

Abstract

Bacteria can exploit gene expression noise to create phenotypic heterogeneity in a population, which may allow a fraction of the population to survive a sudden change in environmental conditions. The phenotypic variations are typically caused by stochastically turning on or off a specific set of genes. Alternative sigma factors (common regulatory proteins of bacterial stress responses) often drive this variability. Recent studies show that sigma factors may adopt dynamical behaviours, including stochastic pulsing and bistability, to modulate their cellular abundance. It has also been shown that biochemical ultrasensitivity is important for several behaviours, such as bistability, but its exact role in the sigma factor circuit remains unclear. Here I simulate a simplified mechanistic model of an alternative sigma factor circuit using the Gillespie algorithm. The model features a mixed self-activation and negative feedback loop with time delay. I first show that a range of dynamical behaviours is produced. Then I observe that without ultrasensitivity, several dynamical behaviours are significantly weakened or cannot be maintained. As many sigma factor circuits do not encode ultrasensitivity, it raises the question of how stochastic switching of gene expression is achieved. Sigma factors must bind to the RNA polymerase (RNAP) core enzymes to function, and recent studies show that the competition between sigma factors for the limited pool of RNAP cores shapes sigma factor dynamics. In light of that, I propose two new mechanisms for a non-ultrasensitive sigma factor to maintain bistability. First, under strong competition, a bistable ultrasensitive circuit may force a non-ultrasensitive circuit to adopt bistability through the limitation of shared resources. Second, for two non-ultrasensitive circuits with low binding affinity of sigma factors to RNAP cores, the circuit is locked to a zero-state and turns on by rare binding events through a self-activation. I show that in the latter scheme, the variations in initial activation times are reduced as sigma factor-core RNAP binding affinity increases, which has implications for a previous model of heterogeneous activation of the alternative sigma factor σ^V in *Bacillus subtilis*. Since the core sigma factor circuit structure is conserved across many bacteria, this research may shed light on a general strategy for the bacteria population to create heterogeneity.

Chapter 1

Introduction

1.1 Gene expression noise

Stochastic fluctuations ("noise") of the number of proteins or mRNAs is common in gene expression. Due to the randomness inherent in the kinetics of single molecules, and environmental fluctuations, noise is generally unavoidable and widely exists in biological processes. Noise in gene expression is prominent since many players of the process are typically of low copy numbers. In a pioneering study, the authors first proposed the concept of intrinsic noise and extrinsic noise, and demonstrated measuring the two categories of noise by examining the correlation between two distinguishable fluorescent reporters driven by identical promoters [1]. Intrinsic noise refers to the noise from the expression mechanism per se and extrinsic noise includes that from upstream components, e.g. the influence of different cell states or the cell cycle, and from the microscopic environment. The study shows that both intrinsic and extrinsic noise are important in setting the cell-cell variations. Single-molecule techniques further revealed the origin of noise in gene expression. Studies with the sensitivity to track single protein production events in *Escherichia coli* show that proteins are expressed in "bursts" [2, 3]. Each protein burst is generated from a single mRNA, the production of which is also noisy. These studies give insight into the mechanism of gene expression noise and provide a theoretical framework to address noise.

Noise can propagate through the gene regulatory network and cells leverage certain network structures, or motifs, to regulate noise. Negative autoregulation filters fluctuations and increases the robustness of gene expression [4]. Coherent feed forward loops (FFL) can filter out either activation pulses or inactivation pulses depending on whether the output node acts as an AND gate or OR gate. Intuitively, since the coherent FFL consists of one direct regulation and one regulation in delay, the input has to persist

longer than the delay to drive the output [5]. However, there are theoretical limitations that prevent the noise of both genes of a two-component system to be suppressed below the uncontrolled level [6]. In another word, the price of noise controlling for one gene is paid by increased noise on other cellular components. On the contrary, a positive feedback loop may amplify noise. This is particularly helpful when the positive feedback loop maintains bistability, since then large enough fluctuation can flip the expression state and drives stochastically and spontaneously switching between activation and inactivation. This amplification of noise can be beneficial to the bacterial population, as discussed in the next section.

1.1.1 Functional roles of noise

Noise in gene expression often impairs the precision of the cellular programme. However, during the past two decades, more and more genetic circuits are found to exploit noise to achieve gene expression dynamics that would otherwise be impossible for deterministic systems. One of the most important functional roles of noise is to differentiate a clonal population (i.e. genetically identical population) in a homogeneous environment [7, 8]. The nature of this differentiation is stochastic, as it is implied when an initially undistinguishable population of cells assume different fates. Stochastic fate decision is important for embryo development in multi-cellular organisms [9]. For bacteria, noise helps to create heterogeneous phenotypes within a population to maximize its chance of survival against unforeseen, fluctuating future environment. Turning on stress response is often a heavy metabolic burden to bacteria [10]. Thus, the benefit is marginal, if any, should bacteria activate its stress response when the stress is only transient. It is shown both experimentally and theoretically that switching on the stress response in only a fraction of the population ("bet-hedging") is the optimal strategy, which the switching frequency should be on par with the changing rate of the environment [11, 12]. For example, the gram-positive bacteria *Bacillus subtilis* will switch on a proportion of its population to take up environmental DNA (i.e. to achieve competence) to increase the fitness in adverse environments. *B. subtilis* switches to competence by expressing the transcription activator ComK. Suel *et al.* demonstrates that a genetic circuit consisting ComK positive autoregulation and a slower negative feedback loop is sufficient to drive the stochastic state-switching [13]. Noise is crucial in such dynamics since reducing expression noise results in a decreased proportion of competent cells [14, 13].

1.1.2 Noise in *B. subtilis* sigB circuit

Noise plays an important role in the dynamics of *B. subtilis* σ^B , the alternative sigma factor (more on Section 1.2) that triggers the general stress response. Locke *et al.* shows that upon energy stress, σ^B is activated as stochastic pulses, representing a scheme for bacteria to hedge their bets against the fluctuating environment [15]. The activation time of each pulse and the interval between pulses are both on the level of hours. Similar to the study of *B. subtilis* ComK pathway, the authors blocked septa formation to create elongated cell phenotypes, which has lower intrinsic noise due to an increased cellular volume. In elongated cells, the frequency of the pulse is reduced, which suggests that noise drives the pulsatile expression. In addition, the frequency of the pulses is positively regulated by the strength of the energy stress, which represents a new regulation paradigm (converting "amplitude modulation" to "frequency modulation"). A later work by Cabeen *et al.* shows different σ^B dynamics, i.e., upon stress response, σ^B is activated as a single, transient pulse, whose amplitude is modulated by the strength of stress [16]. However, when the upstream component (RsbR as part of the stressosome) is altered, environmental stress triggers pulse-like dynamics. A recent study shows that not only σ^B , but also other alternative sigma factors (i.e. sigma factors other than the house-keeping one, which is the primary sigma factor activated during exponential growth), including σ^M , σ^W , σ^X , σ^D , etc., adopt pulsatile expression patterns [17]. In fact, these sigma factors share the same core circuit structure as σ^B , which could account for their similar dynamics (further explained in Section 1.2.2). As a result, different alternative sigma factors take turns to occupy the RNAP cores rather than the conventional static partitioning of the pool of RNAP cores.

1.2 Bacterial alternative sigma factor circuit

1.2.1 The SigB circuit in *B. subtilis*

Sigma factors are the interchangeable components of the RNA polymerase holoenzyme (the rest of the holoenzyme is called the RNA polymerase core) which directs the holoenzyme to recognize different sets of promoters [18]. *B. subtilis* σ^B is the alternative sigma factor that triggers the general stress response by activating a regulon of more than 150 genes, which provides multi-purpose and preventive protection for the cell [19]. Since the activation of σ^B -induced regulon imposes a significant metabolic burden [10], σ^B expression is tightly controlled by a genetic circuit consisting of mixed transcriptional and post-translational regulations. σ^B -dependent general stress response is triggered

by a wide range of stimuli, including environmental stress (ethanol, salt, heat-shock or blue light, etc.) [20, 21], energy stress (starvation or ATP and/or GTP inhibitors, including mycophenolic acid (MPA) and carbonyl cyanide m-chlorophenyl hydrazone (CCCP), etc.) [19] and low temperature stress [22]. The three categories of stress induce σ^B expression through independent pathways [20, 22].

σ^B is regulated by a positive autoregulation and several positive and negative feedback loops with delay [23, 19]. σ^B activates its own expression as *sigB* gene is located in an operon induced by σ^B [24]. During exponential growth, σ^B activity is repressed by binding to the anti-sigma factor RsbW. Upon stress, σ^B is released through a partner-switching mechanism of RsbW. The anti-anti-sigma factor RsbV competes with σ^B to form an alternative complex with RsbW and thus sequesters RsbW from inhibiting σ^B [23]. Both RsbW and RsbV are co-transcribed with σ^B in the same operon driven by a σ^B -dependent promoter, which forms a negative and positive feedback [24]. On top of the regulation via protein-protein interaction is the phosphorylation of RsbV. RsbV can only associate with RsbW in the dephosphorylated state, while RsbW is also a kinase of RsbV [25]. Thus, RsbW inactivates RsbV and contribute to the negative feedback loop of σ^B expression. When RsbV is dephosphorylated, it antagonizes the anti-sigma factor RsbW and releases σ^B to turn on downstream genes. The environmental stress and energy stress uses two independent pathways to release the phosphatase, either RsbU (with co-factor RsbT) or RsbP (with co-factor RsbQ), to dephosphorylate RsbV and initiate stress response [20]. In summary, the regulatory mechanism of σ^B features mixed positive and negative feedback loops involving the anti- and anti-anti-sigma factor (Figure 1.1). This structure of σ^B circuit is conserved across several gram-positive bacteria (though the anti-anti-sigma factor is missing in some species) [19].

