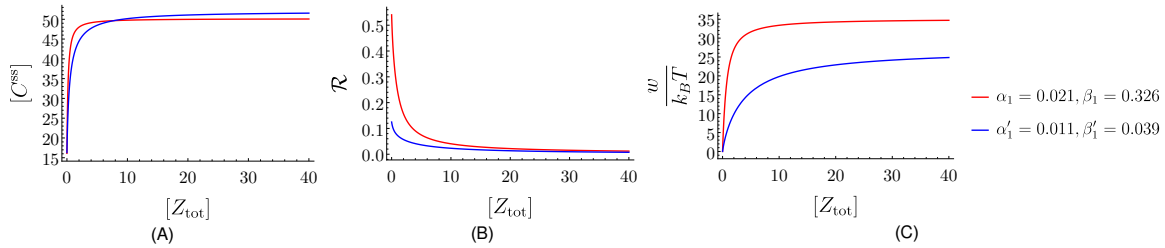


FIGURE 7.6: $[X_{\text{tot}}] = 84.151, [Y_{\text{tot}}] = 100.826, [P_{\text{tot}}] = 56.871, \alpha_1 = 0.021, \alpha_2 = 0.438, \beta_1 = 0.326, \beta_2 = 9.358, k_1 = 8.781, k_2 = 0.145, k_{\text{on}} = 0.323, k_{\text{off}} = 0.389, \epsilon = 0.011.$

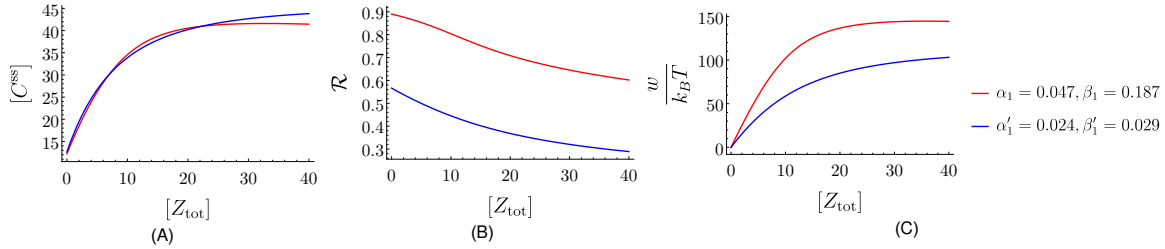


FIGURE 7.7: $[X_{\text{tot}}] = 138.883, [Y_{\text{tot}}] = 67.044, [P_{\text{tot}}] = 67.68, \alpha_1 = 0.047, \alpha_2 = 0.942, \beta_1 = 0.187, \beta_2 = 0.264, k_1 = 2.308, k_2 = 0.717, k_{\text{on}} = 0.041, k_{\text{off}} = 0.822, \epsilon = 0.024.$

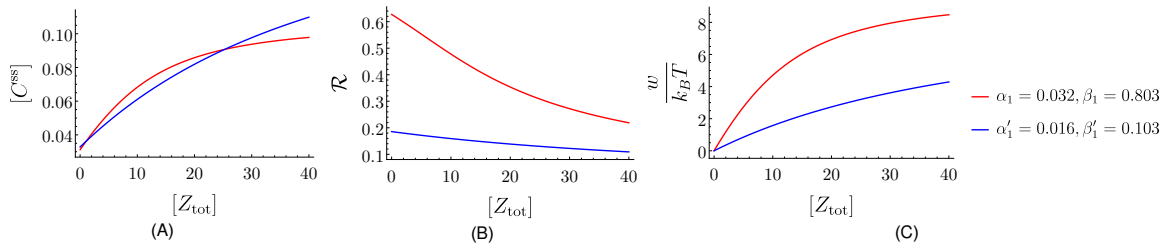


FIGURE 7.8: $[X_{\text{tot}}] = 14.304, [Y_{\text{tot}}] = 58.104, [P_{\text{tot}}] = 6.824, \alpha_1 = 0.032, \alpha_2 = 0.503, \beta_1 = 0.803, \beta_2 = 5.777, k_1 = 0.145, k_2 = 0.671, k_{\text{on}} = 0.069, k_{\text{off}} = 6.124, \epsilon = 0.016.$

