

Engineering sequestration-based biomolecular classifiers with shared resources

Hossein Moghimianavval,^{†,‡,††} Ignacio Gispert,^{¶,‡,††} Santiago R. Castillo,^{§,‡} Olaf B. W. H. Corning,^{||,‡} Allen P. Liu,^{⊥, #, @, Δ} and Christian Cuba Samaniego^{*, ∇, ‡}

[†]*Department of Mechanical Engineering, University of Michigan, Ann Arbor, MI, 48109, USA*

[‡]*CSHL Course in Synthetic Biology 2022, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 11724, United States*

[¶]*Chemical Engineering Department, Imperial College London, London, SW7 2AZ, U.K.*

[§]*Department of Biochemistry and Molecular Biology, Mayo Clinic, Rochester, MN, 55905, USA*

^{||}*Department of Bioengineering, University of Washington, Seattle, WA, 98125, USA*

[⊥]*Department of Mechanical Engineering, University of Michigan, Ann Arbor, MI, 48109, USA*

[#]*Department of Biomedical Engineering, University of Michigan, Ann Arbor, MI, 48109, USA*

[@]*Department of Biophysics, University of Michigan, Ann Arbor, MI, 48109, USA*

^Δ*Cellular and Molecular Biology Program, University of Michigan, Ann Arbor, MI, 48109, USA*

[∇]*Computational Biology Department, Carnegie Mellon University, Pittsburgh, PA 15213, USA.*

^{††}*Equal contribution*

E-mail: ccubasam@andrew.cmu.edu

Abstract

Constructing molecular classifiers that enable cells to recognize linear and non-linear input patterns would expand the biocomputational capabilities of engineered cells, thereby unlocking their potential in diagnostics and therapeutic applications. While several biomolecular classifier schemes have been designed, the effect of biological constraints such as resource limitation and competitive binding on the function of those

classifiers has been left unexplored. Here, we first demonstrate the design of a sigma factor-based perceptron as a molecular classifier working based on the principles of molecular sequestration between the sigma factor and its anti-sigma molecule. We then investigate how the output of the biomolecular perceptron, *i.e.*, its response pattern or decision boundary, is affected by the competitive binding of sigma factors to a pool of shared and limited resources of core RNA polymerase. Finally, we reveal the influence of sharing limited resources on multi-layer perceptron neural networks and outline design principles that enable the construction of non-linear classifiers using sigma-based biomolecular neural networks in the presence of competitive resource-sharing effects.

Keywords

Synthetic biology, Biomolecular neural networks, Molecular sequestration, Competitive binding, Molecular resource sharing.

1 Introduction

Cellular biocomputation is prevalent in nature with examples including activation of genetic circuits during cell proliferation, decision-making in immune response, and a myriad of phosphorylation-based signaling pathways for determining correct response to exogenous signals (1–3). The foundation of such computational processes is often laid on molecular interactions such as protein dimerization or ligand-receptor binding. Thus, the inputs of the computational modules in biological systems are typically the concentration of certain monomeric molecules or ligands and, similarly, the outputs are the concentration of specific dimeric or multimeric molecules.

Drawing inspiration from natural systems, synthetic biologists have been striving to engineer biocomputational schemes in top-down as well as bottom-up synthetic biological systems. While the majority of biocomputational designs rely on utilizing genetic circuits

to engineer basic logic gates and simple computational tasks (4–6), a few studies have demonstrated engineering protein-based circuits that utilize proteolytic or phosphorylation reactions to generate an output (7–9). Although such biocomputational modules enable simple tasks such as biosensing of chemical species and basic computation, they typically generate digital (0 or 1 or "on or off") responses. Furthermore, encoding more sophisticated processing using logic gates demands intricate architectures with many logic gates and computational layers, rendering them convoluted for practical applications for complex tasks.

Therefore, constructing simple signal processing units inside living systems that can perform intricate computational tasks such as classification and decision-making are of great interest (**Fig. 1A**, left). Implementing molecular classifiers in living cells would enable the creation of ultra-sensitive biosensors, programming accurate cellular responses through molecular circuits, and enhanced discrimination of inputs through combinatorial sensing (10). For example, a simple linear classifier (**Fig. 1A**, middle) equips a cell with a signal processing system that ideally allows output generation only in certain input regimes (where x_1 and x_2 approach 1 in the example). Further, combining different molecular classifiers results in more complex, non-linear computation, thus expanding the capabilities of cellular biocomputation.

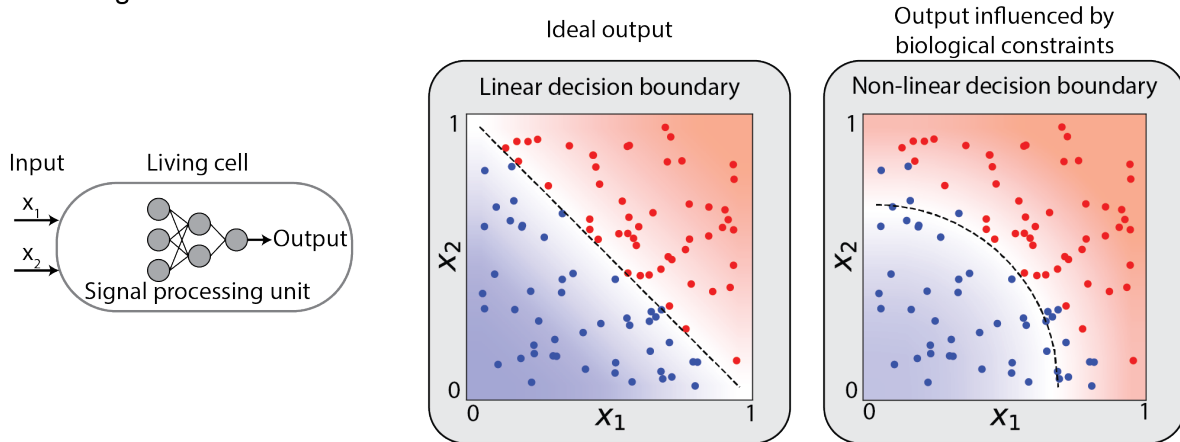
Binary classification using a linear decision boundary was demonstrated in the field of artificial intelligence (AI) in 1958 (11). A simple computational unit called 'perceptron' performs binary classification by computing the linear combination of weighted inputs and passing the summed weighted inputs through an activation function. The most popular activation function in modern perceptrons is a thresholding function called Rectified Linear Unit known as ReLU, in which the output is larger than zero when the input crosses a threshold (**Fig. 1B**). The collective processing of inputs by many layers of multiple perceptrons, known as deep neural networks or artificial neural networks (ANNs), can result in the recapitulation of any continuous function (12), thus making ANNs capable of performing

complicated tasks such as non-linear classification and accurate prediction (13–16).

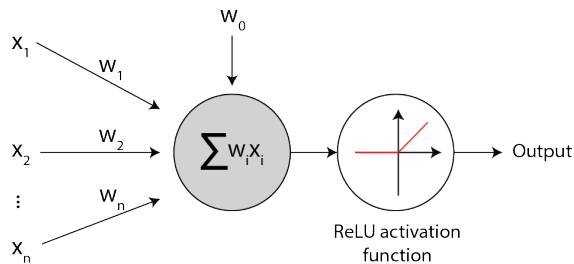
The simple architecture of a perceptron has motivated many efforts towards the creation of a biological perceptron as the signal processing unit for linear classification.(10) The construction of a single biological perceptron can also pave the way for implementing non-linear input classification in living cells by utilizing multiple perceptrons, thus creating biomolecular neural networks (BNNs). A biological perceptron must demonstrate the principal characteristics of an ANN (or computer-based) perceptron, *i.e.*, the biological perceptron must include controllable molecular elements that determine the perceptron’s input weights and activation function.

However, as opposed to ANN perceptrons, biological perceptrons face challenges in linear classification due to biological constraints such as limited resources, competitive binding, and non-specific binding between molecules. Resource constraints in biological systems have been shown to significantly impact the function of biocomputational modules based on molecular interactions. (17–21) For example, the competition between housekeeping sigma factor RpoD with minor sigma factors such as RpoN, RpoH, and RpoF in *Escherichia coli* (*E. coli*) determines the response of *E. coli* to conditions like nitrogen-deficiency, heat shock, or the need for chemotaxis, respectively, by regulating expression of genes important for growth and survival of bacteria.(22) Additionally, competitive binding of only a few promiscuous ligands to various receptors has been demonstrated to accommodate a wide range of signaling activities in multicellular organisms (23). Similarly, it was recently shown that competitive protein dimerization allows small networks made of monomeric proteins to encode an extensive range of homo- or hetero-dimeric outputs through precise adjustments in the concentration of network monomers (24). Likewise, competition between transcription factors that bind to the same RNA polymerase may affect the output of engineered genetic circuits in bacteria. (18, 19, 25).These resource constraints can cause perturbations to the biological perceptron function, thus influencing its decision boundary (**Fig. 1A**, right). Subsequently, BNNs made from the biological perceptron with an altered output will also generate decision

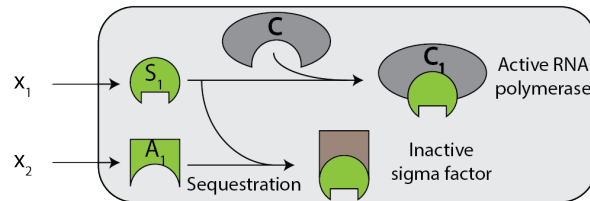
A Biological classifier



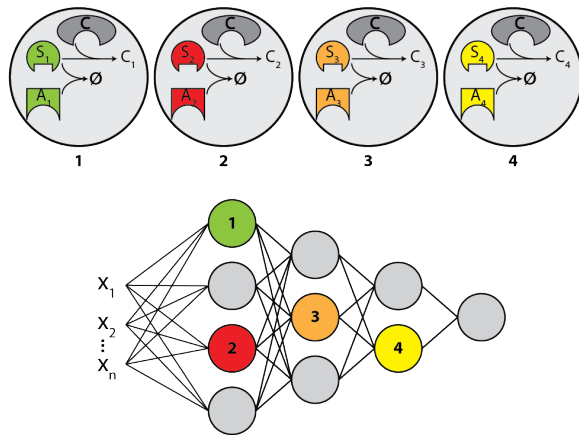
B Single perceptron



C Sigma-based perceptron



D Multi-layer perceptron neural network



E Complexities of multi-layer sigma-based neural networks

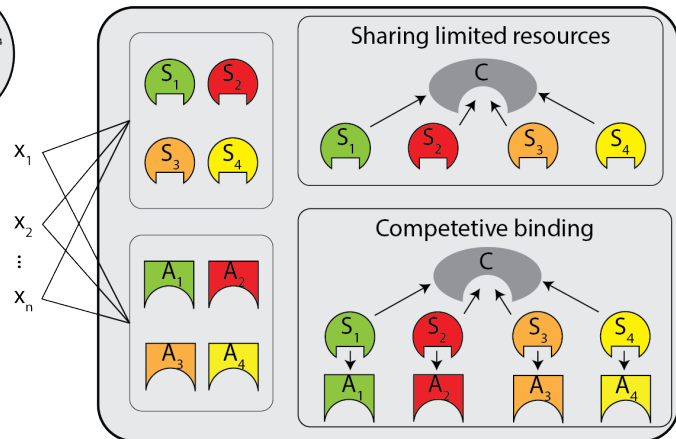


Figure 1: Designing a Biomolecular Neural Network (BNN) utilizing a sigma-based perceptron and multi-layer neural networks with shared limited resources for molecular classification: *A: A molecular classifier as a signal processing unit in living cells enables more sophisticated biocomputation. The resource constraints in biological systems, however, may perturb the designed decision boundary. B: The abstract representation of a perceptron, the building block of deep neural networks. C: The representation of the biological perceptron design in a sigma-based system using sequestration of a sigma factor and its corresponding anti-sigma protein. The inputs of the perceptron are concentrations of the sigma and anti-sigma. D: Abstract depiction of multi-layer neural networks made of multiple perceptrons in each layer.*

E: Schematic design of a multi-layer perceptron in a sigma-based system that poses two major limitations: sharing limited resources and competitive binding. These complexities influence the decision boundary of the multi-layer neural network.

boundaries that may not closely follow their ideal design.

