bioRxiv preprint doi: https://doi.org/10.1101/2020.11.10.377077; this version posted November 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Signaling-based neural networks for cellular computation

Christian Cuba Samaniego[1], Andrew Moorman[2], Giulia Giordano[3], and Elisa Franco[1]

***Abstract*— Cellular signaling pathways are responsible for decision making that sustains life. Most signaling pathways include post-translational modification cycles, that process multiple inputs and are tightly interconnected. Here we consider a model for phosphorylation/dephosphorylation cycles, and we show that under some assumptions they can operate as molecular neurons or perceptrons, that generate sigmoidal-like activation functions by processing sums of inputs with positive and negative weights. We carry out a steady-state and structural stability analysis for single molecular perceptrons as well as for feedforward interconnections, concluding that interconnected phosphorylation/dephosphorylation cycles may work as multi-layer biomolecular neural networks (BNNs) with the capacity to perform a variety of computations. As an application, we design signaling networks that behave as linear and non-linear classifiers.**

I. Introduction

Living cells sense, process, and respond to a multitude of inputs from the environment, collectively using this information to make complex, life-sustaining decisions. To integrate and process several inputs at once, cells rely on signaling pathways that regulate downstream pathways depending on the input patterns they detect [1]. An important building block of signaling pathways is given by protein post-translational modification cycles, of which the best known is phosphorylation [2]. The rich dynamics and steady-state responses of these signaling pathway have been amply studied and are well-understood when they are taken as single-input, single-output isolated systems [3]. However, these pathways not only respond to multiple inputs, but they are also tightly interconnected, making it difficult to clearly identify their input-output function and their full potential for signal processing and computation.

In this manuscript, we consider a model for multi-input, single output phosphorylation/dephosphorylation cycles and describe their ability to operate as molecular perceptrons. Building on previous work by us and others, that proposed a strategy to build biomolecular neural networks (BNNs) using molecular sequestration [4], here we demonstrate that, under some assumptions, these natural signaling pathways can operate as multi-input perceptrons as long as their input-output behavior has a tunable threshold. Further, these molecular perceptrons can process several inputs assigning them positive and negative weights by tuning catalytic rate parameters; the presence of both positive and negative weights is essential for complex classification. We report a *structural* steady-state and stability analysis of the multi-input phosphorylation/dephosphorylation cycle model considered as a candidate perceptron, as well as stability analysis for BNNs resulting from the feedforward interconnection of multiple perceptrons. With simulations, we show that by tuning the weights suitably, and by designing the depth of the BNN, it is possible to build linear and non-linear classifiers.

There is a long history of theoretical and experimental work devoted to discovering the computational power of chemical and biological networks under the lens of neural networks. For example, a chemical implementation of neural networks based on reversible reaction was proposed in [5]; networks in living cells were examined in [6], showing that they present features not found in conventional computer-based neural networks; chemical Boltzmann machines have been described in [7]. It has been proposed that transcriptional networks *in vitro* and *in vivo* may operate as neural networks by tuning various reaction rate parameters [8], [9]. Nucleic acids have been engineered for experimental demonstrations of molecular neural networks and classifiers *in vitro* [10], [11], in cell extracts [12], and in cells [13], [11].

Although in practice any biological network may be tuned to work as a neural network [6], the main challenge is to translate an abstract design into an implementation with predictable behavior. A long-standing challenge, for example, is that of implementing negative weights in a biological neural network that can only admit positive variables (concentrations). In previous work, we demonstrated that molecular sequestration reaction can theoretically implement positive and negative weights because it performs a subtraction operation between its inputs [4]. Because a salient feature of molecular sequestration is that it generates an ultrasensitive response with a tunable threshold, we rationalize that other networks with a tunable threshold may implement molecular neural networks with positive and negative weights. This rationale motivates the present study: we show that when post-translational modification cycles, such as phosphorylation/dephosphorylation, present an ultrasensitive response with a tunable threshold, they can operate as perceptron nodes with positive and negative weights. After providing basic definitions in Section II, in Section III we report mathematical analysis of phosphorylation/dephosphorylation cycles as a perceptrons. In section IV we describe the design of a single-layer linear classifiers, Section V describes multi-layer BNNs, and finally Section VI provides simulation