1.2.2 Shared core circuit structure

The sigma-anti-sigma factor circuit may represent a general regulatory mechanism for bacteria sigma factors. Besides the aforementioned existence of σ^B circuit in the gram-positive relative species, the sporulation-related sigma factors σ^F , σ^E , and σ^G in *B. subtilis* can also be inhibited by corresponding anti-sigma factors [26]. Noticeably, the anti-sigma factor of σ^F , SpoIIAB, is a kinase and can phosphorylate the anti-anti-sigma factor, SpoIIAA. SpoIIAA can only associate with SpoIIAB when dephosphorylated, which simulates the RsbV-RsbW- σ^B circuit [27]. The same structure reoccurs in *E. coli*, e.g., the proteolysis of the general stress response sigma factor RpoS (alias σ^S or σ^{38}) depends on a regulator, RssB, which serves as an anti-sigma factor and is located in a RpoS-induced regulon. In summary, the negative feedback loop of sigma factor conferred by anti-

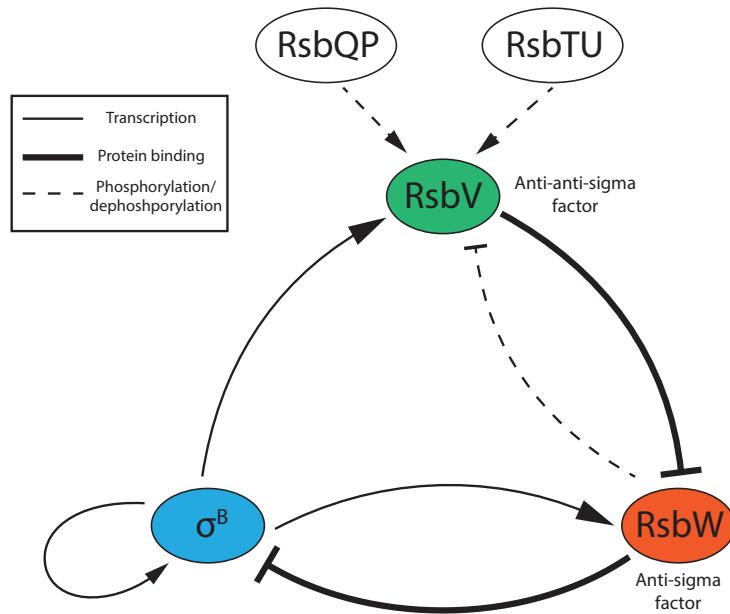


Fig. 1.1 A schematic for σ^B gene regulation circuit. Type of gene/protein interaction are represented by different line styles. Pointed-end arrow denotes activation and blunt-end arrow denotes repression.

sigma factor could be a general regulatory mechanism across different bacteria species and different sigma factors [28]. This shared structure is reflected in the mathematical model for later analysis (Section 2.1.3).

1.2.3 Sigma factor competition

Besides the strict negative regulation shared among alternative sigma factors, shifting away from the housekeeping transcription pattern is also protected against by the competition between sigma factors for limited RNAP core enzymes. Sigma factor competition has been established based on the observation that altering housekeeping sigma factor concentration affects the strength of alternative sigma factor-induced stress response, through overexpression/suppression of the housekeeping sigma factor [29] or modifying the binding affinity of the alternative sigma factor to RNAP core [30]. The alternative sigma factors (e.g. σ^B of *B. subtilis*) are at a disadvantage in the competition against the housekeeping one since (a) The alternative sigma factors have weaker binding affinity to RNAP cores than the housekeeping one, ranging from 1.5- to around 10-fold weaker [31, 32]. (b) The amount of housekeeping sigma factors exceeds significantly that of the alternative ones, even under stress [18, 33]. Also, considering that

RNAP core and housekeeping sigma factor roughly remains a constant level in different cell states [18], the pool of available RNAP cores for the alternative sigma factors is capped and the competition between them can be fierce, which is reflected a competition model that I developed during my research (Section 2.3).

1.3 Biochemical ultrasensitivity

In the context of gene regulation, ultrasensitivity describes a sigmoidal, switch-like increase or decrease of the expression of a gene along increasing concentration of the regulator [34]. Namely, the dose-response curve of gene expression against regulator concentration features a threshold, where the change of expression rate is maximized (or "ultrasensitive" in the vicinity). The most common molecular mechanism of ultrasensitivity is binding cooperativity, where multiple copies of the regulators associate and together plays a role in the regulation (e.g. binds to the promoter to initiate transcription). Conventionally, the strength of activation/repression of gene expression is captured by the mathematical formalization of the Hill equation [35, 5], where the exponent is referred to as the Hill coefficient (n). When $n = 1$, the Hill equation is also referred to as the Michaelis-Menten equation [36, 37]. In equilibrium, Hill coefficient represents the number of regulator monomers that agglomerate to form an multimer [5]. However, often due to the lumped nature of gene expression models (e.g. combined transcription and translation steps) and sources of ultrasensitivity other than binding cooperativity (e.g. multi-site phosphorylation and molecular sequestration [38]), the Hill coefficient is an apparent parameter and is appropriate for quantification of ultrasensitivity in the system [34]. Ultrasensitivity is crucial for a wide range of biochemical processes, e.g., signal transduction and the maintenance of oscillation and bistability [39, 40]. In the context of sigma factor dynamics, ultrasensitivity is needed to explain the heterogeneity of the activation of σ^V , where bistability is implied [41]. However, there is no known binding cooperativity in most sigma factor regulatory circuits, with one of the exceptions of *B. subtilis* σ^B circuit, where the anti-sigma factor RsbW dimerizes and bind to either σ^B or RsbV [25]. The RsbW homodimer can either associate one or two monomers of RsbV [42]. Contradictions between the stochastic state-switching in sigma factor dynamics and the lack of established source of ultrasensitivity is one the main themes that I attempt to address in this thesis.

Chapter 2

Methods

2.1 A general model for alternative sigma factor dynamics

Here I propose a general model to explain the spectrum of dynamical behaviours, e.g. stochastic pulsing [15, 16], transient pulsing, sustained activation [16], etc. of the bacterial alternative sigma factors in the species *B. subtilis*. It also serves as the foundation of the sigma factor competition model (Section 2.3). The model is based on the general topology that is shared by multiple alternative sigma factor networks including σ^D , σ^W , σ^X , etc [27, 43, 44], namely

- The sigma factor is under positive autoregulation, i.e., the sigma factor activates its own expression.
- An anti-sigma factor is co-expressed with the sigma factor as its inhibitor, forming a negative feedback loop.

I formalize a minimal mechanistic model based on these network topologies. Without loss of generality, the following model reflects the σ^B regulatory network but is applicable to other alternative sigma factors. First, the promoter of the operon which contains *sigB* gene is transcribed by the RNA polymerase (RNAP) holoenzyme comprising σ^B . The relative transcription rate to the maximal rate (denoted by v_H) is modelled by Hill kinetics with ultrasensitivity encoded in n :

$$v_H = \frac{\sigma_f^n}{K_S^n + \sigma_f^n} \quad (2.1)$$

where σ_f is the abundance of the unbound sigma factors (e.g. σ^B), as the anti-sigma factor (denoted by A), e.g. RsbW, competes the binding of sigma factors with RNAP cores. K_S is the dissociation constant between the RNAP carrying this particular sigma factor and the promoter. n is the Hill coefficient, which represents the apparent binding cooperativity.

Second, the anti-sigma factor (A) sequesters the sigma factor (σ) from the RNAP core by binding to it and forming a complex (denoted by $A\sigma$). As the binding is on a faster time scale than transcription, the binding dynamics can be considered at steady-state in transcriptional regulation. The binding dynamics is governed by

$$K_D = \frac{A_f \cdot \sigma_f}{A\sigma} \quad (2.2)$$

$$\sigma = \sigma_f + A\sigma \quad (2.3)$$

where σ is the total abundance of sigma factors and A_f is the amount of unbound anti-sigma factors. K_D is the dissociation constant of sigma-anti-sigma factor binding. Together, Eq. 2.2 and Eq. 2.3 bridges the abundance of free sigma factors and the total number of sigma factors as

$$\sigma_f = \frac{K_D}{K_D + A_f} \sigma \quad (2.4)$$

i.e. the binding is subject to Michaelis-Menten kinetics, which, in the general model described here, is considered as non-cooperative. Finally, under the assumption that the anti-sigma factor is in excess to the sigma factor so that $A \approx A_f$, Eq. 2.1 can be re-written as

$$v_H = \frac{(\sigma/K_S)^n}{(\sigma/K_S)^n + (A/K_D + 1)^n} \quad (2.5)$$

2.1.1 Sigma-anti-sigma cooperative binding

Given Eq. 2.5 where ultrasensitivity is derived from the cooperativity of transcription initiation, there are other steps in the core sigma factor network that can potentially contribute to n , e.g., for σ^B network, RsbW dimerizes to bind σ^B , which is captured by as $\sigma_f = K_D^2/(K_D^2 + A_f^2) \cdot \sigma$. Generalization leads to an alternative form of Eq. 2.5 where the

cooperative binding between sigma factors and anti-sigma factors solely accounts for the ultrasensitivity.

$$v_H = \frac{\sigma/K_S}{\sigma/K_S + (A/K_D)^n + 1} \quad (2.6)$$

2.1.2 Activation of the alternative sigma factor network

Here I take the σ^B circuit as an example to explain the model. Upon exposure to the stressor, a phosphatase, either RsbQP or RsbTU is released, depending on whether the environmental stress or energy stress is applied. This in turn dephosphorylates the anti-anti-sigma factor RsbV to its activated form (which the activated form is denoted by V). RsbV competes the binding of RsbW with σ^B , which releases σ^B and, thus, triggers the stress response.