A few biological perceptrons have been demonstrated by using inducible gene expression networks (26–29), enzymatic processing of different metabolites(30), principles of DNA strand-displacement (31–34), and DNA-processing enzymes (35). Relying on the sequestration of two interacting proteins, a biomolecular classifier with tunable positive and negative weights was designed computationally (36) and tested experimentally to achieve non-linear classification in mammalian cells (29). Similarly, a phosphorylation-based neural network with positive and negative weights that perform non-linear classification (*i.e.*, recapitulating XNOR and XOR) was designed by Cuba Samaniego *et al.* (37) Recently, a protein-based neural network that achieves linear classification was implemented by exploiting coiled-coil dimerization of engineered peptides (38). Although these studies utilize different approaches to create a biological perceptron, they all rely on competitive interactions between input-processing molecules with a shared pool of limited resources (*e.g.*, RNA polymerases, ribosomes,(21, 25, 39), co-activators, transcription factors,(40) and enzymes like Cas proteins(41) or proteases(42)). However, the effects of resource constraints on the function of these perceptrons have remained unexplored.

Here, we develop a mathematical model to simulate a biological perceptron based on sigma factors —transcription factors which outnumber the RNA polymerase and hence compete for binding to it (18, 22, 43–45)— that bacteria naturally use to regulate gene expression. Leveraging competitive dimerization between a sigma and either its corresponding anti-sigma molecule or RNA polymerase, we demonstrate the design of a simple perceptron with a non-linear activation function capable of realizing positive and negative weights (**Fig. 1C**). We then impose two physiological requirements on our model to account for both competition between the input sigma factor and other present sigma factors as well

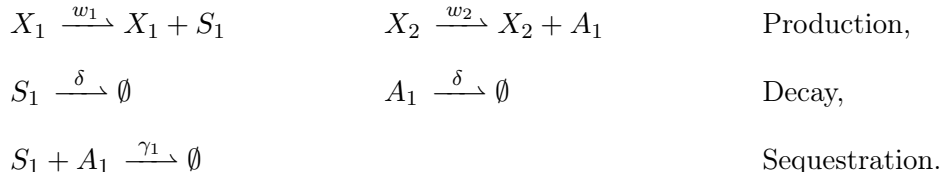
as the limited available resources (*i.e.*, core RNA polymerase (18, 44, 45)). We show that resource sharing reveals its effect on the function of the perceptron by suppressing the output while introducing a slight perturbation to the linear decision boundary. Lastly, since engineering non-linear decision boundaries require multi-layer perceptron networks (as depicted in **Fig. 1D**), we explore designing sequestration-reliant multi-layer sigma-based perceptron networks in the presence of perturbations caused by sharing limited resources (**Fig. 1E**). We demonstrate that resource sharing leads to deviations from ideal design that affect the output of the multi-layer perceptron network. However, despite the non-ideal function of perceptrons due to resource sharing and limited resources, we outline simple design principles for encoding non-linear response patterns such that they closely resemble their ideal design. Our analyses of biological perceptron and BNN function in the presence of resource constraints can also be utilized to model molecular classifiers *in silico* for more precise prediction or design of outputs of these biocomputational systems.

2 Results

Throughout the manuscript, we indicate chemical species with capital letters (*e.g.*, X) and their concentration with the corresponding lowercase letters (*e.g.*, x). Where possible, reaction and network outputs are normalized throughout the manuscript to allow dimension-less comparison between different conditions. Table 1 at the end of the Results section summarizes all parameters used in the manuscript for mathematical analyses and computational simulations.

2.1 Design of a rectified linear activation unit (ReLU) based on sigma factor-anti-sigma factor interaction in the presence of shared limited resources

Molecular sequestration is the stoichiometric binding between two species that results in the formation of a dimeric complex. An example of molecular sequestration is the interaction between sigma factors and their corresponding anti-sigma proteins that leads to the formation of a complex that is unable to promote gene expression (**Fig. 1B**). Such interaction can be modeled assuming the sigma factor S_1 and anti-sigma factor A_1 are produced from species X_1 and X_2 at rate constants w_1 and w_2 , respectively. Additionally, the produced proteins S_1 and A_1 degrade at rate δ and the sequestration occurs with rate constant γ_1 (as shown in **Fig. 2A**). We summarize the list of chemical reactions as follows:



Note that for simplicity, we assume binding reactions are irreversible. This assumption does not affect the steady-state output of the reactions although it cannot be used to analyze the dynamic behavior of the system (see SI section 3 for derivation of the steady-state output of the system assuming reversible reactions). Depending on the sequestration rate (γ_1), the output of the system, \bar{s}_1 , which stands for the steady-state amount of S_1 , follows the input ($w_1 x_1$) in different patterns (**Fig. 2B**). To evaluate two regimes of output at the steady state, we define a dimensionless positive parameter $\xi = \frac{\delta^2}{w_1 x_1 \gamma_1}$. In the fast sequestration regime (where the binding affinity of S_1 and A_1 is large) when $0 < \xi \ll 1$, modeling the interaction between S_1 and A_1 (see section 1.1 in SI for mathematical derivation) leads to a thresholding function (\mathcal{B}) (a function that generates positive outputs only when the input is larger than a threshold) between the output s_1 and input x_1 shown in equation (1) and

depicted in **Fig. 2B**.

$$\lim_{\xi \rightarrow 0} \bar{s}_1 = \frac{1}{\delta} \max(0, w_1 x_1 - w_2 x_2) \quad (1)$$

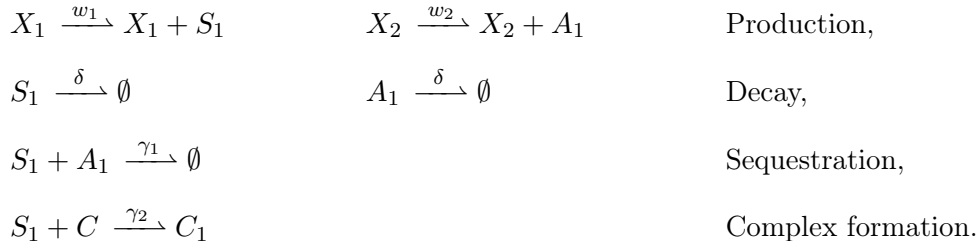
Equation (1) describes a non-linear relationship between \bar{s}_1 and x_1 where \bar{s}_1 is non-zero and proportional to x_1 only when x_1 is larger than a threshold. On the other hand, when $\xi \gg 1$, due to the slow kinetics of $S_1 - A_1$ binding, the relationship between \bar{s}_1 and x_1 becomes non-linear for the whole range of x_1 and the thresholding behavior is lost (**Fig. 2B**).

The assumption of fast sequestration of a sigma factor by its corresponding anti-sigma molecule is valid as previous studies on various anti-sigma molecules have found their rapid effect on transcriptional activity of their target sigma factors (46–48). In addition, molecular controllers based on fast sigma-anti-sigma interaction have been constructed and tested (49–51), thereby providing evidence for fast sequestration assumption.

In equation (1), x_1 and x_2 represent the concentration of input species that generate S_1 and A_1 with production rates of w_1 and w_2 , respectively. At steady state and a fast sequestration regime, equation (1) converges asymptotically to a ReLU function. Therefore, we named the relationship between \bar{s}_1 with the inputs x_1 and x_2 the Asymptotic ReLU (AReLU) function (**Fig. 2C**). Thus, the sequestration relationship between S_1 and A_1 resembles a simple perceptron with an AReLU activation function and weights of w_1 and $-w_2$ for inputs x_1 and x_2 , respectively. The quasi-linear relationship regime between \bar{s}_1 and x_1 (that resembles a soft ReLU function(52)) indicates that the steady-state available amount of S_1 is simply the difference between the total steady-state amounts of S_1 and A_1 , similar to other sequestration-based calculators demonstrated previously (53). In other words, the outcome of the sequestration chemical reaction simply calculates the subtraction of total S_1 and A_1 and does not produce any product if S_1 is lower than A_1 . However, in slow sequestration, the amount of \bar{s}_1 depends on the input x_1 through a non-linear relationship without displaying

thresholding behavior (**Fig. 2B**).

In natural systems, sigma factors either bind to their corresponding anti-sigma or the RNA polymerase (RNAPol) core. In the latter case, the RNAPol-sigma factor complex binds to a specific sigma promoter in the DNA sequence and drives the expression of the downstream genes. When free S_1 is present, it binds to the RNAPol to initiate transcription. Therefore, when designing sigma-based neural networks, since the amount of RNAPol-sigma factor complex directly influences protein expression, the interaction between S_1 and the RNAPol core should be considered. Assuming that the sigma factor S_1 binds to a limited amount of available RNAPol core C , with total concentration denoted as c^{tot} , at rate γ_2 , the total amount of RNAPol-sigma factor complex, denoted as C_1 , can be calculated by solving the system of ordinary differential equations (ODEs) representing the following chemical reactions which are illustrated in **Fig. 2D**:



To analyze the input-output relationship of the chemical reactions that consider limited resources, we introduce a dimensionless variable $r = \frac{\gamma_2}{\gamma_1}$ referred to as competitive binding ratio that represents the ratio of sigma factor-RNAPol affinity to the sequestration rate. Solving the ODEs representing the above reactions (see section 1.2 in SI for derivation) reveals that the relationship between steady-state C_1 denoted as \bar{c}_1 (and consequently its normalized value denoted as $\bar{c}_1^n = \bar{c}_1/c^{tot}$) and x_1 depends on the competitive binding ratio. When $r \gg 1$ (slow competitive binding regime) the input-output relationship deviates from the ideal behavior (thresholding function) observed in **Fig. 2B**. However, when $r \ll 1$ (fast