[1]Department of Mechanical and Aerospace Engineering, University of California Los Angeles, California 90095, USA. ccubasam@g.ucla.edu,efranco@seas.ucla.edu

[2]Department of Biological Engineering, Massachusetts Institute of Technology, Massachusetts 02139, USA. amoorman@mit.edu

[3]Department of Industrial Engineering, University of Trento, 38123 Povo (TN), Italy. giulia.giordano@unitn.it

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.10.377077; this version posted November 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

examples of non-linear classifiers.

II. Background: Artificial Neural Networks and Perceptrons

Artificial neural networks (ANNs), also called perceptron networks, take inspiration from the architecture of the human brain. ANNs consist of multiple, interconnected artificial neurons, or perceptrons, which represent the nodes of the network. Each perceptron performs a different computation of the same functional form, whose outcome depends on the perceptron's inputs and parameters. We adopt the following model (Fig. 1A) for a single perceptron that takes n inputs:

$$f\left(\sum_{i=1}^{n} w_i x_i\right),\ x_i, w_i \in \mathbb{R},$$

where f is a non-linear activation function, x_i are the perceptron inputs, $i = 1, ..., n$, and w_i is the parameter or weight associated to the ith input. By connecting multiple perceptrons, each with potentially distinct inputs and weights, ANNs achieve the emergent ability to carry out complex computations far more sophisticated than those realizable by an individual perceptron [14]. In the next sections, we discuss how a well-known class of (post-translational) signaling networks can be viewed as molecular perceptrons.

III. Phosphorylation-dephosphorylation cycles can operate like biomolecular perceptrons

Here we argue that post-translational modification cycles such as phosphorylation-dephosphorylation cycles can operate like biomolecular perceptrons (Fig. 1). Let us denote chemical species with uppercase letters, and their concentrations with the corresponding lowercase letter, so that species X has concentration x.

In a phosphorylation-dephosphorylation cycle (scheme in Fig 1B), a target protein P^* (inactive state, unphosphorylated) is converted into P (active state, phosphorylated) in the presence of a kinase Z, which adds a phosphate group to P^* so as to produce P. A phosphatase Y can reverse this process by removing the phosphate group from protein P, thus producing P^*.

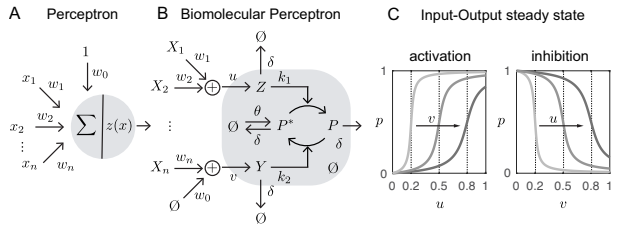

Fig. 1. **Phosphorylation-based perceptron**. A: Perceptron scheme. B: Biomolecular implementation of the perceptron, including at its heart a phosphorylation-dephosphorylation cycle. C: Sigmoidal-like activation and inhibition functions.