Similar to Eq. 2.4, at steady-state, the abundance of RsbW free to RsbV binding is captured by Michaelis-Menten kinetics

$$A_f = \frac{K_A}{K_A + V} A \quad (2.7)$$

Notice that in Eq. 2.5, A is essentially the unbound anti-sigma factor, considering the effect of V , I have

$$v_H = \frac{(\sigma/K_S)^n}{(\sigma/K_S)^n + \left(\frac{A}{(1+V/K_A)K_D} + 1\right)^n} \quad (2.8)$$

Here, I define the apparent K_D as $K'_D = (1 + V/K_A) \cdot K_D$, which then keeps the mathematical form of Eq. 2.5 unchanged. Thus, upon stress, the step-increase in the abundance of dephosphorylated RsbV is modelled by a step-increase of the apparent K_D . For the general model, the step-increase of K_D reflects the activation of an alternative sigma factor network.

2.1.3 Modelling the stochastic fluctuation under low copy number

The alternative σ factors typically exist in low copy number and, thus, their abundance is subject to stochastic fluctuation. In *Escherichia coli*, while there can be thousands of housekeeping σ^{70} factor per cell, the number of alternative σ^E factor is only about 160

even under stress [45]. As the noise can be crucial to the pulsing dynamics [17], I model the system with the exact molecular number of the alternative σ factor (σ) and the anti-sigma factor (A) as the state variables, and the random walk through state-space with jump events (based on the unpublished work of Torkel Loman):



where the relative transcription rate, ν_H is given by Eq. 2.5. The transcription rate is generalized which also accounts for translation. ν_0 is the basal transcription activity and β is the maximal transcription rate, which determines the steady-state abundance. If steady-state is reached, the average abundance of both species equals to $\beta\nu_H$ (ignoring basal transcription), i.e. $\langle\sigma\rangle = \langle A\rangle$. The same primary transcription rate is shared between σ and A since they are co-expressed in the same operon.

If not specifically specified, I use $\beta = 50$, $\tau_A = 10$, and $\tau_S = 50$ in simulations. A previous study showed that the number of alternative sigma factors in a cell during exponential growth ranges from < 1 to over 100, depending on the type of the sigma factor. Thus, β , which represents the maximum steady-state abundance of sigma factors, is set to 50 to reflect the uncertainty. More on the molecular lifetime (τ_A and τ_S) are discussed in Section 4.1.1.

The parameters used in the general model or the derivation of which are listed in Table 2.1.

2.2 Simulation of the model

The model is simulated by the Gillespie algorithm, i.e. the Stochastic Simulation Algorithm [46], which generates statistically accurate trajectories of the jump processes described in Section 2.1.3 as a Markov process. The algorithm I used is provided by the Catalyst.jl [47] package, which an interface around DifferentialEquations.jl [48], of the Julia language [49].

Parameter	Description
K_S	Dissociation constant between the sigma factor- RNAP core complex and the promoter
K_D	Dissociation constant between the sigma factor and the anti-sigma factor
K_A	Dissociation constant between the anti-sigma factor and the anti-anti-sigma factor
τ_S	Molecular lifetime of the sigma factor
τ_A	Molecular lifetime of the anti-sigma factor
β	Maximal transcription rate of the sigma factor-activated operon
v_0	Basal transcription rate of the sigma factor-activated operon

Table 2.1 Parameter list of the sigma factor model

2.3 The sigma factor competition model

To explore the emerged dynamics from the competition of different sigma factors for limited RNAP core enzymes, I built a model consisting of two sigma factor circuits, whose biochemical parameters (n , K_S , K_D , etc.) are independent. For simplicity, the model only focuses on two of the various alternative sigma factors (for context, *B. subtilis* has 17 different alternative sigma factors [17]). The influence of the remaining sigma factors, including the housekeeping one, is averaged and reflected by the limited number of available RNAP cores. In addition to the model of the core circuit (Section 2.1.3), the competition model links the two sigma factors by the association to and dissociation from a shared pool of RNAP cores:



Where E is the RNAP core enzyme and σ represents either of the two competing sigma factors. The reaction rate follows the law of mass reaction (considering elementary reactions) with the association rate constant k_{on} and dissociation rate constant k_{off} . The rate constants relate to the equilibrium binding affinity (expressed by the dissociation constant $K_{E\sigma}$) by [37]

$$K_{E\sigma} = \frac{k_{off}}{k_{on}} \quad (2.15)$$

The total amount of RNAP cores, captured by $E + E\sigma$ in the model, is essentially the pool of RNAP cores that are accessible to the two sigma factors here subtracting the proportion already bound by the housekeeping sigma factor and other alternative ones. This value is assumed constant since both the total RNAP cores (shared by all sigma factors) and the housekeeping sigma factors remain relatively constant in different growth conditions [18]. Admittedly, the assumption is impaired by the fluctuating amount of other alternative sigma factors, which motivates a multi-sigma factor competition model in future study.

2.4 An automatic classifier for different dynamical behaviours

To effectively explore how the different behaviours generated by the model depend on the parameters, I developed this automatic classifier. The classification algorithm described here takes the discrete-valued trajectories of a two-species reaction system as the input and classifies the system as one of the dynamical regimes. First, the algorithm recreates the phase-plane paths and the vector field of time-derivatives from the trajectories. The intensity (I) of phase-plane path of each lattice point on the phase plane is calculated as the total number of times that the path passes the point. The time-derivative at the point (σ_k, A_k) is approximated by

$$\dot{\sigma} = \frac{1}{I(\sigma_k, A_k)} \sum_{t_k} \frac{\sigma(t_k + \Delta t) - \sigma(t_k - \Delta t)}{2\Delta t} \quad (2.16)$$

$$\dot{A} = \frac{1}{I(\sigma_k, A_k)} \sum_{t_k} \frac{A(t_k + \Delta t) - A(t_k - \Delta t)}{2\Delta t} \quad (2.17)$$

where t_k are the time points after exposure to stress which satisfies $\sigma(t_k) = \sigma_k$ and $A(t_k) = A_k$. Second, the algorithm identifies the stable fixed points of the system, if existing, by finding the lattice point which is both a minimum of the magnitude of the derivative vector and a maximum of the path intensity. To exclude false positives due to the trivial local extrema caused by stochastic fluctuation, the vector field is smoothed and the magnitude of the derivative vector and the path intensity at the point are tested against certain thresholds.

Finally, Figure 2.1 shows the flowchart to decide the classification, where N_{small} is the number of stable fixed points where the abundance of both species are greater than the threshold T_{fluc} , and similarly, N_{large} is the number of stable fixed points which the

abundance of both species are less than T_{fluc} . $I_{forward}$, or the forward flow, is defined as the sum of intensity on the vertical segment from $(T_{fluc}, 0)$ to (T_{fluc}, T_{fluc}) . Similarly, the reverse flow $I_{reverse}$ is the sum of intensity on the horizontal segment from $(0, T_{fluc})$ to (T_{fluc}, T_{fluc}) .