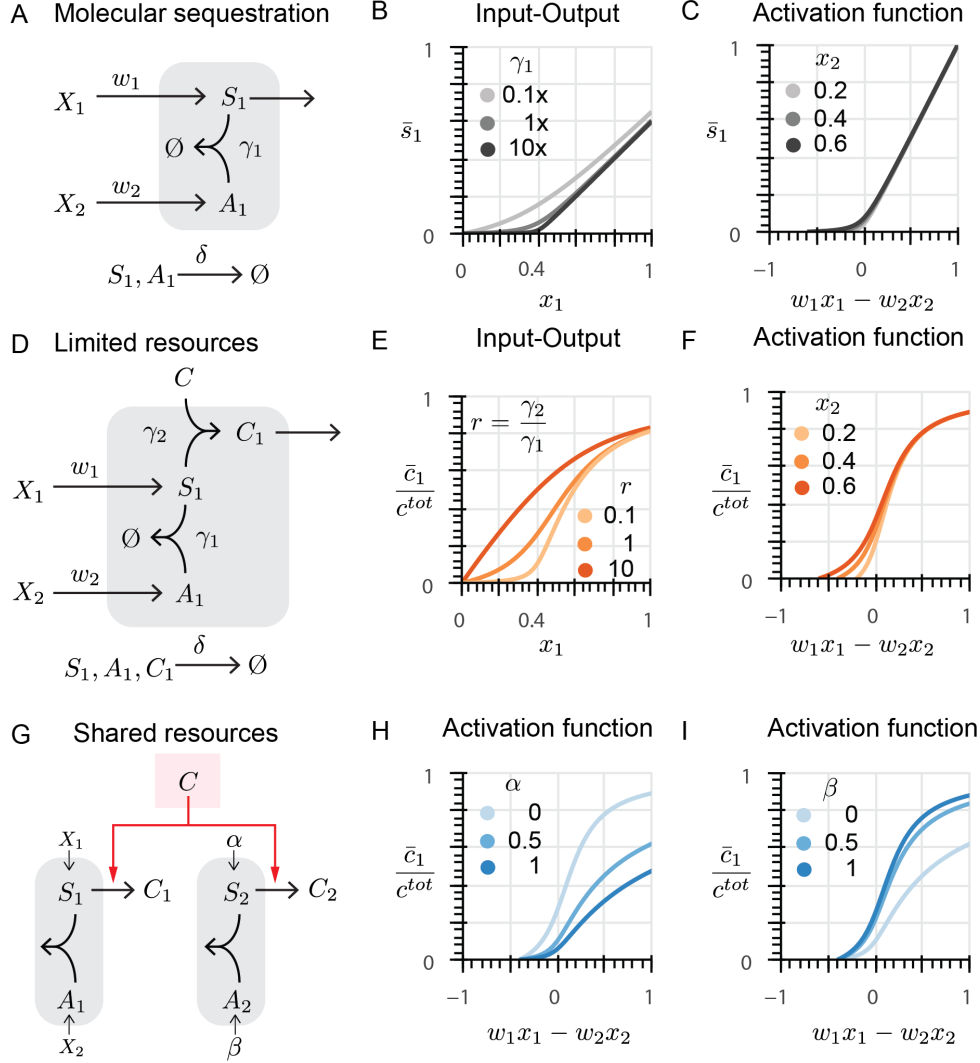


Figure 2: Molecular sequestration of a sigma factor (S_1) and its anti-sigma (A_1) recapitulates variants of ReLU activation function. **A:** Schematic representation of chemical reactions depicting molecular sequestration of S_1 factor and A_1 . The inputs are concentrations of species X_1 and X_2 that generate S_1 and A_1 , respectively. The output is steady-state concentration of S_1 shown as \bar{s}_1 . **B:** The effect of sequestration rate (γ_1) on input-output relationship of the sequestration reaction. **C:** The relationship between \bar{s}_1 and the inputs x_1 and x_2 resembles a soft ReLU function in the slow sequestration regime. **D:** Schematic representation of chemical reactions depicting molecular sequestration of S_1 and A_1 as well as complex formation of S_1 with a limited amount of core RNA polymerase (C). The inputs are concentrations of species X_1 and X_2 that generate S_1 and A_1 , respectively. The output is steady-state normalized concentration of S_1 - C dimer (C_1) shown as $\frac{\bar{c}_1}{c^{tot}}$. **E:** The ratio of RNAPol- S_1 complex formation rate to sequestration rate (referred to as competitive binding ratio r) influences the input-output relationship in the sequestration system. Lower r allows construction of a thresholding function. **F:** In fast sequestration regime and slow complex formation ($r \rightarrow 0$ referred to as fast competitive binding regime), the output of the sequestration reactions in the presence of limited resources resembles an asymptotic saturated ReLU (ASReLU). **G:** Schematic representation of chemical reactions depicting molecular sequestration of S_1 and A_1 as well as complex formation of S_1 with a limited amount of core RNA polymerase (C) in the presence of a competing sigma factor (S_2).

The inputs are concentrations of species X_1 and X_2 that generate S_1 and A_1 with rates w_1 and w_2 , respectively. The output is steady-state normalized concentration of S_1 - C dimer (C_1) shown as $\frac{\bar{c}_1}{c^{tot}}$. **H**: The effect of concentration of species α that generates the competing sigma factor S_2 on the activation function of sequestration reaction. The addition of competitive binding reduces the amount of total C available, thus lowering the saturation level of the ASReLU function. **I**: The effect of concentration of species (β) that generates A_2 that sequesters the competing S_2 on the activation function of sequestration reaction. Higher sequestration of S_2 results in a higher concentration of available C , thus increasing the saturation level of the ASReLU function.

competitive binding regime), \bar{c}_1 follows the thresholding function (**Fig. 2E**). Hence, when input x_1 is lower than the threshold, there is no response, but when input is higher than the threshold, the response is non-zero and ultimately saturates at 1 (c^{tot}).

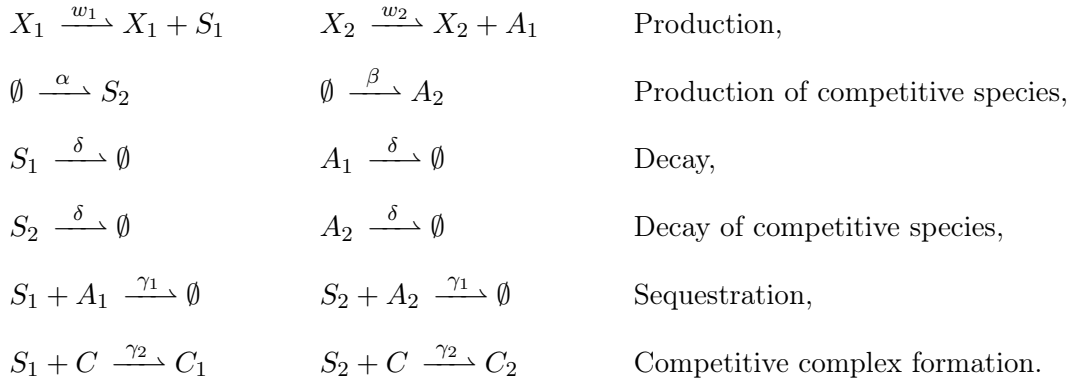
Equation (2) describes the input-output function in the fast competitive binding regime ($r \ll 1$).

$$\lim_{r, \xi \rightarrow 0} \bar{c}_1 = \max \left[0, \min \left[c^{tot}, \frac{x_1 w_1 - x_2 w_2}{\delta} \right] \right] \quad (2)$$

Further, equation (2) is closely similar to AReLU, equation (1), with the difference of having a limit on the output (c^{tot}) which is introduced due to the limited resources. Thus, the effect of limited resources, in this case, total RNAPol core, causes the deviation of \bar{c}_1 from a linear trend when x_1 and x_2 vary. Nevertheless, in the fast competitive binding regime even with a limited RNAPol core, the activation function still performs as a subtraction calculator although it is saturated at c^{tot} (**Fig. 2F**). Equation (2) characterizes this non-linear relationship of \bar{c}_1 with inputs x_1 and x_2 . We call this activation function asymptotic saturated ReLU (referred to as ASReLU hereafter). It is worth noting that while the value of c^{tot} in *in vitro* systems can be controlled by simply diluting the reaction, *in vivo* regulation of total amount of resources (C) could be achieved *via* inhibition of RNAPol activity by a small peptide called Gb2(54) or overexpression of RNAPol subunits β and β' as described by Izard *et al.*(55)

Despite the ASReLU maintaining its function as a subtractor when the resources are

limited, it is rarely the case for substrates to interact with their binding partners without other competing factors. For example, various sigma factors may compete with each other to bind to the available RNAPol core molecules.(18, 22, 43–45) Thus, when this competition is considered, the function of ASReLU may change. A simplified model consisting of two competing sigma factors (S_1, S_2) and their corresponding anti-sigma factors can be utilized to predict the function of ASReLU in the presence of competing sigma factors (**Fig. 2G**). For simplicity, here we assume that the competing sigma factors have identical affinities for their corresponding anti-sigma proteins as well as C . The effect of different kinetics will be investigated in the next section. The following chemical reactions represent the model:



Solving at steady state, the set of ODEs modeling above equations yields equation (3) (for mathematical derivation, see SI section 1.3), which describes an expression of \bar{c}_1 as a function of inputs (x_1 and x_2) and c^{tot} . Although it is challenging to find a closed-form expression of \bar{c}_1 at the steady state, we can find an expression of \bar{c}_1 as a function of the inputs x_1, x_2 , and \bar{c}_2 (unknown). This expression is useful to understand the effect of coupling between the two competing sigma factors on their binding to the RNAPol. Equation (3) shows that the available RNAPol is depleted by the other sigma factor reflected in the term $c^{tot} - \bar{c}_2$. Additionally, the amount of depletion of available resources will depend on the production rates of both the competing sigma factor and its anti-sigma protein. Yet, in

the fast sequestration regime ($\xi \rightarrow 0$), and fast competitive binding ($r \rightarrow 0$), the system converges to an ASReLU function with a lower saturation magnitude captured by $c^{tot} - \bar{c}_2$.

$$\lim_{r, \xi \rightarrow 0} \bar{c}_1 = \max \left[0, \min \left[c^{tot} - \bar{c}_2, \frac{x_1 w_1 - x_2 w_2}{\delta} \right] \right] \quad (3)$$

Since \bar{c}_2 only influences the ASReLU saturation level, the performance of the ASReLU function in calculating the difference between S_1 and A_1 as a thresholding function is sustained under the effect of a competing sigma factor (**Fig. 2H**). However, the limit of output \bar{c}_1 (the saturation level) further decreases as the amount of competing sigma factor s_2 increases which corresponds with higher α (**Fig. 2H**). This trend reverses back to the non-competing ASReLU (equation (2)) when the amount of anti-sigma factor A_2 increases with higher β since less S_2 is available to compete with S_1 (**Fig. 2I**).

Overall, molecular sequestration in a sigma factor-dependent translation system models variations of the ReLU function. In ideal conditions where the RNAPol core is unlimited, this dependency is perfectly quasi-linear and reflects the difference between S_1 and A_1 , enabling an AReLU function in a fast sequestration regime. However, in the presence of limited as well as shared resources, the trend between the sigma factor-RNAPol complex and the differential value of S_1 and A_1 deviates from a linear trend and this deviation becomes more intense as the total amount of free S_2 increases. This relationship in fast sequestration and competitive binding regimes is captured by the ASReLU function. However, since ASReLU still represents the difference of S_1 and A_1 , we reasoned that the activation function can be used to create a perceptron, called sigma-based perceptron hereafter, that can pave the way for constructing multi-layer neural networks for generating non-linear outputs.

2.2 Design and analysis of a sigma-based perceptron with ASReLU activation function and sharing limited resources

After confirming the output of the sigma-based perceptron with an ASReLU activation function, we seek to determine whether this system demonstrates typical characteristics of a single node or perceptron in a neural network in the presence of a competing node. We test various input ranges of the perceptron and its competing counterparts to investigate the perceptron’s linear decision boundary as well as weight-tuning for manipulating its decision boundary. Thus, we model the binding of a sigma factor (S_1) to RNAPol in the presence of its anti-sigma (A_1) as well as a competing sigma-anti-sigma pair (S_2 and A_2) and look at the steady-state total amount of RNAPol bound to S_1 (denoted as \bar{c}_1) as the output of the node in response to a wide range of inputs (**Figs. 3A and B**).