We model this process through the following biochemical reactions. The kinase Z, the phosphatase Y and the target protein P^* are produced at a constant rate, respectively u, v and θ:

$$\emptyset \xrightarrow{u} Z \qquad \emptyset \xrightarrow{v} Y \qquad \emptyset \xrightarrow{\theta} P^*.$$

All species are subject to first order degradation with rate constant δ:

$$Z, Y, P, P^*, C_1, C_2 \xrightarrow{\delta} \emptyset.$$

Z and P^* bind with rate constant a_1 to form an intermediate complex C_1, which dissociates at rate constant d_1. The complex C_1 produces P and Z with rate parameter k_1.

$$Z + P^* \underset{d_1}{\overset{a_1}{\rightleftharpoons}} C_1 \xrightarrow{k_1} P + Z.$$

Analogously, Y and P bind at rate constant a_2 to form an intermediate complex C_2, which dissociates with parameter d_2. In addition, the complex C_2 can produce P^* and Y with rate parameter k_2.

$$Y + P \underset{d_2}{\overset{a_2}{\rightleftharpoons}} C_2 \xrightarrow{k_2} P^* + Y.$$

Assuming that the reaction kinetics follow the law of mass action, the dynamic evolution of the species concentrations can be expressed by the following system of Ordinary Differential Equations (ODEs):

$$\begin{aligned}
\dot{z} &= u - a_1 z p^* + (d_1 + k_1) c_1 - \delta z, && (1)\\
\dot{p^*} &= \theta - a_1 z p^* + d_1 c_1 + k_2 c_2 - \delta p^*, && (2)\\
\dot{c}_1 &= a_1 z p^* - (d_1 + k_1) c_1 - \delta c_1, && (3)\\
\dot{y} &= v - a_2 y p + (d_2 + k_2) c_2 - \delta y, && (4)\\
\dot{p} &= -a_2 y p + d_2 c_2 + k_1 c_1 - \delta p, && (5)\\
\dot{c}_2 &= a_2 y p - (d_2 + k_2) c_2 - \delta c_2. && (6)
\end{aligned}$$

We can show that the system is positive, which is consistent with the fact that the state variables are concentrations of chemical species, hence cannot take negative values.

Proposition 1: The system (1)–(6) is positive: given a positive initial condition, all the state variables have nonnegative values at all future times.

Proof: The result follows by noticing that, for each state variable $x \in \{z, p^*, c_1, y, p, c_2\}$, $\dot{x} \geq 0$ when $x = 0$, hence the variable cannot become negative. ■

A. Stability analysis without production and degradation

If we neglect production and degradation reactions ($u = \theta = v = 0$ and $\delta = 0$), assuming that they compensate for each other, then the system simplifies to:

$$\begin{aligned}
\dot{z} &= -a_1 z p^* + (d_1 + k_1) c_1, && (7)\\
\dot{p^*} &= -a_1 z p^* + d_1 c_1 + k_2 c_2, && (8)\\
\dot{c}_1 &= a_1 z p^* - (d_1 + k_1) c_1, && (9)\\
\dot{y} &= -a_2 y p + (d_2 + k_2) c_2, && (10)\\
\dot{p} &= -a_2 y p + d_2 c_2 + k_1 c_1, && (11)\\
\dot{c}_2 &= a_2 y p - (d_2 + k_2) c_2, && (12)
\end{aligned}$$

where the sums $z + c_1$, $y + c_2$ and $p^* + p + c_1 + c_2$ are constant quantities, hence the system can be associated with a reduced-order system with three differential equations and three conservation laws. Since each species is involved at least in one conservation law, this is a *conservative* chemical reaction network. The steady-state value and its global stability can only be assessed within the *stoichiometric*

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.10.377077; this version posted November 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

compatibility class associated with an assigned value of the conserved quantities.

Theorem 1: The system (7)–(12) admits a unique steady state within each stoichiometric compatibility class, and such a steady state is structurally globally asymptotically stable.