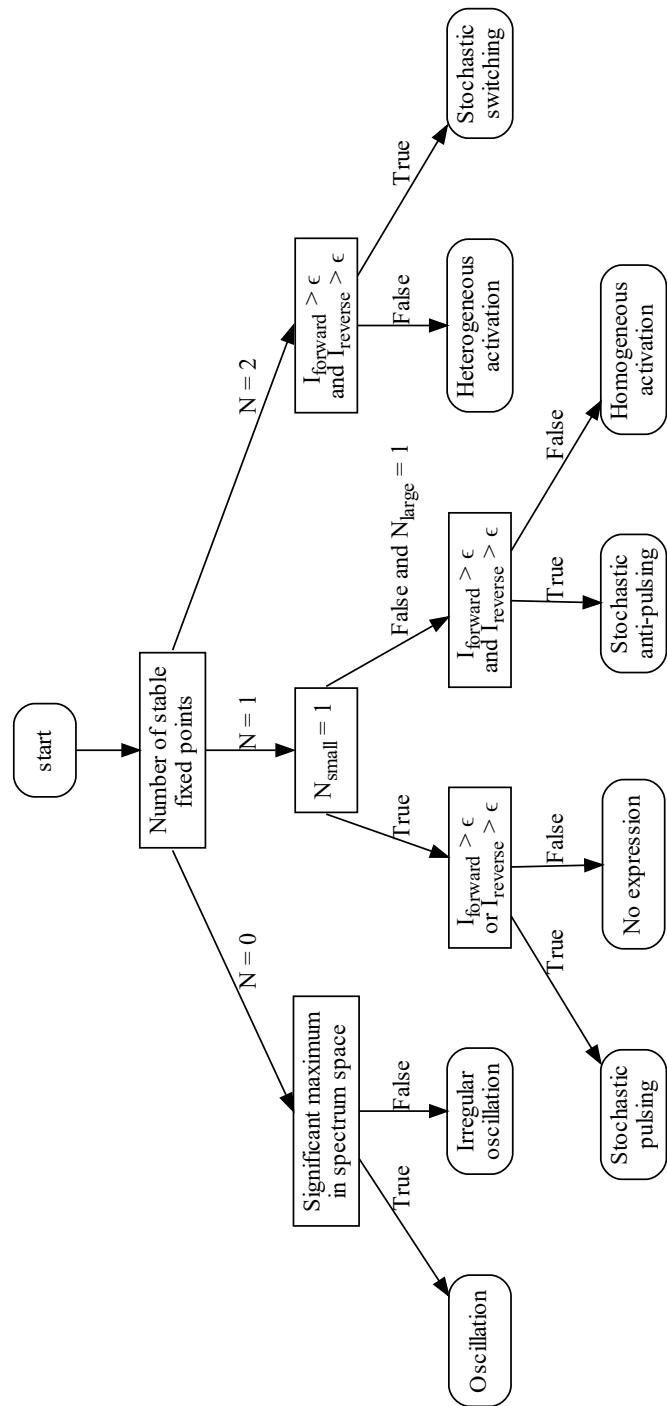


Fig. 2.1 Flowchart for the classification algorithm

Chapter 3

Results

3.1 Mathematical model predicts different dynamical behaviours

To validate the mathematical model which is based on the biochemical processes (Section 2.1.3), I simulated the time evolution of the system using the Gillespie algorithm (i.e. stochastic simulation algorithm) [46]. Previous work shows that a similar model simulated by the chemical Langevin equation (CLE) exhibits diverse dynamical behaviours (unpublished work by Torkel Loman). However, CLE only holds on the condition that, in a time step, the expected occurrence of the reaction is large enough and the change of reaction rate is small enough, both of which are facilitated by a large number of molecules present in the system [50]. Since the simulation focuses on dynamical behaviours such as stochastic pulsing and stochastic switching between activation and inactivation where low-copy number is important, these assumptions may be violated. The Gillespie algorithm generates statistically correct trajectories from the chemical master equations [46], which is important for the low-copy number regime here and also for establishing the bistability in Section 3.2.2 which exploits single-molecule dynamics.

Simulated with the Gillespie algorithm, the model (Section 2.1.3) displays a wide range of dynamical behaviours when changing the activating and repressive forces (Figure 3.1). The parameters that tune the strength of activating and repressive forces are K_S and K_D , respectively. The activating force is inversely scaled by K_S and, thus, K_S is called the activation threshold (which makes sense especially in an ultrasensitive system). Similarly, the repressive force is inversely scaled by K_D , which serves as the repression threshold (biochemical significance see Table 2.1). Seven primary types of dynam-

ical behaviours emerge from this simulation of an ultrasensitive (Hill coefficient $n = 3$) sigma factor circuit. To illustrate how the strength of activating and repressive forces determine the type of dynamical behaviours, I mapped the behaviours to the parametric space of K_S and K_D , where the different behaviours are colour-coded (Figure 3.1 centre). The seven types of dynamical behaviours are:

- (a) No expression.
- (b) Stochastic pulsing, characterized by the dominance of OFF-states interspersed with narrow peaks of expression.
- (c) Oscillation, characterized by approximately periodic switching between ON and OFF states.
- (d) Stochastic anti-pulsing, characterized by the dominance of ON-states with stochastically timed transient shutdowns.
- (e) Homogeneous activation, characterized by immediate activation upon exposure to stress.
- (f) Stochastic switching, characterized by stochastically timed switching between ON- and OFF-states.
- (g) Heterogeneous activation, characterized by heterogeneously delayed activation.

And (h) irregular fluctuation, which typically exists in non-ultrasensitive ($n = 1$) regime. It is an intermediate dynamical behaviour between stochastic pulsing and activation (more discussed in Section 3.1.1). A short trajectory along time for each behaviour is shown in Figure 3.1. The behaviour mapping in Figure 3.1 shows a wide range for stochastic pulsing (yellow) and homogeneous activation (deep blue), suggesting that these behaviours are robust to the biochemical parameters and thus are more likely to be conserved through neutral selection. Certain values of the ratio K_D/K_S separates dynamical behaviours (e.g. the vertical boundary of oscillation in Figure 3.1). This can be understood by rewriting the Hill term from Eq. 2.5 as:

$$v_H = \frac{\sigma^n}{\sigma^n + \left[A / \left(\frac{K_D}{K_S} \right) + K_S \right]^n} \quad (3.1)$$

$$\approx \frac{\sigma^n}{\sigma^n + \left[A / \left(\frac{K_D}{K_S} \right) \right]^n} \quad (3.2)$$

Where ν_H is determined by the ratio K_D/K_S and notice that ν_H is the decisive factor for the expression rate of both the sigma and anti-sigma factor. The approximation holds when it satisfies $A/(\frac{K_D}{K_S}) \gg K_S$, as is often the case since most of the dynamical behaviours happen when K_D/K_S is around 1 and K_S is small, and when the circuit is activated, the abundance of the anti-sigma factor is often dozens to hundreds. The absolute value of K_S or K_D is connected with whether the system being bistable. When K_D/K_S is approximately greater than 1, as K_S increases, the dynamics transitions from monostability (including stochastic anti-pulsing and homogeneous activation) to bistability (including stochastic switching and heterogeneous activation) (Figure 3.1). In bistable dynamics, state-flipping is triggered by stochastic fluctuations [8]. Thus, as the activation threshold K_S (or repression threshold K_D) increases, the event of noise breakthrough becomes more rare and tends to lock the circuit in one of the states, which explains the transition from monostability to bistability. In summary, the behaviour mapping (Figure 3.1 centre) can be roughly understood by independently tuning bistability/monostability by K_S and the relative activating force by K_D/K_S .

The significance of the dynamical behaviours predicted by the model is supported by experiments. Single-cell measurements of the activation of σ^B in *B. subtilis* shows stochastic pulsing [15, 16], while the heterogeneous activation delay is observed in *B. subtilis* σ^V dynamics [41].

3.1.1 Ultrasensitivity is important for bistability and oscillation

To ask what role ultrasensitivity of the circuit plays in its dynamics, I simulated the range of behaviours across the K_S - K_D/K_S parametric space with either $n = 1$ (non-ultrasensitive) or $n > 1$ (ultrasensitive). The loss of ultrasensitivity significantly changes the landscape of the behaviour mapping, with the stochastic anti-pulsing region much expanded and the oscillation and heterogeneous activation region completely lost (Figure 3.2 A). A dynamical behaviour is considered stochastic pulsing if the abundance of molecules (σ and A) is attracted to the ON-state, but large fluctuations persist (Section 2.4). Thus, the stochastic anti-pulsing area under $n = 1$ reflects a range of dynamics that the amount of sigma factors random-walk between the ON- and OFF-states without establishing oscillation or showing bistability.

To explain how some behaviours are lost in the non-ultrasensitive regime, I examined the non-ultrasensitive counterparts of the oscillation and stochastic switching dynamics (Figure 3.3). In Figure 3.3, oscillation (E) and its non-ultrasensitive counterpart (G) show similar vector fields and the existence of fixed points, so do stochastic switching (F) and its counterpart (H). However, their time-dependent trajectories are distinct

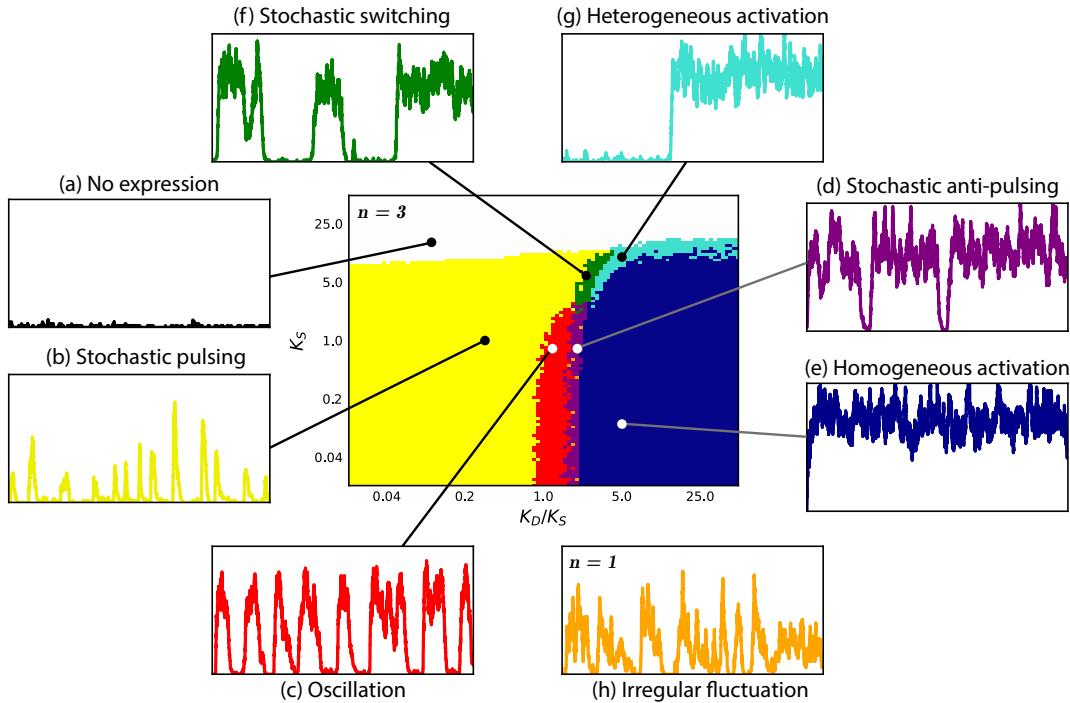


Fig. 3.1 Different dynamical behaviours generated by the mathematical model and their mapping to the parametric space. (a-g) shows the time-trajectory (x-axis: time and y-axis: the amount of sigma factors) of the types of behaviours. The trajectory starts immediately after exposure to stress. (middle) the types of behaviours mapped to a K_S - K_D/K_S parametric space, simulated under ultrasensitive Hill coefficient $n = 3$.

from each other. Under $n = 1$, oscillations collapse into irregular fluctuations, characterized by no fixed points, but not being periodic either. Non-ultrasensitivity stochastic switching shows reduced separation between the ON- and OFF-states and diminished stability of both states. The lack of ultrasensitivity may promote random walks between ON- and OFF-states, which makes oscillations and bistability difficult to maintain. In summary, the loss of ultrasensitivity in the circuit significantly changes the behaviour mapping across varying K_S and K_D and makes some behaviours unobtainable. I showed the importance of ultrasensitivity in maintaining oscillations and bistability, which is supported by observations of other genetic circuits [39, 40] (Section 1.3). However, the competition between alternative sigma factors may be an alternative source of ultrasensitivity that is not accounted for in this model of a single sigma factor circuit.