First, we investigate the output of the single node in isolation (without competing factors) to understand the effect of binding kinetics between S_1 and A_1 (γ_1) as well as S_1 and C (γ_2) on the perceptron’s decision boundary. Our findings show that the competitive binding ratio plays a critical role in determining the response pattern or the decision boundary of the perceptron. Equation (4) below provides insight on how r and total resources c^{tot} affect the decision boundary (see SI section 1.2 for derivation):

$$\bar{c}_1 = \frac{w_1 x_1}{\delta} - \left(\frac{\bar{c}_1}{\bar{c}_1 + r(c^{tot} - \bar{c}_1)} \right) \frac{w_2 x_2}{\delta} - \left(\frac{\delta}{\gamma_2} \right) \frac{\bar{c}_1}{c^{tot} - \bar{c}_1} \quad (4)$$

The last term in the equation (4) simply introduces a bias to the decision boundary which only influences the response amplitude while leaving the response pattern, dictated by weights w_1 and w_2 , intact. Additionally, the coefficient of x_2 depends on both variables w_2 and $\frac{\bar{c}_1}{\bar{c}_1 + r(c^{tot} - \bar{c}_1)}$. Therefore, the competitive binding ratio, depending on its magnitude, can change the slope of the decision boundary. When $r \rightarrow 0$ (**Fig. 3C**, right), the coefficient of $\frac{w_2 x_2}{\delta}$ in equation (4) converges to 1. Hence, in this regime, the pattern of decision boundary remains linear across the input ranges although its bias changes depending on the inputs.

However, as r becomes greater, the effect of limited resources on the decision boundary strengthens (**Fig. 3C**, left). In such a condition, the decision boundary still remains linear, but since \bar{c}_1 is a function of x_1 and x_2 , the slope of the decision boundary varies across the range of the inputs as an inverse function of $c^{tot} - \bar{c}_1$. Intuitively, when $r \rightarrow 0$, the sequestration rate γ_1 is higher than γ_2 . Thus, the available amount of S_1 becomes equal to $x_1w_1 - x_2w_2$ which corresponds to excess amount of S_1 after binding to A_1 . In this case, regardless of the binding rate between S_1 and the RNAPol C , the output will be a simple linear function of excess S_1 . On the other hand, when the sequestration rate γ_1 is slower than the binding rate between S_1 and C (γ_2), the dynamics of binding between S_1 and A_1 in conjunction with binding of S_1 and C dictates non-linearity on the output of the sequestration and complex formation reactions (seen in **Fig. 2F**). The decision boundary of the perceptron, consequently, consists of linear decision boundaries with slopes that change as the inputs vary. Nevertheless, as long as $r \ll 1$, the perceptron still functions similarly to the ideal condition (when $r \rightarrow 0$) and generates a decision boundary which resembles the ideal linear condition and can be used for the construction of more complicated architectures.

Importantly, in the construction of our model, we assumed that the sequestration species S_1 and A_1 are generated from an orthogonal RNA polymerase. *i.e.*, production of S_1 and A_1 does not consume the shared resources C . Such a scenario corresponds to systems in which input species promote the production of S_1 and A_1 through polymerases such as T7 or SP6 RNA polymerase (we call this scenario uncoupled input-output relationship) instead of a housekeeping sigma factor (*e.g.*, sigma 70. We call this scenario coupled input-output relationship). To investigate the effect of resource consumption by input species (*i.e.*, if S_1 and A_1 are generated by complexes of housekeeping sigma factors with core RNAPol), we modified our model to include coupling between the resource consumption of S_1 and input species X_1 and X_2 (**Fig. S2A**). Assuming that housekeeping sigma factor S_C is responsible for the production of S_1 and A_1 via $S_C - C$ complexes denoted as C_S and C_A , respectively, our model shows the effect of resource consumption by S_C in different input

regimes (**Figs. S2B and C**). Specifically, when $[x_1 \rightarrow 1, x_2 \rightarrow 1]$, both C_S and C_A use the limited resource C , thus causing a reduction in the production amount of sigma factor and anti-sigma protein compared to the uncoupled case. Consequently, this effect leaves less resources for S_1 . Therefore, the response magnitude of the perceptron with a coupled input-output relationship becomes lower than the uncoupled case while the decision boundary remains similar (**Figs. S2D and E**). Therefore, while the perceptron still demonstrates the linear decision boundary when input species consume limited resources, sharing limited resources between the perceptron sigma factor and the perceptron inputs suppresses the output magnitude. Therefore, in practice, production from input species through orthogonal RNA polymerases like T7 and SP6 would be preferred as it circumvents the limitations of the input-output coupling.

Next, we seek to investigate how the total amount of RNAPol, C , affects the perceptron function. Since C does not play a role in the dynamics of the ASReLU function, its variation reveals itself as a simple increase or decrease in perceptron response amplitude (**Fig. 3D**). This effect occurs because the normalized output of the ASReLU, \bar{c}_1^n , is completely independent of C and varies only by x_1 and x_2 (see equation (23) in the SI). Therefore, a single perceptron node in the absence of any competing sigma factors and in the presence of limited resources still acts as a linear classifier. However, the dynamic effects of anti-sigma-sigma binding on the system effectively change the weights of the perceptron decision boundary. Such changes, however, are minimal when $r \ll 1$.

So far, we have focused on the decision boundary of the biological perceptron made by a sigma factor and its corresponding anti-sigma protein. Next, we consider whether the presence of another sigma factor (S_2) and its anti-sigma protein (A_2) changes the pattern of the perceptron decision boundary. Equation (5) provides an expression for the output of the perceptron as a function of inputs, \bar{c}_1 , and \bar{c}_2 . (See SI section 1.3 for derivation.)

$$\bar{c}_1 = \frac{w_1 x_1}{\delta} - \left(\frac{\bar{c}_1}{\bar{c}_1 + r(c^{tot} - \bar{c}_1 - \bar{c}_2)} \right) \frac{w_2 x_2}{\delta} - \left(\frac{\delta}{\gamma_2} \right) \frac{\bar{c}_1}{c^{tot} - \bar{c}_1 - \bar{c}_2} \quad (5)$$

Equation (5) can be interpreted as an alternative form of equation (4) if c^{tot} in equation (4) is replaced with $c^{tot} - \bar{c}_2$. In other words, since the presence of a competing perceptron shrinks the amount of available resources, it amplifies the bias introduced to the perceptron output (by lowering the magnitude of the denominator in the last term in equation (5)) and strengthens the input-dependent slope variation in decision boundary of the main perceptron (by increasing the coefficient of $\frac{w_2 x_2}{\delta}$ in equation (5)). Note that \bar{c}_1 in equation (5) is a function of inputs x_1 and x_2 . Therefore, the magnitude of introduced bias and change in the slope of the decision boundary will be varied in different input regimes. However, if $r \ll 1$, the effect of competition and resource sharing on tuning decision boundary becomes negligible as the coefficient of $\frac{w_2 x_2}{\delta}$ in equation (5) converges to 1.

Our simulations for a perceptron in a fast competitive binding regime over a range of different input concentrations verify our mathematical analysis by demonstrating that in the presence of another sigma factor (produced by input α_2) competing to bind to the RNAPol, the perceptron response is suppressed due to the bias introduced by the competing perceptron which lowers the amount of available RNAPol (**Fig. 3E**). The response pattern, on the other hand, remains mainly intact due to the low competitive binding ratio.

We also investigate how anti-sigma A_2 (produced by input β_2) influences the response of the perceptron in a resource-sharing system. Binding of A_2 to S_2 disables S_2 from binding to RNAPol, thereby increasing the total amount of RNAPol available for S_1 to bind. Therefore, we expect that the introduction of A_2 to the system suppresses the bias effect of S_2 on the perceptron response pattern (last term in equation (5)). Indeed, our simulations show that an increase in production of A_2 increases the perceptron output (**Fig. 3F**), thus confirming our hypothesis.

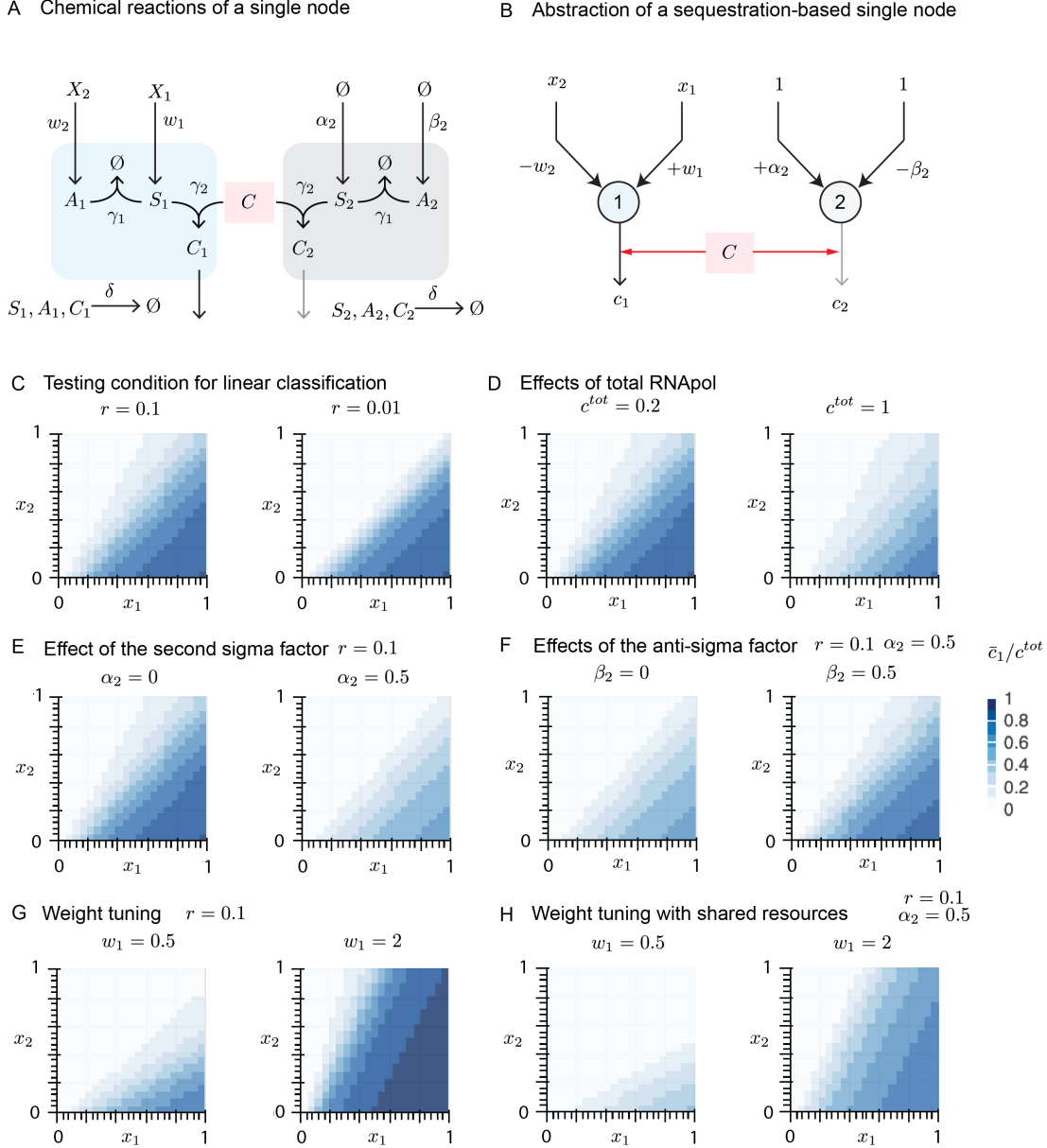


Figure 3: Sigma-based biomolecular network with thresholding response demonstrate characteristics of a perceptron. **A:** Chemical reactions representing a single perceptron in the presence of a competing node. See table 1 for the value of input weights. **B:** The equivalent representation of the chemical reactions in the context of artificial neural networks. **C:** Linear decision boundary of a single node in the absence of competition. Lower r demonstrates linear decision boundary while higher r generates linear decision boundaries with weights that vary as a function of input. **D:** Linear decision boundary of a single node with different amounts of total C in the absence of competition. A higher concentration of C allows a higher range of responses and lack of saturation observed with lower c^{tot} . **E:** Linear decision boundary of a perceptron in the presence of a competing node. α_2 represents the concentration of species that produces competing S_2 . A higher amount of S_2 reduces the available concentration of C , thus suppressing the response of the decision boundary while leaving its pattern intact. **F:** The effect of concentration of A_2 (controlled by β_2) which sequesters competing S_2 on the linear decision boundary of a perceptron in the presence of limited resources and a competing node. Higher sequestration yields a higher available concentration of C , thus increasing the amplitude of perceptron response.