Proof: If we set $z = z^{tot} - c_1$ and $y = y^{tot} - c_2$, and neglect the dynamic equations for z and y, from (7)–(12) we get the reduced-order system

$$\begin{aligned}
\dot{p}^* &= -a_1 p^*(z^{tot} - c_1) + d_1 c_1 + k_2 c_2,\\
\dot{c}_1 &= a_1 p^*(z^{tot} - c_1) - (d_1 + k_1)c_1,\\
\dot{p} &= -a_2 p(y^{tot} - c_2) + d_2 c_2 + k_1 c_1,\\
\dot{c}_2 &= a_2 p(y^{tot} - c_2) - (d_2 + k_2)c_2,
\end{aligned}$$

where $p^* + p + c_1 + c_2$ is constant. The Jacobian is

$$J_r = \begin{bmatrix} -\alpha_1 & \beta_1 + d_1 & 0 & k_2 \\ \alpha_1 & -(\beta_1 + d_1 + k_1) & 0 & 0 \\ 0 & k_1 & -\alpha_2 & \beta_2 + d_2 \\ 0 & 0 & \alpha_2 & -(\beta_2 + d_2 + k_2) \end{bmatrix},$$

where $\alpha_1 = a_1(z^{tot} - \bar{c}_1)$, $\alpha_2 = a_2(y^{tot} - \bar{c}_2)$, $\beta_1 = a_1\bar{p}^*$ and $\beta_2 = a_2\bar{p}$. J_r is singular and weakly column diagonally dominant, with strictly negative diagonal entries and non-negative off-diagonal entries. Therefore, the system admits the 1-norm

$$\|[p^* - \bar{p}^* \quad c_1 - \bar{c}_1 \quad p - \bar{p} \quad c_2 - \bar{c}_2]^\top\|_1 \doteq \|x\|_1$$

as a weak Lyapunov function that certifies the *structural* marginal stability of its steady state $(\bar{p}^*, \bar{c}_1, \bar{p}, \bar{c}_2)$ [15], [16, Section 4.5.5], regardless of the values of the system parameters. The only eigenvalue on the imaginary axis is $\lambda = 0$, associated with the left eigenvector $[1, 1, 1, 1]$ that identifies an *invariant* subspace for the system. In this invariant subspace, the system admits a polyhedral Lyapunov function whose unit ball is given by the intersection between the shifted diamond centred at the steady state $(\bar{p}^*, \bar{c}_1, \bar{p}, \bar{c}_2)$ and the hyperplane orthogonal to the vector $[1, 1, 1, 1]$ and passing through $(\bar{p}^*, \bar{c}_1, \bar{p}, \bar{c}_2)$. Since the system Jacobian is structurally nonsingular within the stoichiometric compatibility class, this Lyapunov function guarantees that the system equilibrium is *structurally* asymptotically stable. Note that, since the original (non-linearized) system can be rewritten as $\dot{x} = A(x)x$, where A is related to the integral of the system Jacobian [15], [17], the system admits the same Lyapunov function at all points, hence a global (not only local) stability analysis can be conducted. In particular, in view of the structural nonsingularity of the Jacobian, the results in [17] also guarantee the uniqueness of the steady state and its *structural global asymptotic stability* within the stoichiometric compatibility class. See also [18], where the system (7)–(12) is shown to be *structurally attractive*. ■

It is worth stressing that Theorem 1 does not rely on the mass-action-kinetics assumption: the same result would hold true for any functional form (and parameter values) of the reaction kinetics, as long as the reaction rates are monotonic functions of the reagent concentrations.

B. Quasi-steady-state approximation

In the presence of production and degradation reactions, the quantities $z^{tot} = z + c_1$, $y^{tot} = y + c_2$, and $p^{tot} = p^* + p + c_1 + c_2$ are not constant. From Equations (1)–(6), it follows that these quantities evolve over time as

$$\dot{z}^{tot} = u - \delta z^{tot}, \tag{13}$$
$$\dot{y}^{tot} = v - \delta y^{tot}, \tag{14}$$
$$\dot{p}^{tot} = \theta - \delta p^{tot}. \tag{15}$$

Therefore, at the steady state, $\bar{z} + \bar{c}_1 = \bar{z}^{tot} = u/\delta$, while $\bar{y} + \bar{c}_2 = \bar{y}^{tot} = v/\delta$ and $\bar{p}^* + \bar{p} + \bar{c}_1 + \bar{c}_2 = \bar{p}^{tot} = \theta/\delta$.