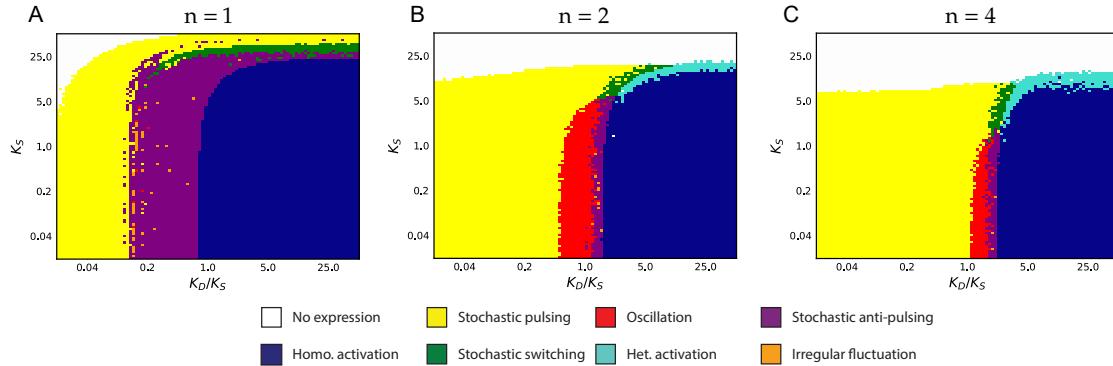


Fig. 3.2 The mapping of dynamical behaviours across the parametric space under different Hill coefficients. The mathematical model for the alternative sigma factor circuit is simulated under three different Hill coefficients ($n = 1, 2$ or 4). The landscape of the non-ultrasensitive behaviour mapping is qualitatively different from the other two. K_S : the activation threshold (inversely correlated with activating strength). K_D : the repression threshold (inversely correlated with the repressive strength).

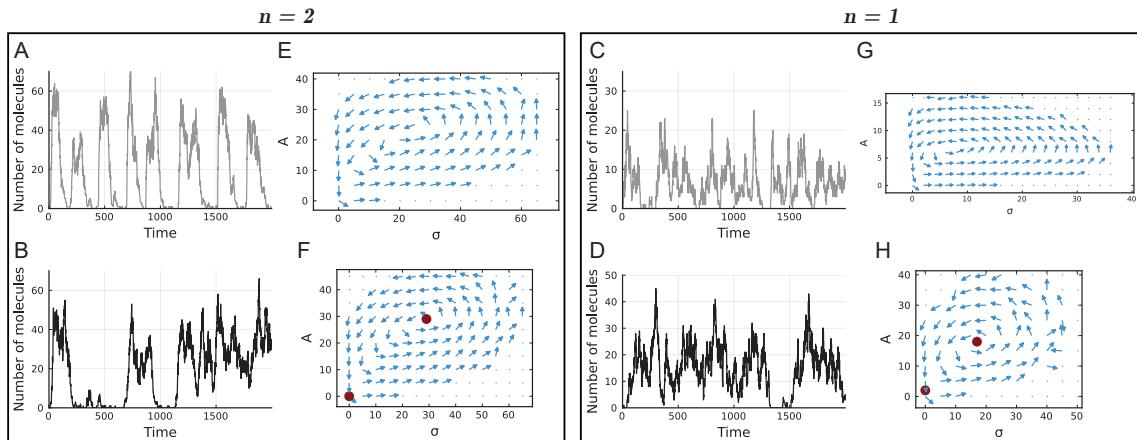


Fig. 3.3 The counterparts of oscillation and stochastic switching in non-ultrasensitive regime. (A-D) The time-dependent trajectories of oscillation (A) and stochastic switching (B) of the ultrasensitive circuit and their counterparts in the non-ultrasensitive circuit (C and D). The system is exposed to stress at time point 0. (E-H) The corresponding vector field calculated based on the trajectories. Red dots mark the fixed points. (A-B and E-F) simulation under $n = 2$, with ultrasensitivity. (C-D and G-H) simulation under $n = 1$, without ultrasensitivity.

3.2 Competition leads to bistability of non-cooperative sigma factor circuit

A previous study shows that *B. subtilis* σ^V activates heterogeneously upon stress, but there is no known cooperativity (thus, also no known source of ultrasensitivity) in the circuit [41]. I have shown that ultrasensitivity is essential for bistable dynamics, e.g., heterogeneous activation, which suggests that there could be other source of ultrasensitivity besides binding cooperativity in the σ^V circuit. Here, I propose new mechanisms (Section 3.2.1 and Section 3.2.2) for the bistability of non-cooperative alternative sigma factor circuits through the competition of several circuits for limited RNA polymerase (RNAP) core enzymes. To validate the mechanisms, I built a mathematical model describing two competitive alternative sigma factor circuits and a finite amount of RNAP cores, where the two sigma factor circuits share the same topology and the contribution from the housekeeping sigma factor is reflected by the reduced number of RNAP cores (Section 2.3).

To visualize the competition between the sigma factors, I simulated a system of two identical sigma factor circuits, in terms of their activating/repressive strength and cooperativity, etc., using the Gillespie algorithm with varying K_S and K_D . The activity of the two sigma factors (in terms of the amount of sigma factors bound to the RNAP cores) show anti-correlations in all of the dynamical behaviours, including stochastic pulsing, stochastic switching, and activations (Figure 3.4). Depending on the activation threshold (K_S), the two sigma factors can either be activated in parallel (concurrent activation) or the activation of one completely repress the other (exclusive activation). The anti-correlations are important to explain forced bistability (Section 3.2.1) and, especially the stochastic switching dynamics here, represents time sharing activation pattern of the sigma factors, by which the bacteria may generate heterogeneity in a population [41].

3.2.1 Forced bistability in asymmetric dual-sigma factor circuits

First, I examined a dual-sigma factor competition network where the Hill coefficients are asymmetric, i.e., one of the circuits is without binding cooperativity ($n = 1$, denoted as σ_1) while the other one has binding cooperativity ($n = 3$, denoted as σ_2). This model represents the actual biochemistry of, e.g., the competition between *B. subtilis* sigma factors σ^V and σ^B , since the anti-sigma factor of σ^B , RsbW, dimerizes and cooperatively binds to σ^B , while the σ^V circuit has no known binding cooperativity [42, 41]. The sim-

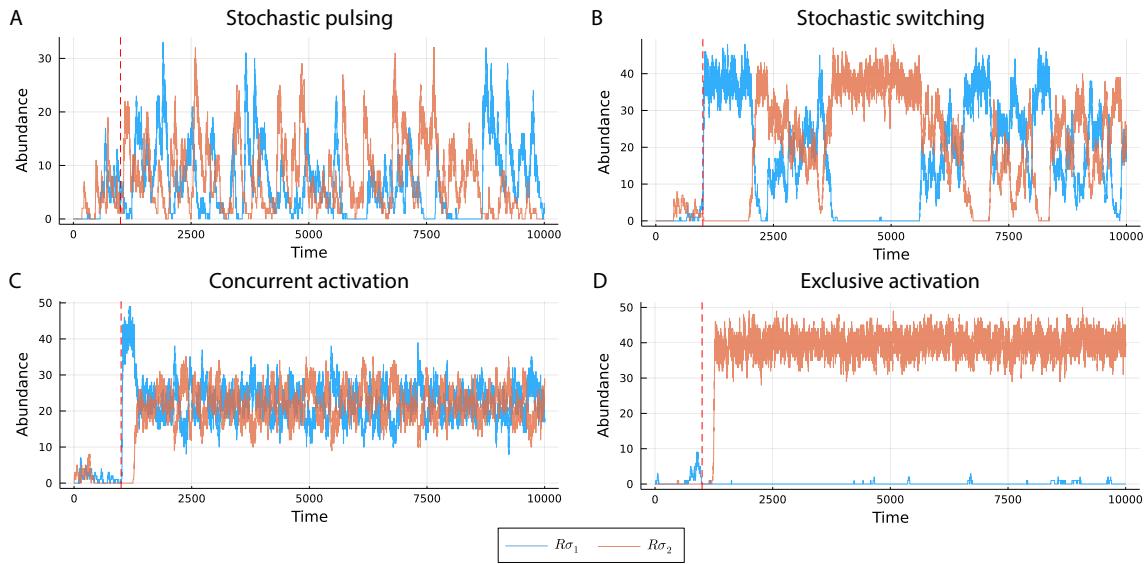


Fig. 3.4 Different behaviours of the dual-sigma factor network show anti-correlations. (A-D) The trajectories of the amount of sigma factors bound to the RNAP cores (denoted as R), which serves as the activity of the sigma factors. The total amount of RNAP cores is kept constant at 50. The two sigma factor circuits are modelled with the same parameters. The red dashed line is the time exposed to stress. Binding rate constant between RNAP core and the sigma factor $k_1 = 0.005$, the dissociation rate constant $k_2 = 0.05$.

ulation of the model shows that when the cooperative sigma factor circuit is in bistable dynamics and when competition is strong, the non-cooperative sigma factor circuit is forced to be bistable, while when in weak competition, the dynamics of the cooperative circuit does not have significant impact on the non-cooperative one (Figure 3.5 A). I suggest that the bistability of σ_2 causes the amount of free (unbound) RNAP cores to also switch between high and low states. Since steady-state expression level of σ_1 relies on the amount of free RNAP cores, it is induced to be bistable. Unlike σ_2 , the lower state of the non-cooperative σ_1 bistability is not necessarily around 0 (Figure 3.5 B-D).