G: Linear decision boundary of a single node in the absence of a competing factor can be tuned by adjusting the weight of the input even when $r = 0.1$. **H:** Tuning the linear decision boundary of the perceptron in the presence of a competing node can still be realized by changing the input weight. The amplitude of the response, however, is suppressed due to the lower availability of C .

Our mathematical analysis of the perceptron output in the presence of a competing perceptron leads to an expression for \bar{c}_1 that is dependent on inputs and \bar{c}_2 (equation (48) in the SI), indicating that kinetics of binding between S_2 and A_2 would reveal its effect on perceptron output by simply varying its response amplitude while leaving its response pattern intact as long as the perceptron functions in the fast competitive binding regime (see SI section 1.3 for mathematical derivation. See **Figs. S3 and S4** for simulations over a wide range of different S_2 and A_2 concentrations and different kinetics of S_2 - A_2 binding, respectively).

Lastly, as weight tuning is a fundamental characteristic of nodes in neural networks that defines their individual decision boundaries, we determine the feasibility of tuning the weights applied to our biological perceptron in the presence or absence of resource sharing. We first analyze an individual node with a limited amount of C and observe the change in the output pattern as w_1 increased (**Fig. 3G**). Similar to perceptrons used in ANNs, adjusting the weight applied to the input results in a discernible change in the slope of the response pattern which is in agreement with our mathematical analysis (see 1.3 in SI). Aligned with our expectation, introducing a second node S_2 that imposes resources sharing on the system preserves the weight-tuning characteristic of the perceptron and only affects its response amplitude by lowering the saturation limit (**Fig. 3H**).

In addition to steady-state analysis, we investigated the transient response of the sigma-based perceptron to study its function before it reaches its stable decision boundary. We observed that the perceptron output reaches the steady state shortly after the beginning of the simulation regardless of the parameters defining the concentration of species or the kinetics of the sequestration reaction (**Fig. S5A**). Additionally, the perceptron decision boundary forms even in the transient state (**Fig. S5B**). While in all scenarios with different

values of α and β in either slow or fast sequestration regimes, the decision boundary is similar to the steady-state response pattern (*i.e.*, maximum response generation at $[x_1 \rightarrow 1, x_2 \rightarrow 0]$), the response magnitude is smaller than the steady-state response.

Taken together, our simulations elucidate the effect of a competing sigma factor on the perceptron output magnitude in the fast competitive binding regime and also demonstrate the weight-tuning ability of perceptron with or without resource sharing. Therefore, we concluded that by operating in a fast competitive binding regime, sequestration-based perceptrons can demonstrate linear decision boundaries despite sharing limited resources which in turn allows the construction of multi-layer perceptron networks.

2.3 Biomolecular neural networks generate non-linear classifiers in the presence of shared resources

While a single sigma-based perceptron generates a tunable decision boundary even with sharing limited resources (**Fig. 3**), most biologically relevant biocomputation processes such as competitive ligand binding and protein dimerization generate sophisticated non-linear responses that rely on the protein-protein interactions and the concentration of competing dimerizing proteins or ligands (23, 24). Therefore, we aim to utilize the sigma-based perceptron as a basic building block of more intricate networks made of multiple perceptrons that are capable of generating non-linear outputs.

First, we investigate a two-node network where each node representing a sigma-based perceptron receives two inputs with unique weights (**Fig. 4A**). This simple network allows us to study the effect of sharing limited resources on the output of nodes in the same layer (see SI section 2.1 for mathematical representation of the network). We note that the output of each node or perceptron represents the molecular complex made by binding of the perceptron sigma factor to the shared resources RNAPol (as depicted in **Fig. 2D**) which in turn can promote expression of another sigma factor (corresponding to positive weight) or another anti-sigma molecule (corresponding to negative weight). Therefore c_i stands for the complex

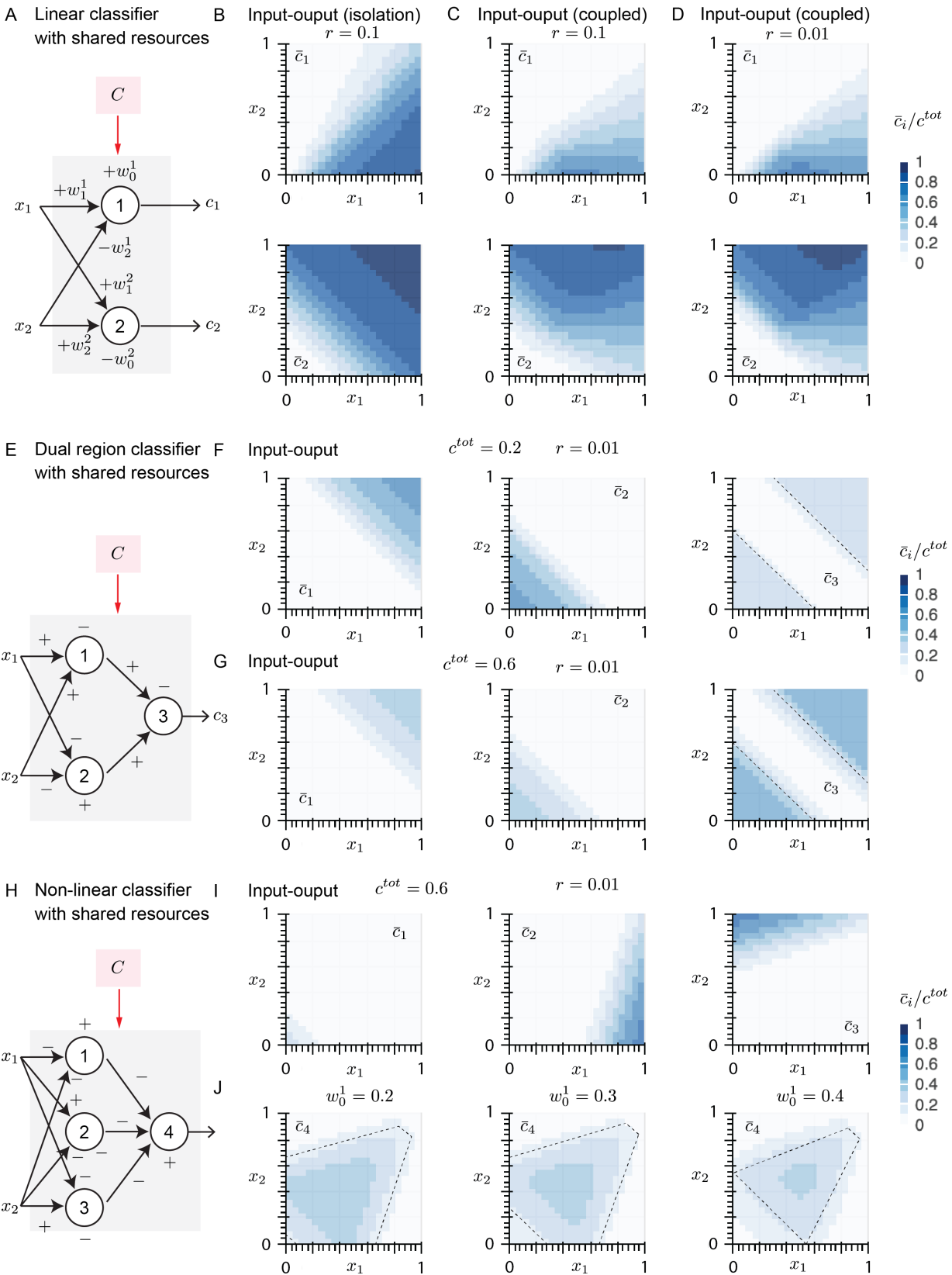


Figure 4: Biomolecular neural networks enable implementation of linear and non-linear classifiers: **A:** Schematic representation of a single-layer neural network that generates a linear classifier. Both perceptrons share the limited resources C . See table 1 for the value of input weights. **B:** Linear decision boundary of perceptron 1 (top) and 2 (bottom) in the absence of competitive binding. **C:** Decision boundary of perceptrons 1 (top) and 2 (bottom) when perceptrons compete to bind to the limited resources C . The competitive binding ratio r is 0.1. **D:** Decision boundary of perceptrons 1 (top) and 2 (bottom) when perceptrons compete to bind to the limited resources C . The competitive binding ratio r is 0.01. Competitive binding introduces non-linearity to the classifier by reducing the output amplitude in input regimes in which both perceptrons require resources. **E:** Schematic representation of the architecture of the dual region classifier multi-layer neural network with shared resources. + and - signs indicate the weights of inputs and whether they promote the expression of sigma or anti-sigma proteins. See table 1 for the value of input weights. **F:** Output maps of the perceptrons in the first layer (left and middle for C_1 and C_2 , respectively) each representing a linear classifier with minimal interference. Right panel depicts network output (c_3) which comprises the dual region classifier with $c^{\text{tot}} = 0.2$. The dashed lines represent the decision boundary of an ideal perceptron with the same weights which function in the absence of resource constraints. **G:** Output maps of the perceptrons in the first layer (left and middle) and network output (right) which comprises the dual region classifier with $c^{\text{tot}} = 0.6$. The pattern of the classifier is invariant to c^{tot} while its amplitude varies with c^{tot} . The dashed lines represent the decision boundary of an ideal perceptron with the same weights which function in the absence of resource constraints. **H:** Schematic representation of the multi-layer neural network with a non-linear classifier output. + and - signs indicate the weights of inputs and whether they promote the expression of sigma or anti-sigma proteins. See table 1 for the value of input weights. **I:** The output of perceptrons in the first layer of the network. Perceptrons 1 (left), 2 (middle), and 3 (right) each recapitulate a linear classifier with zero to little interference in their linear decision boundaries. **J:** The output of the network depicted in **H** (c_4) is a non-linear classifier. The classifier boundary can be tuned by adjusting the biases of nodes in the first layer. Shown are examples of different decision boundaries by changing the bias of perceptron 1 to 0.2 (left), 0.3 (middle), and 0.4 (right). For each case, the dashed lines represent the decision boundary of an ideal perceptron with same weights which functions in the absence of resource constraints.

between S_i and C as the output of each node in the networks illustrated in **Fig. 4**.