Let us adopt a commonly accepted assumption in enzyme kinetics (cf. Michaelis-Menten kinetics) and consider that the concentrations of the intermediate complexes, c_1 and c_2, do not change on the time scale of product formation, thus reaching a quasi steady state:

$$\dot{c}_1 = a_1(z^{tot} - c_1)p^* - d_1 c_1 - k_1 c_1 - \delta c_1 \approx 0, \tag{16}$$
$$\dot{c}_2 = a_2(y^{tot} - c_2)p - d_2 c_2 - k_2 c_2 - \delta c_2 \approx 0. \tag{17}$$

Therefore

$$c_1 \approx \frac{p^*}{p^* + K_1} z^{tot}, \quad \text{with} \quad K_1 = \frac{d_1 + k_1 + \delta}{a_1}, \tag{18}$$

and

$$c_2 \approx \frac{p}{p + K_2} y^{tot}, \quad \text{with} \quad K_2 = \frac{d_2 + k_2 + \delta}{a_2}. \tag{19}$$

If we plug the expression of c_1 and c_2 into the equation for $\dot{p}$, we obtain

$$\dot{p} = -a_2 p \frac{K_2}{p + K_2} y^{tot} + d_2 \frac{p}{p + K_2} y^{tot} + k_1 \frac{p^*}{p^* + K_1} z^{tot} - \delta p,$$

which can rewrite as

$$\dot{p} = -(k_2 + \delta)\frac{p}{p + K_2} y^{tot} + k_1 \frac{p^*}{p^* + K_1} z^{tot} - \delta p. \tag{20}$$

Similarly, the equation for $\dot{p}^*$ can be rewritten as

$$\dot{p}^* = \theta - (k_1 + \delta)\frac{p^*}{p^* + K_1} z^{tot} + k_2 \frac{p}{p + K_2} y^{tot} - \delta p^*. \tag{21}$$

C. Steady-state analysis

To perform the steady-state analysis of the phosphorylation-dephosphorylation network, we make the following simplifying assumptions, which are typically verified in practice. Phosphorylation and dephosphorylation are assumed to be much faster than degradation: $k_1, k_2 \gg \delta$. Also, the total amount of the target protein is assumed to be much larger than the total amount of the enzymes: $p^{tot} \gg z^{tot}, y^{tot}$. Then, since $c_1 < z^{tot} \ll p^{tot}$ and $c_2 < y^{tot} \ll p^{tot}$, we can approximate $p^* + p = p^{tot} - c_1 - c_2 \approx p^{tot}$. Then, by considering $k_1 + \delta \approx k_1$ and $k_2 + \delta \approx k_2$ (since $k_1, k_2 \gg \delta$), we can rewrite equations (20) and (21) as

$$\dot{p} = -k_2 \frac{p}{p + K_2} y^{tot} + k_1 \frac{p^*}{p^* + K_1} z^{tot} - \delta p \tag{22}$$

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.10.377077; this version posted November 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

and

$$\dot{p}^* = \theta - k_1 \frac{p^*}{p^* + K_1} z^{tot} + k_2 \frac{p}{p + K_2} y^{tot} - \delta p^*. \quad (23)$$

Under these approximations, $p + p^* = p^{tot}$ and consistently $\dot{p} + \dot{p}^* = \dot{p}^{tot}$.

Also, under the above approximations, by rearranging equation (22) at steady state (i.e., $\dot{p} = 0$) and neglecting the term δp, we can write the ratio

$$r = \frac{k_1 \bar{z}^{tot}}{k_2 \bar{y}^{tot}} \approx \frac{\bar{p}/\bar{p}^{tot} \left(\hat{K}_1 + 1 - \bar{p}/\bar{p}^{tot} \right)}{\left(\hat{K}_2 + \bar{p}/\bar{p}^{tot} \right) (1 - \bar{p}/\bar{p}^{tot})}, \quad (24)$$

where $\hat{K}_1 = K_1/\bar{p}^{tot}$ and $\hat{K}_2 = K_2/\bar{p}^{tot}$.