To ask how the strength of the competition affects the dynamics of the non-cooperative sigma factor, I simulated the system against different values of the binding rate constant k_1 of σ_1 . The fixed points are detected as per the classification algorithm (Section 2.4) and are shown as the bifurcation-like diagram (Figure 3.5 B). The binding rate constant for the cooperative σ_2 circuit is fixed. To keep σ_2 in bistability, I first examined the region of the bistable dynamics across the $k_1 - K_S$ space (Figure 3.6). Then, the red line is chosen to assign K_S as a linear function of k_1 across different values of k_1 to keep σ_2 in the bistable regime, which is $K_S(k_1) = -0.004 \cdot k_1 + 25$. The bifurcation-like diagram shows

that as the hypothetical binding strength between the sigma factor and the RNAP core increases, the non-cooperative circuit first transits from the OFF-state to a short period of bistability, then into tristability (Figure 3.5 A middle panel and C). The tristable dynamics features two non-zero fixed points induced by the bistable σ_2 expression and a fixed point at 0 presumably due to the balance between RNAP holoenzyme formation and sigma factor degradation (further discussed in Section 3.2.2). As k_1 of the σ_1 circuit continues to increase, the zero fixed point diminishes and the system maintains bistability for a rather wide range of parameters. In summary, in a system of alternative sigma factors competing for a limited amount RNAP core enzymes, the bistable circuit can force the non-cooperative circuit to show bistability, or even tristability.

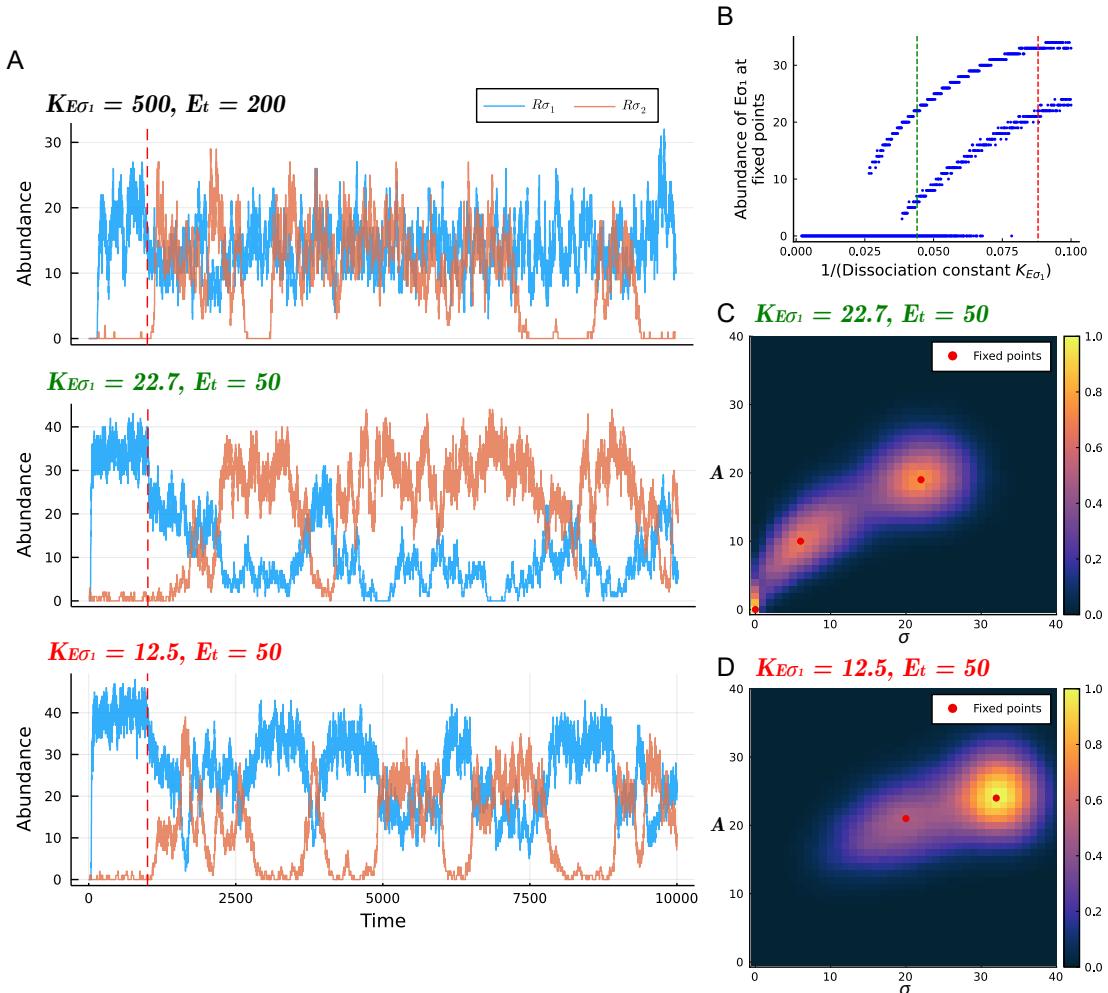


Fig. 3.5 Non-cooperative sigma factor circuit shows forced bistability due to the rival bistable circuit. (A top) Bistable σ_2 (cooperative) has no significant influence on the expression of σ_1 (non-cooperative) under weak competition (due to small binding affinity between σ_1 and RNAP cores and rather large amount of RNAP cores available). (A middle and bottom) the activity of σ_1 displays forced bistability in strong competition with bistable σ_2 . The red dashed line marks exposure to stress. $K_{E\sigma_1}$: the dissociation constant of the binding between σ_1 and the RNAP core (E). E_t : the total amount of RNAP core enzymes. (B) The bifurcation-like diagram of the RNAP core- σ_1 complex with the binding rate constant of σ_1 as the bifurcation parameter. The green and the red dashed line corresponds to the system shown in (A middle and C) and (A bottom and D). (C-D) The density of the phase paths corresponding to (A middle) or (A bottom), respectively showing tristability or bistability. $K_{E\sigma_2}$ is fixed at 10 molecules/cell.

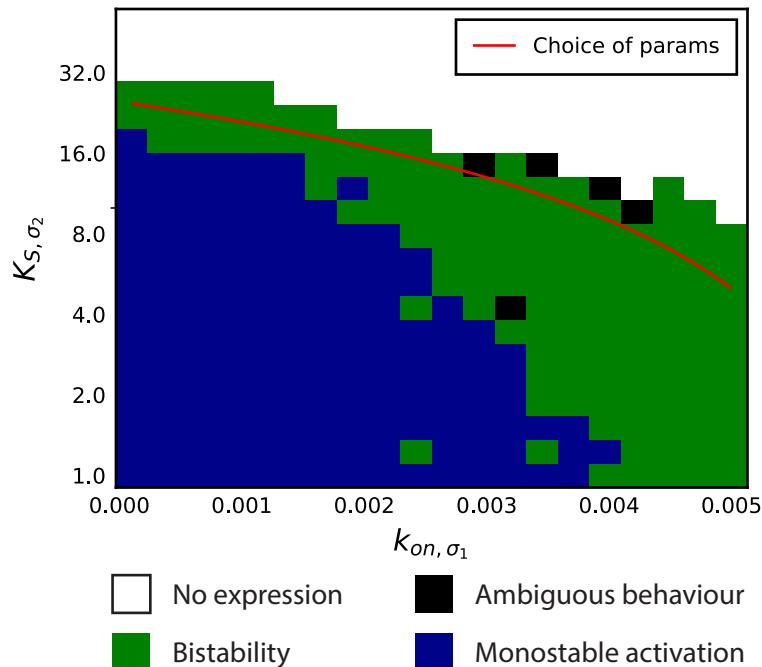


Fig. 3.6 The mapping of the behaviours of non-cooperative sigma factor circuit in the competition model to $k_1 - K_S$ space. The red line shows the function of K_S given k_1 to keep the cooperative sigma factor circuit in bistability, especially, $K_S(k_1) = -0.004 \cdot k_1 + 25$. K_1 : the binding rate constant of the non-cooperative sigma factor. K_S : the activation threshold of the cooperative circuit.