In the absence of resource sharing, each node's output is a linear combination of its inputs (**Fig. 3C**). Therefore, with given weights denoted in **Fig. 4A**, linear patterns for outputs of nodes 1 and 2 are observed (**Fig. 4B**). However, when the binding competition between sigma factors is taken into account, the effect of limited resources and competition on the network output is elucidated as the attenuation of outputs for each node in certain input regimes (**Fig. 4C**) resulting in a non-linear response pattern. Notably, the response of each node is maximally affected where the competing node is consuming most of the resources. For instance, in isolation, the output of node 2 (C_2) has its highest expression

level when $[x_1, x_2] \rightarrow 1$ (**Fig. 4B**, bottom). Consequently, the response pattern of C_1 is highly diminished in that region (**Fig. 4C**) due to the effect of resource sharing with node 2. In a similar fashion, the output pattern of C_2 demonstrates its highest attenuation where C_1 is expressed the most in isolation ($x_1 \rightarrow 1$), thus leaving a smaller amount of available RNAPol (C) for production of C_2 . Hence, the concerted effect of resource sharing in a wide range of inputs imposes non-linearity in the output of nodes in the same layer due to the emergence of a dominant perceptron that consumes most of the resources. This effect can also be deduced from equation (5) knowing that both \bar{c}_1 and \bar{c}_2 are functions of x_1 and x_2 . Therefore, in certain input regimes where either C_1 or C_2 are strongly expressed, the effects of competitive resource sharing detailed in the previous section become strengthened.

Interestingly, the network output is influenced differently in fast and slow competitive binding regimes. In a fast competitive binding regime, \bar{c}_1 and \bar{c}_2 are linear functions of x_1 and x_2 . Therefore, the outputs C_1 and C_2 display sharper deviations when both are dependent on inputs (**Fig. 4D**) due to the stronger effect of the bias term in equation (5) imposed by sharp reduction of resources (see **Fig. S6** for isolated perceptron responses in fast competitive binding regime.). Overall, we conclude that resource sharing can significantly induce non-linearity to the network output pattern and the highest impact of competitive resource sharing reveals itself where the response of isolated nodes overlap with each other, corresponding to input regimes with maximum resource sharing.

Knowing that overlapping responses can cause non-linearity in the network outputs, we next consider whether we can still implement non-linear networks with decision boundaries that are distorted minimally due to resource sharing. We reason that if the outputs of the same layer do not overlap with each other, the overall output of the network will follow the design principles dictated by the input weights and will not have undesired non-linearity. To test this hypothesis, we look at a simple network made of three nodes that reconstitutes a dual region classifier (band-stop filter) represented in **Fig. 4E** (see SI section 2.2 for mathematical representation of the network and **Fig. S7** for the ideal response of this

network in the absence of sharing limited resources). In order to avoid non-linearity induced by resource sharing, we design the network such that the node outputs have minimal overlap with each other (**Fig. 4F**, left and middle panels). We expect that the lack of overlap between node outputs would prevent unwanted non-linearity in the network output (C_3). Aligned with our expectation, our simulations demonstrate the network response consisting of two separate linear regions despite resource sharing (**Fig. 4F**, right). We also confirmed that an increase in the amount of available resources (C) decreases the effect of competition for resources and does not alter the decision boundary of first layer nodes (**Fig. 4G**, left and middle) as well as overall network (**Fig. 4G**, right). Therefore, we conclude that linear responses can be engineered in a network, despite the presence of resource sharing, by tuning input weights such that the outputs of the nodes in the same layer do not overlap.

Finally, we seek to create a biomolecular neural network that generates a non-linear classifier. Such non-linear outputs are key features of many biological processes and drivers of cell decision-making (56–58). We design a simple network with specific input weights based on the interaction of multiple sigma factors and their corresponding anti-sigma molecules that generates a non-linear classifier resembling a band-pass filter (**Fig. 4H**, see SI section 2.3 for ODEs describing the network). As in the previous design, we want the outputs of the first layer to have minimal overlap with each other to avoid non-linearity induced by resource sharing. Our simulations depict that the first layer outputs, C_1 , C_2 , and C_3 have minimal overlap with each other given the particular weights and biases (**Fig. 4I**). Consequently, our simulations demonstrate that network output (C_4) creates the non-linear classifier (**Fig. 4J**) that is aligned with our ideal design expectations (**Fig. S8**). Consistent with our findings in the previous section, changing the c^{tot} does not change the non-linear pattern of the decision boundary but tunes its amplitude (**Fig. S9**).

We also look at the effect of input bias on the first layer of the output (**Fig. 4J**). Indeed, in accordance with the ideal design (**Fig. S8**), increasing the bias of node 1 amplifies its steady-state output (\bar{c}_1) which subsequently further sequesters the amount of available

S_4 , thus suppressing the network response in its corresponding input regime ($[x_1, x_2] \rightarrow 0$) without significant deviation from design expectations (**Fig. S8**). Therefore, we showed that even in the presence of resource sharing, we can develop non-linear biological neural networks to realize dual region and non-linear classifiers. We also demonstrated that we can modulate the network response by tuning the model parameters that are independent of limited resources.

Table 1: List of parameters used in computational simulations. The value of parameters indicated with * are described in their corresponding figures. The value for rates of production, decay, sequestration, and sigma factor-RNAPol binding are determined based on measurements found in the corresponding reference listed in the column Ref.

Parameter	Symbol	Unit	Value	Figure	Ref.
Perceptron input species	$[x_1, x_2]$	μM	$[0 \rightarrow 1]$	2,3,4	
Competing perceptron input species	$[\alpha, \beta]$	$\mu M h^{-1}$	$[0, 0.5, 1]$	2	(59)
Core RNA polymerase (RNAP)	c	μM	$[0 \rightarrow 1]$	2,3,4	
Sigma-anti-sigma sequestration rate	γ_1	$\mu M^{-1} h^{-1}$	*	2	(60, 61)
Sigma-RNAP complex formation rate	γ_2	$\mu M^{-1} h^{-1}$	10	2	(18, 61)
Degradation rate	δ	h^{-1}	1	2,3,4	(59)
Node 1 input weights	$[w_1, w_2]$	h^{-1}	$[1, 1]$	3	(59)
Node 2 input weights	$[\alpha_2, \beta_2]$	h^{-1}	$[0, 0]$	3	
			$[0.5, 0.5]$	3	
Node 1 input weights	$[w_0^1, w_1^1, w_1^2]$	h^{-1}	$[1, 1, 1]$	4A	
Node 2 input weights	$[-w_0^2, -w_2^1, w_2^2]$	h^{-1}	$[0.5, 1, 1]$	4A	
Node 1 input weights	$[-w_0^1, w_1^1, w_1^2]$	h^{-1}	$[1.2, 1, 1]$	4E	
Node 2 input weights	$[w_0^2, -w_2^1, -w_2^2]$	h^{-1}	$[0.7, 1, 1]$	4E	
Node 3 input weights	$[-w_0^3, w_1^3, w_2^3]$	h^{-1}	$[0.15, 4, 4]$	4E	
Node 1 input weights	$[w_0^1, -w_1^1, -w_1^2]$	h^{-1}	$[0.4, 0.5, 0.5]$	4H	
Node 2 input weights	$[-w_0^2, w_2^1, -w_2^2]$	h^{-1}	$[0.8, 1.5, 0.5]$	4H	
Node 3 input weights	$[-w_0^3, -w_1^3, w_2^3]$	h^{-1}	$[0.8, 0.5, 1.5]$	4H	
Node 4 input weights	$[w_0^4, -w_1^4, -w_2^4, -w_3^4]$	h^{-1}	$[0.3, 1, 1, 1]$	4H	

3 Discussion

Biological signal processing units that reconstitute molecular linear and non-linear classifiers are powerful tools that enable cellular decision-making, precise cell programming, highly

discriminatory input processing, and ultra-sensitive molecular biosensors for applications such as CAR T-cell engineering or *in vitro* diagnostics. A biological perceptron ideally allows linear classification while a combination of biological perceptrons can create biomolecular neural networks that compute non-linear classification.

Among different approaches that are used to construct biochemical neural networks and classifiers, sequestration-based networks are of significant interest because of their simplicity and compatibility with *in vivo* systems (29, 36). However, the effect of physiological constraints such as a limitation on resources, as well as competitive binding between elements of molecular classifier networks, specifically sequestration-based networks, on the classification function have remained unexplored.

We mathematically modeled sequestration-based biochemical neural networks and investigated how sharing limited resources, a ubiquitous feature of physiological systems, influences the function of the neural network decision boundary. We chose a network of sigma factors, their corresponding anti-sigma proteins, and core RNA polymerase as our model system. Our analyses demonstrated that a single perceptron, the basic building block of neural networks, with a ReLU-like activation function is recapitulated by modeling the interaction of one sigma factor with its anti-sigma molecule. We further showed that modifying the model to include a limited amount of core RNAPol in the system changes the activation function of the sigma-based perceptron to an asymptotic saturated ReLU. Drawing inspiration from natural systems where multiple sigma factors compete to bind to a limited pool of RNAPol, we altered the model to include another sigma factor and found that the decision boundary of the sigma-based perceptron remains the same although its output is suppressed.

While conditions of biochemical *in vitro* reactions are primarily controlled, in living organisms, endogenous factors can cause perturbations and deviations from the ideal design. For instance, in bacteria, although a certain sigma factor might be designed to control gene expression, multiple other sigma factors compete with the engineered sigma factor to bind to a limited amount of RNAPol. To include competitive binding to shared resources, we

increased the number of perceptrons, controlled by the same inputs, to two. We found that in specific input regimes where the outputs of the perceptrons interfere with each other, one dominant perceptron emerges and consumes most of the resources whereas the decision boundary of the non-dominant perceptron significantly deviates from its ideal design.

Given that engineering any kind of non-linear response by neural networks requires multi-layer perceptrons, we investigated conditions where despite resource sharing, non-linear decision boundaries could be designed. Knowing that interference of outputs of perceptrons in the same layer raises deviations of the linear decision boundary, we engineered particular multi-layer perceptrons in which non-linear decision boundaries were successfully demonstrated. We showed that despite sharing limited resources, dual region and non-linear classifiers resembling band-stop and band-pass filters can be implemented in sigma-based neural networks using different architectures of 2-layer neural networks with minimal deviation from ideal design.

Since sharing limited resources is not exclusive to the sigma-based neural networks, our findings can be extended to other biocomputational input-processing systems used in living cells. Taking Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) gene editing as an example, the sequestration of a single guide RNA (sgRNA) strand by its complementary RNA, or anti-guide RNA, is analogous to the interaction of a sigma factor with its anti-sigma protein. If the sgRNA is engineered to drive a CRISPR reaction, the Cas protein will consequently be the limited resource that all sgRNAs will compete to bind to. Similarly, a limited pool of proteolytic substrates that are catalyzed by a protease-based neural network (38, 62) will follow the functional principles outlined in this work. Therefore, the principles of input classification under the effects of sharing limited resources outlined in this paper can be extended and used for designing other molecular classifiers.