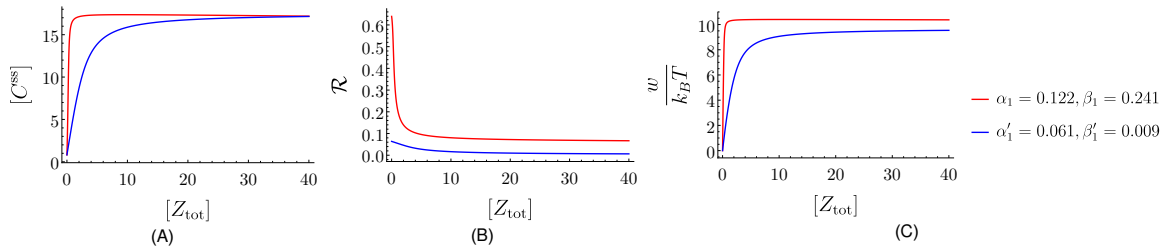


FIGURE 7.9: $[X_{\text{tot}}] = 58.564, [Y_{\text{tot}}] = 4.7, [P_{\text{tot}}] = 23.929, \alpha_1 = 0.122, \alpha_2 = 0.128, \beta_1 = 0.240, \beta_2 = 1.023, k_1 = 6.555, k_2 = 0.252, k_{\text{on}} = 0.680, k_{\text{off}} = 9.066, \epsilon = 0.061.$

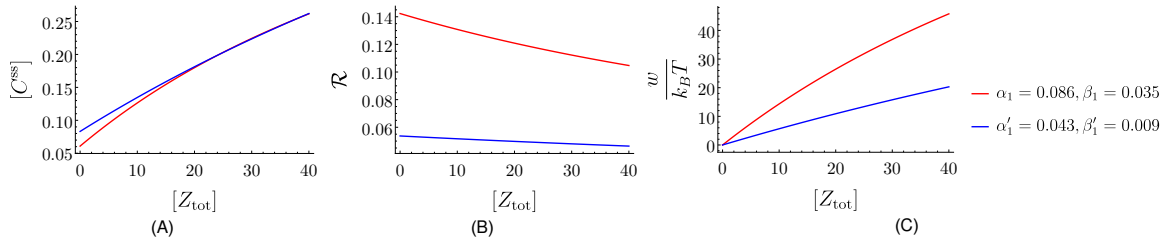


FIGURE 7.10: $[X_{\text{tot}}] = 18.954, [Y_{\text{tot}}] = 166.734, [P_{\text{tot}}] = 4.529, \alpha_1 = 0.086, \alpha_2 = 0.131, \beta_1 = 0.035, \beta_2 = 0.524, k_1 = 1.245, k_2 = 7.211, k_{\text{on}} = 0.438, k_{\text{off}} = 3.268, \epsilon = 0.043.$

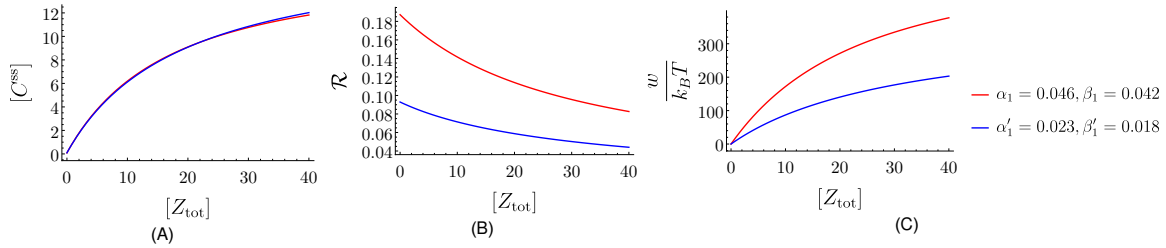


FIGURE 7.11: $[X_{\text{tot}}] = 54.201, [Y_{\text{tot}}] = 50.055, [P_{\text{tot}}] = 39.431, \alpha_1 = 0.045, \alpha_2 = 0.295, \beta_1 = 0.042, \beta_2 = 0.135, k_1 = 9.537, k_2 = 6.173, k_{\text{on}} = 0.038, k_{\text{off}} = 1.337, \epsilon = 0.023.$