3.2.2 Non-cooperative bistability arises from low binding affinity

Here I explored the situation where none of the sigma factor circuits in the system are cooperative. Using the dual-sigma factor competition model with the Hill coefficient of both circuits set to one, I found a new bistable behaviour with an almost strict-zero OFF-state and a fluctuating ON-state (Figure 3.7 A). The bistability only holds when (a) the amount of available RNAP cores are limited, and (b) the binding affinity between sigma factors and RNAP cores is considerably low, e.g., the trajectory in Figure 3.7 A is simulated under $E_t = 50$ and $K_{E\sigma} = 80$ (units: molecules per cell). In comparison, the binding affinity for the non-ultrasensitive sigma factor in forced bistability is moderate $K_{E\sigma_1} = 10$ (for parameter specifications see Section 4.1.2). The dynamics of the rival sigma factors have no significant effect on each other (data not shown), which could be the consequence of low binding affinity. To reflect this, all simulations that contribute to Figure 3.7 are made with one of the sigma circuits turned off. The density of the occupancy of RNAP cores is clearly bimodal, with the first peak at zero (silenced) and the second one around 10%, which reflects the low binding affinity (Figure 3.7 B). As the relative strength of activating force increases, the dynamics of the circuit transitions from stochastic switching to heterogeneous activation (Figure 3.7 C). Multiple trajectories in Figure 3.7 C simulates a genetically identical bacterial population displaying heterogeneity in terms of the start of stress response. Based on the trajectories I plotted the accumulated fraction of activation in the population. I found that the variability of the start of the activation reduces against increasing binding affinity $K_{E\sigma}$ (Figure 3.7 D). Such observation also supports that the weak binding between sigma factors and RNAP cores accounts for heterogeneous activations. As the OFF-states feature strict zero-amount of $E\sigma$ complex, one hypothesis of the mechanism of bistability is that the low abundance of E and the low affinity of σ to it combined makes the production of a single holoenzyme very rare. Due to the positive autoregulation of the sigma factor, the formation of a single holoenzyme can switch on the circuit and then, it is turned off by random fluctuations. In summary, I found that in a regime of low RNAP core availability and low binding affinity of sigma factors to the cores, sigma factor dynamics exhibits bistability. This new mechanism may address the activation heterogeneity of the non-cooperative σ^V circuit in *B. subtilis*, and may represent a general mechanism for the generation of dynamics that require ultrasensitivity from a non-cooperative circuit.

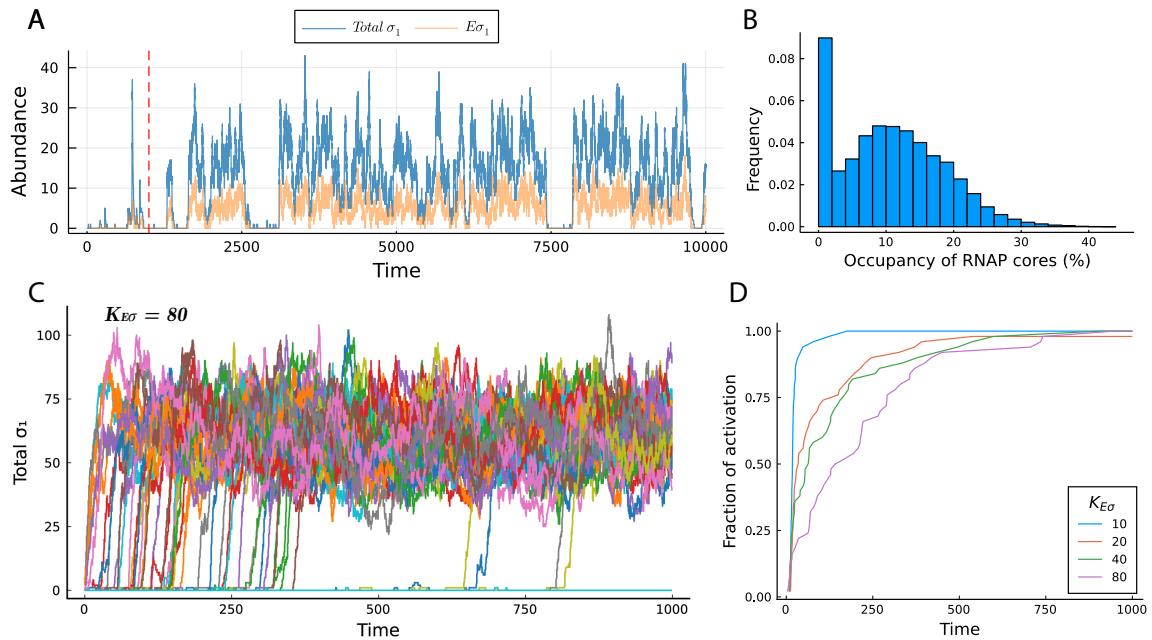


Fig. 3.7 The bistable dynamics of non-cooperative alternative sigma factor circuit under low RNAP core affinity. (A) Trajectories of total sigma factors and the bound sigma factors. The dashed red line indicates exposure to stress. In the simulation, $K_{E\sigma} = 80$ and $K_D/K_S = 0.22$. (B) Distribution of the occupancy ratio of RNAP cores by sigma factors cumulated along the time course of the simulation. The distribution is bi-modal, which centers at 0 and around 10%. (C) Trajectories of total sigma factors from 50 repeated simulations. In the simulation, $K_{E\sigma} = 80$ and K_D/K_S is increased to 0.5 to stabilize the activation state. (D) Accumulated fraction of activation across different binding affinity, where activation is defined as the time when the amount of sigma factors reaches ON-state.

Chapter 4

Discussion

4.1 Considerations on the model

The change of environment can be drastic and frequent considering small size of bacteria, which often demands the use of alternative sigma factors to adapt to the new environment through a shift in the transcription profile. It is established that bacteria leverage noise to dynamically modulate the expression of sigma factors and, thus, create phenotypic heterogeneity inside the population to gain a survival advantage [15–17, 41]. Theoretical studies show that a common network motif combining a positive autoregulation and a negative feedback loop with delay underpins these dynamics ([41] and Torkel Loman’s unpublished work). In this study, I used a simplified, mechanistic model of the same general structure to reproduce diverse dynamical behaviours of the sigma factors. I used the discrete Gillespie algorithm to correctly predict the low-molecule dynamics and properly reflect the role of noise in the circuit. I also proposed two new mechanisms to address the bistability suggested in non-cooperativity circuits [41] through competition, namely (a) A competing cooperative sigma factor circuit forces bistability via shared pool of RNAP. (b) Bistability in a non-cooperative circuit maintained by rare sigma factor-RNAP core binding event due to low binding affinity. The theoretical work here could motivate future studies to alter the sigma factor-RNAP core binding affinity (e.g. via sigma factor mutants) to discover potential new dynamical behaviours. It may also provide new insights for synthetic biologists to engineer new oscillators or bistable switches by rewiring the bacteria sigma factor circuits. Nonetheless, the model is built under considerable uncertainty of the actual biochemical processes and trade-off between generality and the amount of molecular details encoded. Here, I will elaborate some important decisions and assumptions for the model, and several future directions.

4.1.1 Arbitrary units of time

Although the model uses the absolute, discrete number of molecules per cell as the concentration for all chemical species to accurately account for stochastic fluctuations, the units of time are arbitrary. Notice that in the model (Section 2.1.3), the maximum steady-state concentration and the molecular lifetime is tuned independently through β and τ . When the circuit is activated, the expression rate factor v_H is high and can approximate to 1. Thus, the steady-state abundance of σ is roughly β . The degradation rate $1/\tau$ gives the molecule an average lifetime of τ . To support the argument, consider the (deterministic) first-order degradation reaction of molecule x with degradation rate constant $1/\tau$:

$$\frac{dx}{dt} = -\frac{1}{\tau} \cdot x \quad (4.1)$$

Given the initial concentration x_0 at $t = 0$, the evolution of x along time is expressed as

$$x = x_0 \cdot e^{-t/\tau} \quad (4.2)$$

When x_0 is large enough, the average molecular lifetime (\bar{t}) can be approximated by the area under the t - x curve divided by the total number of molecules, x_0 , namely

$$\bar{t} = \frac{1}{x_0} \int_{t=0}^{\infty} x_0 e^{-t/\tau} dt \quad (4.3)$$

$$= \tau \quad (4.4)$$

Which shows that in first-order degradation kinetics, the reciprocal of the degradation rate is the average molecular lifetime.

In the model, time is measured relative to the average lifetime of different molecules. In all simulations, the lifetime of the sigma factor is 10 time units. To reflect the delay from the anti-sigma factor, the lifetime of A is set to 50, 5-times of that of the sigma factor. In the case of competition against the RNAP cores, the dissociation rate constant k_{off} is reported to be $1 \times 10^{-3} \text{ s}^{-1}$, equivalent to the average lifetime of $E\sigma$ being 1,000 s [51]. The average lifetime of sigma factors is required to normalize k_{off} to the time units I use. To begin with, the half-life of the major pool of proteome in *E.coli* is reported to

be around 1 hour. From first-order degradation kinetics, we can derive that half-life is $\ln 2 \cdot \tau$ (proof not shown) and thus, the average lifetime is around 5,000 s [52] However, another study reports the half-life of bacteria sigma factor under protease negative regulation to be 1 minute, 60-times faster than the average proteome [53]. Given the scarcity of measurement, I approximately set the average lifetime of the holoenzyme to be the same as sigma factor (1,000 s). The predictability of the model can be improved by dedicated measurements of the half-lives of the holoenzyme and the sigma factor. Also, altering the relative lifetime of the anti-sigma factor compared to the sigma factor could reveal interesting dynamics, which contributes to future research directions. However, I argue that the accuracy of the simulation is not harmed by the use of arbitrary time units, since I inspected that the change of absolute lifetime of molecules does not change the steady-state distribution of sigma factors, given a small time step used (data not shown).