An advantage of sigma-based sequestration-based neural networks over DNA-based neural networks is that they originate from endogenous proteins in bacteria. Therefore, sigma-based neural networks can be implemented both in bacterial systems as well as bacterial lysate-

based cell-free expression systems, thus expanding their applications as living computers as well as *in vitro* biosensors and synthetic cell-based biocomputers(63, 64). To indicate the possible implementation of neural networks demonstrated here (**Fig. 4**), we present schematics of required DNA sequences for each network in **Figs. S10** and **S11**. Given that construction and function of sigma-based perceptron relies on transcription-translation as well as protein degradation, the perceptron parameters (*i.e.*, input weights), which collectively control the whole network behavior, can be determined with common tools of tuning genetic circuits.(4, 65, 66) For example, changing the strength of ribosome-binding site or promoters that control expression of sigma and anti-sigma proteins can tune the slope of the perceptron decision boundary. Other controllable elements such as plasmid copy number (changed by switching origin of replication), degradation rate (changed by addition of degradation tags to the perceptron proteins), or sequestration rate (tuned via mutagenesis in sigma and anti-sigma proteins) may regulate the function of perceptron by effectively changing its weights as well, the practical demonstration of which awaits future studies.

Additionally, unlike neuromorphic systems made of inducible genetic circuits, in sequestration-based networks, both positive and negative weights can be engineered, thereby making them more flexible and applicable for generating non-linear response patterns. Since outputs of sigma-based perceptrons are transcriptionally active sigma-RNAPol complexes, they can be programmed to drive the expression of other sigma factors. Relying on this characteristic of sigma-based perceptrons, multi-layered perceptrons that are capable of recapitulating sophisticated non-linear response patterns can be designed readily. However, as shown in this work, for the neural network to function as designed and avoid perturbations caused by sharing limited resources among perceptrons, the outputs of perceptrons in the same layer must have minimal interference.

Although using sigma factors —that are endogenous molecules in bacterial systems— as components of a biocomputational module offers flexibility and simple implementation, we note that utilizing sigma factors as building blocks of BNNs could face challenges originating

from their native roles in physiology of bacteria. Overexpression of endogenous sigma factors may affect native biological processes in bacteria, cause sigma factor cross-talk with other proteins, or introduce growth issues. Of these, unintentional binding of sigma factors to other proteins could specifically change the behavior of BNN due to unwanted perturbations that effectively change the weights of the network. An avenue to explore in order to circumvent this challenge is to use orthogonal exogenous sigma factors to construct the BNN. Exploiting exogenous sigma factors would prevent unwanted binding of perceptor sigma factors to native molecules as well as interference with biological events in bacteria. In fact, sequestration of *Bacillus subtilis* sigma and anti-sigma proteins sigW and rsiW, respectively, was previously used in *E. coli* for construction of a closed loop biomolecular integral feedback control circuit.⁽⁵¹⁾ Interestingly, the output of this circuit resembles the behavior of a sigma-based perceptor presented in this study, thus promising the potential of exogenous sigma factors as building blocks of BNNs. Further, the challenges of using endogenous sigma factors may be less problematic when BNNs are implemented using *in vitro* bacterial cell-free expression systems due to the absence of endogenous biological processes.

While we demonstrated designs of linear and non-linear classifiers in this work with predetermined weights that generate desirable decision boundaries, we note that sigma-based neural networks are not capable of learning through common algorithms like backpropagation. *i.e.*, the input weights that directly determine the decision boundary are chosen by the designer. However, utilizing the mathematical analysis presented here, one can implement optimization approaches such as particle swarm optimization or other heuristic algorithms to find the appropriate weights for the generation of the desired decision boundary *in silico* prior to testing them *in vivo* or *in vitro*. Such optimization algorithms can also be implemented to expand the function of BNNs to regimes where non-linear effects of resource constraints or perturbations due to non-specific or unwanted bindings described above appear. Therefore, developing BNNs with biological resource constraints capable of recognizing any arbitrary input pattern using optimization methods is an avenue worth exploring in the future.

Recently, it was shown that many cancer cell types can be recognized with higher precision if a combination of two antigens is used to identify them instead of traditionally using one biomarker(67). Implementing non-linear output patterns with sequestration-based neural networks could increase the recognition ability of engineered cellular systems like CAR T-cells by equipping them with information-processing neural networks that generate desired outputs only in designed antigen concentration regimes. Similarly, by coupling inputs to the expression of sigma factors and their anti-sigma molecules, more sensitive, precise, and versatile *in vitro* biosensors for the detection of pathogens, substances, and biomarkers can be constructed.

With the nanofabrication technology in the semiconductor industry approaching its physical limits of manufacturing smaller and smaller elements (68), alternative computational devices with biological components are gaining increasing interest. However, current biocomputational systems are only in their infancy. Although biocomputation in living systems was initially shown more than two decades ago using genetic circuits, the limited range of computational tasks that genetic circuits can perform as well as the digital nature of their input-outputs makes their application limited. With the recent booming advance of AI, biochemical approaches that recapitulate biological neural networks holding the potential to perform intricate computational tasks have gained attention (3, 69). Our study provides a general framework for designing biological perceptron or linear classifiers using existing biomolecular tools in the presence of resource constraints that are ubiquitous in physiological conditions. This framework, thus, is the first step towards designing sophisticated biomolecular neural networks that equip engineered cells with high-level computational and decision-making abilities.

In addition to transformative applications of sigma-based neural networks used in a forward-engineering manner in both cellular and cell-free systems, the fact that sigma and anti-sigma molecules can construct complicated computational modules elucidates the hidden capabilities of these rather simple transcriptional regulation molecules in cells. It was

revealed by Park *et al.* that sigma factors share the resources in a pulsatile manner (18). However, here we demonstrated that sharing limited resources influences sigma-based processes beyond time sharing. Bacterial cells have many sigma factors, some of which are activated only when their inputs meet certain conditions. However, how bacteria utilize the principles of molecular sequestration as well as sharing limited resources to respond differently to input combinations awaits future studies. Additionally, if bacteria are able to process certain input patterns using their endogenous sigma-based neural networks, the nature of these patterns and their role in guiding bacteria to make particular decisions remain unclear. In conclusion, our investigation demonstrates the effects of resource-sharing on sigma-based sequestration-based neural networks with up to four sigma factors and provides an outline for designing sigma-based non-linear neural networks in bacterial systems.

4 Methods

Sequestration and binding reactions were modeled into ODEs using the law of mass action. See SI for the derivation of ODEs describing sigma-antisigma molecular sequestration and the construction of a perceptron based on their behavior. Where possible, analytical expressions are derived for steady-state outputs of reactions. All ODEs were modeled and the steady-state solutions were solved by a custom Python code using the numpy and scipy packages. ODEs were integrated using odeint function from scipy package to obtain transient and steady-state response of reactions. The figures were plotted using Bokeh and Matplotlib packages. The codes used for solving ODEs and generating figures can be found here: [GitHub - mhossein7/BNN_shared_resources](https://github.com/mhossein7/BNN_shared_resources)

Acknowledgement

We thank Jean-Baptiste Lugagne, Noah Olsman, and Elisa Franco for helpful discussions. We thank Cold Spring Harbor Labs (CSHL) and instructors and TAs of CSHL synthetic

biology course during summer 2022. Work in the A.P.L.'s laboratory has been supported by the National Science Foundation (EF1935265). H.M. thanks support from the University of Michigan Rackham Predoctoral Fellowship. I.G. acknowledges assistance from Boehringer Ingelheim Fonds. S.R.C. acknowledges funding from the Mayo Foundation for Medical Education and Research and Mayo Clinic's Department of Molecular Medicine Small Grant. The 2022 CSHL Synthetic Biology course was funded by the National Science Foundation (MCB 2207222), with additional support from the Helmsley Charitable Trust and the Howard Hughes Medical Institute.

Supporting Information Available

The supplementary file contains the derivation of ODEs describing sigma-anti-sigma molecular sequestration at the steady state and the construction of a perceptron based on their behavior.

5 Author Contribution

C.C.S. conceptualized the research. H.M., I.G., and C.C.S. wrote the manuscript. H.M. and C.C.S. performed the mathematical analysis and simulations. All authors revised and approved the manuscript. C.C.S. and A.P.L. provided funding.

6 Conflict of interest

The authors declare no competing financial interest.

References

1. Goñi-Moreno, A., and Nikel, P. I. (2019) High-performance biocomputing in synthetic biology-integrated transcriptional and metabolic circuits. *Frontiers in Bioengineering and Biotechnology* 7, 443035.
2. Dagdia, Z. C., Avdeyev, P., and Bayzid, M. S. (2021) Biological computation and computational biology: survey, challenges, and discussion. *Artificial Intelligence Review* 2021 54:6 54, 4169–4235.
3. Manzoni, R., Urrios, A., Velazquez-Garcia, S., Nadal, E. D., and Posas, F. (2016) Synthetic biology: insights into biological computation. *Integrative Biology* 8, 518–532.
4. Brophy, J. A., and Voigt, C. A. (2014) Principles of genetic circuit design. *Nature Methods* 2014 11:5 11, 508–520.
5. Sedlmayer, F., Aubel, D., and Fussenegger, M. (2018) Synthetic gene circuits for the detection, elimination and prevention of disease. *Nature Biomedical Engineering* 2018 2:6 2, 399–415.
6. Weinberg, B. H., Pham, N. T., Caraballo, L. D., Lozanoski, T., Engel, A., Bhatia, S., and Wong, W. W. (2017) Large-scale design of robust genetic circuits with multiple inputs and outputs for mammalian cells. *Nature Biotechnology* 2017 35:5 35, 453–462.
7. Chen, Z. et al. (2020) De novo design of protein logic gates. *Science (New York, N.Y.)* 368, 78.
8. Chen, Z., and Elowitz, M. B. (2021) Programmable protein circuit design. *Cell* 184, 2284–2301.
9. Yang, X., Rocks, J. W., Jiang, K., Walters, A. J., Rai, K., Liu, J., Nguyen, J., Olson, S. D., Mehta, P., Collins, J. J., Daringer, N. M., and Bashor, C. J. (2023)

- Engineering synthetic phosphorylation signaling networks in human cells. *bioRxiv* 2023.09.11.557100.
10. Kieffer, C., Genot, A. J., Rondelez, Y., and Gines, G. (2023) Molecular computation for molecular classification. *Advanced Biology* 7, 2200203.
 11. Rosenblatt, F. (1958) The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological review* 65, 386.
 12. Cybenko, G. (1989) Approximation by superpositions of a sigmoidal function. *Mathematics of control, signals and systems* 2, 303–314.
 13. Novakovsky, G., Dexter, N., Libbrecht, M. W., Wasserman, W. W., and Mostafavi, S. (2023) Obtaining genetics insights from deep learning via explainable artificial intelligence. *Nature Reviews Genetics* 24, 125–137.
 14. Kim, J., McFee, M., Fang, Q., Abdin, O., and Kim, P. M. (2023) Computational and artificial intelligence-based methods for antibody development. *Trends in Pharmacological Sciences*
 15. Wong, F., Zheng, E. J., Valeri, J. A., Donghia, N. M., Anahtar, M. N., Omori, S., Li, A., Cubillos-Ruiz, A., Krishnan, A., Jin, W., and others (2023) Discovery of a structural class of antibiotics with explainable deep learning. *Nature* 1–9.
 16. Yang, K. K., Wu, Z., and Arnold, F. H. (2019) Machine-learning-guided directed evolution for protein engineering. *Nature methods* 16, 687–694.
 17. Rondelez, Y. (2012) Competition for catalytic resources alters biological network dynamics. *Physical review letters* 108, 018102.
 18. Park, J., Dies, M., Lin, Y., Hormoz, S., Smith-Unna, S. E., Quinodoz, S., Hernández-Jiménez, M. J., Garcia-Ojalvo, J., Locke, J. C., and Elowitz, M. B. (2018) Molecular Time Sharing through Dynamic Pulsing in Single Cells. *Cell Systems* 6, 216–229.e15.