In addition, to simplify our analysis, let us assume that $\hat{K} = \hat{K}_1 = \hat{K}_2$. When the steady-state amount of active p is larger than half the total amount, $\bar{p} > \bar{p}^{tot}/2$, the ratio is larger than one: $r > 1$. Conversely, when $\bar{p} < \bar{p}^{tot}/2$, the ratio is smaller than one: $r < 1$. Finally, when $\bar{p} = \bar{p}^{tot}/2$, this results in $r = 1$ (the inputs $k_1 \bar{z}^{tot}$ and $k_2 \bar{y}^{tot}$ are approximately the same in magnitude). In short,

$$\begin{cases} \bar{p} \geq 1/2 & \text{if } r \geq 1 \\ \bar{p} < 1/2 & \text{if } r < 1. \end{cases}$$

We can rewrite this expression as a function of $\bar{z}^{tot}$ and $\bar{y}^{tot}$ as follows:

$$\begin{cases} \bar{p} \geq 1/2 & \text{if } k_1 \bar{z}^{tot} - k_2 \bar{y}^{tot} \geq 0 \\ \bar{p} < 1/2 & \text{if } k_1 \bar{z}^{tot} - k_2 \bar{y}^{tot} < 0 \end{cases} \quad (25)$$

Since at steady state $\bar{z}^{tot} = u/\delta$ and $\bar{y}^{tot} = v/\delta$, we can rewrite the condition as a function of the inputs u and v:

$$\begin{cases} \bar{p} \geq 1/2 & \text{if } k_1 u - k_2 v \geq 0 \\ \bar{p} < 1/2 & \text{if } k_1 u - k_2 v < 0 \end{cases}$$

This mechanism allow us to implement both positive and negative weights, thanks to the fact that the input u appears in the condition with a positive sign, while the input v appears with a negative sign. Then, defining the input $u = \sum_{i=1}^{m} w_i^* x_i$, associated with a positive sign, and the input $v = \sum_{j=1}^{n-m} w_j^* x_j$, associated with a negative sign, the effective input (whose sign determines whether the value of $\bar{p}$ is above or below the threshold $1/2$) becomes

$$\sum_{q=1}^{n} w_q x_q = \sum_{i=1}^{m} (k_1 w_i^*) x_i + \sum_{j=1}^{n-m} (-k_2 w_j^*) x_j,$$

which corresponds to the weighted sum of inputs with both positive and negative coefficients. The coefficients w_k^* can be interpreted biologically as production rates of the enzymes Z and Y. Then, the steady-state expression for multiple inputs can be rewritten as

$$\begin{cases} \bar{p} \geq 1/2 & \text{if } \sum_{q=1}^{n} w_q x_q \geq 0 \\ \bar{p} < 1/2 & \text{if } \sum_{q=1}^{n} w_q x_q < 0 \end{cases} \quad (26)$$

D. Stability analysis

In this section, we show structural stability of the reduced order model including equations (13), (14) and (15), as well as equations (22) and (23). Since $p + p^* = p^{tot}$, we can consider a system with only four equations:

$$\dot{z}^{tot} = u - \delta z^{tot} \quad (27)$$

$$\dot{y}^{tot} = v - \delta y^{tot} \quad (28)$$

$$\dot{p}^{tot} = \theta - \delta p^{tot} \quad (29)$$

$$\dot{p} = -k_2 \frac{p}{p + K_2} y^{tot} + k_1 \frac{p^{tot} - p}{p^{tot} - p + K_1} z^{tot} - \delta p \quad (30)$$

Theorem 2: The system (27)–(30) admits a unique steady state, which is structurally globally asymptotically stable.