4.1.2 Derivation of model parameters

In this section, I will discuss the rationale of choosing other model parameters besides the aforementioned degradation/dissociation rates. The dissociation constant between RNAP cores and sigma factors, $K_{E\sigma}$, quantifies the binding affinity and plays an important role in tuning the strength of sigma factor competition and, thus, the existence of bistability in the system. The dissociation constant has the units of concentration [5]. Since I used molecules per cell as the units for concentration as many literature report dissociation constants in molar concentrations, I will first determine the conversion between the two. Nanomolar (nm) is the common units for the physiological concentrations of macromolecules. Multiplying the unit concentration (1 nm) by the average cellular volume of bacteria (1.32 fL) [51] and the Avogadro constant gives approximately 0.8 (molecules/cell). Maeda *et al.* reported that the dissociation constants of the alternative sigma factors in *E. coli* range from 0.3 nm to 4.26 nm, which translates to 0.24 to 3.41 molecules/cell [31]. However, a more recent study measures the binding affinity by a different method in various cellular conditions. The study displays that the dissociation constant between RpoH (alias σ^H or σ^{32}) and RNAP cores may range from 18 to 100 molecules/cell across different temperatures [32]. Though both studies consistently report the binding affinity to the housekeeping sigma factor to be higher than any of the alternative ones, the absolute values can vary 10- to 100-fold, depending on different methods and measuring conditions. I chose $K_{E\sigma}$ between 10 to 100 molecules/cell for most simulations, in comply with the newer measurements. Notice that both studies measure the dissociation constants *in vitro*. It would be interesting to see *in vivo* mea-

surements, which may solve the conflicts and improve model accuracy. Also, since the $E\text{-}\sigma$ dissociation constant is shown to change according to ionic strength and temperature. Future study may explore whether the change in binding affinities contribute to sigma factor competition.

The concentration of RNAP core enzyme in bacteria is also an unsettled case, with estimates ranging from approximately 1,500 to 13,000 copies per cell [33, 54]. To ask the proportion of RNAP cores associated with the alternative sigma factors, which is the number used in the model to reflect housekeeping sigma factor competition, I designed an equilibrium binding model. The total number of RNAP enzyme cores is set to $E_t = 2,000$. Since the amount of housekeeping sigma factors (denoted by σ^A) is reported to significantly exceed at approximately 3-fold, I set $\sigma_t^A = 6,000$. Finally, I chose the number of all alternative sigma factors to be $\sigma_t^B = 2,000$ to reflect the observation that the alternative sigma factors are fewer than the housekeeping ones. From the above discussion of binding affinities, I set $K_{E\sigma^A} = 1$ and $K_{E\sigma^B} = 10$ to show the weaker binding affinities of the alternative sigma factors. The model is given by:

$$K_{E\sigma^A} = \frac{E \cdot \sigma^A}{E\sigma^A} \quad (4.5)$$

$$K_{E\sigma^B} = \frac{E \cdot \sigma^B}{E\sigma^B} \quad (4.6)$$

$$E_t = E + E\sigma^A + E\sigma^B \quad (4.7)$$

$$\sigma_t^A = \sigma^A + E\sigma^A \quad (4.8)$$

$$\sigma_t^B = \sigma^B + E\sigma^B \quad (4.9)$$

Where E represents only the free RNAP cores, while $E\sigma^A$ are the ones bound to the housekeeping sigma factors, and E_t is the total amount. The sigma factors adopt the same rules of notations. Solving the equations gives almost saturated RNAP core binding, with $E\sigma^A \approx 1,910$ and $E\sigma^B \approx 89$ (molecules/cell). In the core model for the sigma factor circuit, I used a smaller $E = 50$ (RNAP cores available to alternative sigma factors) in most settings to simulate a fierce competition scenario. Altering the amount RNAP cores available to the alternative sigma factors is shown to weaken the competition (Section 3.5). Future study may further quantify the effects and explore whether bistability can sustain with changing total number of RNAP cores.

4.2 Future work

By simulating a simplified two-component mathematical model of the alternative sigma factor circuit, I produced a range of behaviours, some of which have been experimentally observed. From this model, I examined the scenario where two alternative sigma factor circuits compete for the limited amount of RNAP cores. From there I discovered two new bistable patterns of the sigma factor circuit. First, when the binding between sigma factors and RNAP cores is strong, and if a ultrasensitive circuit exists in the cell, another non-ultrasensitive circuit can be induced to be bistable. The bistability could be transferred through the shared, limited pool of RNAP cores. Second, when the binding affinity is weak (and if the amount of RNAP cores are limited due other sigma factors, especially the housekeeping one), the event of holoenzyme formation may be rare. Then, the system would be turned on by the rare binding event and be turned off due to random walk to zero from a relative small number of holoenzymes.

Considering that the research is theoretical, it would be strengthened by evidence from experiments. Furthermore, several experimental observations may provide the model with new insights and in turn refine the theory. To begin with, there are still a few interesting directions that one may explore under the current modelling framework. First, since the Gillespie algorithm generates correct low-copy number dynamics, it would be interesting to ask what is the effect of absolute number of molecules in the system to different dynamical behaviours. Experiments have shown that when noise is reduced due to an increase in the absolute number of molecules, some dynamical behaviours, e.g., stochastic switching and stochastic pulsing, can be weakened [13, 15]. It would be important to validate and quantify this effect in theory. Second, I observed that, similar to the effects of ultrasensitivity, tuning the molecular lifetime of the anti-sigma factor to be less than that of the sigma factor eliminates some behaviours, including oscillation (data not shown). Intuitively, since the lifetime of anti-sigma factors determines the time delay of the negative feedback in the model and the lifetime of sigma factors determine the speed of positive autoregulation, establishing the oscillation may require a relative slower negative feedback. It will be interesting to confirm this from simulation or theory.

Several experiments may strengthen the foundation of the model, including measuring the average lifetime of different sigma factors (considering both protein degradation and dilution effects) and the dissociation rate of the holoenzyme. A more important aspect is to measure the sigma factor-RNAP core binding affinity (in the form of dissociation constant $K_{E\sigma}$) *in vivo*, since this binding affinity has major implications to sigma factor behaviours under competition. Moreover, there is still huge uncertainty in

its measurement, which is complicated by the dependence of $K_{E\sigma}$ on different cellular conditions (Section 4.1.2) [32]. Future research may pivot on whether stress influences the binding affinity and, thus, alters sigma factor dynamics. The behaviour mapping in Figure 3.1 predicts several dynamical behaviours of sigma factors that have not yet been experimentally observed, e.g. oscillation, stochastic anti-pulsing, etc. It would be possible to observe these behaviours using single-cell techniques, by either constructing sigma factor or anti-sigma factor mutants, or simply by tuning the strength of stress. The model also suggests that the strength of competition can modify the bistable dynamics of non-ultrasensitive sigma factors. Relevant experiments can not only serve as a confirmation of the proposed molecular mechanisms, but also help to discover new ways to tune the phenotypic variability of a bacterial population. For example, the strategy which *B. subtilis* tunes the variability of σ^V by the strength of stress is qualitatively reproduced by my model [41]. Finally, the range of dynamical behaviours shown in sigma factors may contribute to the synthetic biology toolbox. Orthogonal sigma factor circuits from other species can be leveraged as a frequency-tunable stochastic pulsing generator or as a source of heterogeneity in a population.

Chapter 5

Conclusion

Noise unavoidably arises from the thermal motion of molecules and from the fluctuating environment where the cells reside. However, more and more evidence has shown that noise may play functional role in biology [7, 8]. Bacteria can utilize noise to activate a specific set of genes in a fraction of the population to create phenotypic heterogeneity. In this way, bacteria may "hedge their bets" to make some individuals survive a sudden change in the environment or benefit the individuals that do not switch on stress response (thus less metabolic burden) if the favourable environment recovers (Section 1.1.1). In bacteria, the activation of a certain group of genes is typically directed by the expression of alternative sigma factors. Previous studies have shown that sigma factors may exhibit dynamical behaviours upon stress, e.g. stochastic pulsing (σ^B of *B. subtilis*) and heterogeneous start of activation (σ^V of *B. subtilis*) [15, 41]. In this study, I simulated a simplified mechanistic model with the Gillespie algorithm and produced a wide range of behaviours, e.g., stochastic pulsing, stochastic state-switching, stochastic anti-pulsing (i.e. inactivation pulsing), oscillation, etc (Section 3.1). The model features mixed positive autoregulation and negative feedback loop ([41] and Torkel Loman's unpublished work) (Section 2.1.3). Biochemical ultrasensitivity has been shown to facilitate some dynamical behaviours, including oscillation and stochastic state-switching [39]. However, the role of ultrasensitivity in the alternative sigma factor circuit remains unclear. By simulating a non-ultrasensitive circuit ($n = 1$), I observed that not only is oscillation lost, but also other behaviours, e.g. stochastic switching and stochastic anti-pulsing, are significantly weakened by increased fluctuations (Section 3.1.1). Though ultrasensitivity can be important for bistability, many sigma factor circuits show no known source of ultrasensitivity from binding cooperativity. This suggests that there can be other sources of ultrasensitivity. It has also been shown that sigma factors are in competition with each other for the limited amount of RNAP cores (Section 1.2.3). In

light of that, I simulated a dual sigma factor model to reflect the competition between them (Section 2.3). Based on the model, I proposed two new mechanisms for bistability to emerge from a non-ultrasensitivity circuit. First, when competition is strong, one ultrasensitive, bistable sigma factor circuit can force another non-ultrasensitive circuit to show bistability through the limitation of a shared pool of RNAP cores (Section 3.2.1). Second, when the binding affinity between the non-ultrasensitive sigma factor and the RNAP core is low, the abundance of sigma factors is trapped in a strict-zero state. The circuit is then turned on by rare binding events and turned off by the random walk from a low copy number to zero (Section 3.2.2). In the second scheme, I also showed that the activation variability caused by bistability increases as the binding affinity decreases, which qualitatively reflects the observations of σ^V in *B. subtilis* [41]. As the core structure used in the model is shared among various bacterial alternative sigma factor circuits, these mechanisms may represent a general strategy to generate heterogeneity in a population. In future studies, it would be interesting to experimentally explore other dynamical behaviours predicted by the model. It would also be important to validate the proposed mechanisms of non-ultrasensitive bistability by experiments. Finally, considering the wide range of dynamics produced from the rather simple sigma-anti-sigma circuit, engineering orthogonal sigma factor circuits may enrich the toolbox for synthetic biology.