19. Caveney, P. M., Norred, S. E., Chin, C. W., Boreyko, J. B., Razooky, B. S., Retterer, S. T., Collier, C. P., and Simpson, M. L. (2017) Resource Sharing Controls Gene Expression Bursting. *ACS Synthetic Biology* 6, 334–343.
20. Goetz, H., Stone, A., Zhang, R., Lai, Y.-C., and Tian, X.-J. (2022) Double-Edged Role of Resource Competition in Gene Expression Noise and Control. *Advanced genetics* 3, 2100050.
21. Gyorgy, A., Jiménez, J. I., Yazbek, J., Huang, H.-H., Chung, H., Weiss, R., and Del Vecchio, D. (2015) Isocost lines describe the cellular economy of genetic circuits. *Biophysical journal* 109, 639–646.
22. Ishihama, A. (2000) Functional modulation of Escherichia coli RNA polymerase. *Annual Reviews in Microbiology* 54, 499–518.
23. Su, C. J., Murugan, A., Linton, J. M., Yeluri, A., Bois, J., Klumpe, H., Langley, M. A., Antebi, Y. E., and Elowitz, M. B. (2022) Ligand-receptor promiscuity enables cellular addressing. *Cell Systems* 13, 408–425.e12.
24. Parres-Gold, J., Levine, M., Emert, B., Stuart, A., and Elowitz, M. B. (2023) Principles of Computation by Competitive Protein Dimerization Networks. *bioRxiv* 2023.10.30.564854.
25. Qian, Y., Huang, H.-H., Jiménez, J. I., and Del Vecchio, D. (2017) Resource competition shapes the response of genetic circuits. *ACS synthetic biology* 6, 1263–1272.
26. Rizik, L., Danial, L., Habib, M., Weiss, R., and Daniel, R. (2022) Synthetic neuromorphic computing in living cells. *Nature Communications* 2022 13:1 13, 1–17.
27. Li, X., Rizik, L., Kravchik, V., Khoury, M., Korin, N., and Daniel, R. (2021) Synthetic neural-like computing in microbial consortia for pattern recognition. *Nature Communications* 2021 12:1 12, 1–12.

28. Bonnerjee, D., Chakraborty, S., and Bagh, S. (2023) Lego-like Mixing and Matching of Engineered Bacteria Configure Full Subtractor and Adder Through an Artificial Neural Network Type Architecture. *bioRxiv* 2023.06.15.545057.
29. Moorman, A. A. R. Machine learning inspired synthetic biology: neuromorphic computing in mammalian cells. Ph.D. thesis, Massachusetts Institute of Technology, 2020.
30. Pandi, A., Koch, M., Voyvodic, P. L., Soudier, P., Bonnet, J., Kushwaha, M., and Faulon, J. L. (2019) Metabolic perceptrons for neural computing in biological systems. *Nature Communications* 2019 10:1 10, 1–13.
31. Xiong, X., Zhu, T., Zhu, Y., Cao, M., Xiao, J., Li, L., Wang, F., Fan, C., and Pei, H. (2022) Molecular convolutional neural networks with DNA regulatory circuits. *Nature Machine Intelligence* 2022 4:7 4, 625–635.
32. Linden, A. J. V. D., Pieters, P. A., Bartelds, M. W., Nathalia, B. L., Yin, P., Huck, W. T., Kim, J., and Greef, T. F. D. (2022) DNA Input Classification by a Riboregulator-Based Cell-Free Perceptron. *ACS Synthetic Biology* 11, 1510–1520.
33. Cherry, K. M., and Qian, L. (2018) Scaling up molecular pattern recognition with DNA-based winner-take-all neural networks. *Nature* 2018 559:7714 559, 370–376.
34. Lopez, R., Wang, R., and Seelig, G. (2018) A molecular multi-gene classifier for disease diagnostics. *Nature Chemistry* 2018 10:7 10, 746–754.
35. Okumura, S., Gines, G., Lobato-Dauzier, N., Baccouche, A., Deteix, R., Fujii, T., Rondelez, Y., and Genot, A. J. (2022) Nonlinear decision-making with enzymatic neural networks. *Nature* 2022 610:7932 610, 496–501.
36. Moorman, A., Samaniego, C. C., Maley, C., and Weiss, R. A dynamical biomolecular neural network. 2019 IEEE 58th conference on decision and control (CDC). 2019; pp 1797–1802.

37. Samaniego, C. C., Moorman, A., Giordano, G., and Franco, E. Signaling-based neural networks for cellular computation. 2021 American Control Conference (ACC). 2021; pp 1883–1890.
38. Chen, Z., Linton, J. M., Zhu, R., and Elowitz, M. B. (2022) A synthetic protein-level neural network in mammalian cells. *bioRxiv* 2022.07.10.499405.
39. Ceroni, F., Algar, R., Stan, G.-B., and Ellis, T. (2015) Quantifying cellular capacity identifies gene expression designs with reduced burden. *Nature methods* 12, 415–418.
40. Frei, T., Cella, F., Tedeschi, F., Gutiérrez, J., Stan, G.-B., Khammash, M., and Siciliano, V. (2020) Characterization and mitigation of gene expression burden in mammalian cells. *Nature communications* 11, 4641.
41. Chen, P.-Y., Qian, Y., and Del Vecchio, D. A model for resource competition in CRISPR-mediated gene repression. 2018 IEEE Conference on Decision and Control (CDC). 2018; pp 4333–4338.
42. Ogle, C. T., and Mather, W. H. (2016) Proteolytic crosstalk in multi-protease networks. *Physical biology* 13, 025002.
43. Shin, J., and Noireaux, V. (2012) An E. coli cell-free expression toolbox: application to synthetic gene circuits and artificial cells. *ACS synthetic biology* 1, 29–41.
44. Mauri, M., and Klumpp, S. (2014) A model for sigma factor competition in bacterial cells. *PLoS computational biology* 10, e1003845.
45. Farewell, A., Kvint, K., and Nyström, T. (1998) Negative regulation by RpoS: a case of sigma factor competition. *Molecular microbiology* 29, 1039–1051.
46. Lambert, L. J., Schirf, V., Demeler, B., Cadene, M., and Werner, M. H. (2001) Flipping a genetic switch by subunit exchange. *The EMBO journal* 20, 7149–7159.

47. Bae, J.-B., Park, J.-H., Hahn, M.-Y., Kim, M.-S., and Roe, J.-H. (2004) Redox-dependent changes in RsrA, an anti-sigma factor in *Streptomyces coelicolor*: zinc release and disulfide bond formation. *Journal of molecular biology* 335, 425–435.
48. Ades, S. E., Connolly, L. E., Alba, B. M., and Gross, C. A. (1999) The *Escherichia coli* σ^E -dependent extracytoplasmic stress response is controlled by the regulated proteolysis of an anti- σ factor. *Genes & development* 13, 2449–2461.
49. Huang, H.-H., Qian, Y., and Del Vecchio, D. (2018) A quasi-integral controller for adaptation of genetic modules to variable ribosome demand. *Nature communications* 9, 5415.
50. Agrawal, D. K., Marshall, R., Noireaux, V., and Sontag, E. D. (2019) In vitro implementation of robust gene regulation in a synthetic biomolecular integral controller. *Nature communications* 10, 5760.
51. Aoki, S. K., Lillacci, G., Gupta, A., Baumschlager, A., Schweingruber, D., and Khammash, M. (2019) A universal biomolecular integral feedback controller for robust perfect adaptation. *Nature* 570, 533–537.
52. Zhang, C. Joint Training Methods for Tandem and Hybrid Speech Recognition Systems using Deep Neural Networks. Ph.D. thesis, 2017.
53. Samaniego, C. C., Qian, Y., Carleton, K., and Franco, E. (2023) Building subtraction operators and controllers via molecular sequestration. *IEEE Control Systems Letters*
54. Stargardt, P., Feuchtenhofer, L., Cserjan-Puschmann, M., Striedner, G., and Mairhofer, J. (2020) Bacteriophage inspired growth-decoupled recombinant protein production in *Escherichia coli*. *ACS Synthetic Biology* 9, 1336–1348.
55. Izard, J., Gomez Balderas, C. D., Ropers, D., Lacour, S., Song, X., Yang, Y., Lindner, A. B., Geiselman, J., and de Jong, H. (2015) A synthetic growth switch based on controlled expression of RNA polymerase. *Molecular systems biology* 11, 840.

56. Helikar, T., Konvalina, J., Heidel, J., and Rogers, J. A. (2008) Emergent decision-making in biological signal transduction networks. *Proceedings of the National Academy of Sciences* 105, 1913–1918.
57. Bowsler, C. G., and Swain, P. S. (2014) Environmental sensing, information transfer, and cellular decision-making. *Current opinion in biotechnology* 28, 149–155.
58. Barua, A., and Hatzikirou, H. (2023) Cell Decision Making through the Lens of Bayesian Learning. *Entropy* 25, 609.
59. Milo, R., and Phillips, R. *Cell biology by the numbers*; Garland Science, 2015.
60. Buchler, N. E., and Cross, F. R. (2009) Protein sequestration generates a flexible ultrasensitive response in a genetic network. *Molecular systems biology* 5, 272.
61. Qian, Y., and Del Vecchio, D. (2018) Realizing ‘integral control’ in living cells: how to overcome leaky integration due to dilution? *Journal of The Royal Society Interface* 15, 20170902.
62. Cuba Samaniego, C., Wallace, E., Blanchini, F., Franco, E., and Giordano, G. (2024) Neural networks built from enzymatic reactions can operate as linear and nonlinear classifiers. *bioRxiv* 2024–03.
63. Stano, P. (2023) Chemical Systems for Wetware Artificial Life: Selected Perspectives in Synthetic Cell Research. *International Journal of Molecular Sciences* 24, 14138.
64. Stano, P., Gentili, P. L., Damiano, L., and Magarini, M. (2023) A role for bottom-up synthetic cells in the internet of bio-nano things? *Molecules* 28, 5564.
65. Bradley, R. W., Buck, M., and Wang, B. (2016) Tools and principles for microbial gene circuit engineering. *Journal of Molecular Biology* 428, 862–888.

66. Arpino, J. A., Hancock, E. J., Anderson, J., Barahona, M., Stan, G.-B. V., Papachristodoulou, A., and Polizzi, K. (2013) Tuning the dials of synthetic biology. *Microbiology* 159, 1236–1253.
67. Dannenfels, R., Allen, G. M., VanderSluis, B., Koegel, A. K., Levinson, S., Stark, S. R., Yao, V., Tadych, A., Troyanskaya, O. G., and Lim, W. A. (2020) Discriminatory power of combinatorial antigen recognition in cancer T cell therapies. *Cell systems* 11, 215–228.
68. Haselman, M., and Hauck, S. (2009) The future of integrated circuits: A survey of nanoelectronics. *Proceedings of the IEEE* 98, 11–38.
69. Grozinger, L., Amos, M., Goroehowski, T. E., Carbonell, P., Oyarzún, D. A., Stoof, R., Fellermann, H., Zuliani, P., Tas, H., and Goñi-Moreno, A. (2019) Pathways to cellular supremacy in biocomputing. *Nature communications* 10, 5250.