Proof: We start by observing that the system trajectories are asymptotically bounded in the open set $\mathcal{S} = \{0 < z^{tot} < z^{MAX}, 0 < y^{tot} < y^{MAX}, 0 < p^{tot} < p^{MAX}, 0 < p < p^{tot}\}$, hence at least one steady state must exist in $\mathcal{S}$. In particular, it can be seen that $\bar{z}^{tot} = u/\delta$, $\bar{y}^{tot} = v/\delta$ and $\bar{p}^{tot} = \theta/\delta$, while $\bar{p} \in (0, \bar{p}^{tot})$ solves the equilibrium condition

$$g(p) \doteq k_2 \frac{p}{p + K_2} \bar{y}^{tot} + \delta p = k_1 \frac{\bar{p}^{tot} - p}{\bar{p}^{tot} - p + K_1} \bar{z}^{tot} \doteq h(p).$$

In the interval $[0, \bar{p}^{tot}]$, $g(p)$ is a continuous strictly increasing function with $g(0) = 0$ and $g(\bar{p}^{tot}) > 0$, while $h(p)$ is a continuous strictly decreasing function with $h(0) > 0$ and $h(\bar{p}^{tot}) = 0$, hence an intersection must exist in view of continuity, and it is unique in view of monotonicity.
Structural local asymptotic stability of the equilibrium can be proven by noticing that the Jacobian matrix

$$J = \begin{bmatrix} -\delta & 0 & 0 & 0 \\ 0 & -\delta & 0 & 0 \\ 0 & 0 & -\delta & 0 \\ f_1 & -f_2 & f_3 & -\delta - f_4 \end{bmatrix}, \quad (31)$$

where the quantities $f_1 = k_1 \frac{p^{tot} - p}{p^{tot} - p + K_1}$, $f_2 = k_2 \frac{p}{p + K_2}$, $f_3 = k_1 \frac{K_1}{(p^{tot} - p + K_1)^2} z^{tot}$ and $f_4 = k_1 \frac{K_1}{(p^{tot} - p + K_1)^2} z^{tot} + k_2 \frac{K_2}{(p + K_2)^2} y^{tot}$ are positive, is a lower triangular matrix, whose eigenvalues are the negative diagonal entries.
Structural asymptotic stability can also be guaranteed by the existence of a polyhedral Lyapunov function [15]: the similarity transformation TJT^{-1}, with $T = \text{diag}(1, 1, 1, \rho)$, turns J into a column diagonally dominant matrix with strictly negative diagonal entries, if ρ is chosen small enough; then, the weighted 1-norm $\|x\|_{1,T} = \|Tx\|_1$, with $x = [z^{tot} - \bar{z}^{tot}, y^{tot} - \bar{y}^{tot}, p^{tot} - \bar{p}^{tot}, p - \bar{p}]^\top$, is a structural Lyapunov function for the system [16, Proposition 4.58]. Finally, since the Jacobian is structurally nonsingular, the results in [17] guarantee the *structural global asymptotic stability* of the equilibrium. ■

Also Theorem 2 would hold for any kinetics (not necessarily mass-action or Michaelis-Menten), as long as monotonicity with respect to species concentrations can be assumed.

Remark 1: Since all the variables of system (27)–(30) are asymptotically bounded in an open set, as shown in the proof of Theorem 2, the following fundamental result from

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.10.377077; this version posted November 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

degree theory can also be applied to prove uniqueness of the equilibrium.

Theorem 3: [19], [20] Assume that the system $\dot{x} = f(x)$, with $f : \mathbb{R}^n \to \mathbb{R}^n$ sufficiently regular, has solutions that are globally uniformly asymptotically bounded in an open set $\mathcal{S}$ and admits $N < \infty$ equilibrium points $\bar{x}^{(i)}$, $i = 1, \ldots, N$, each contained in $\mathcal{S}$ and such that the determinant of the system Jacobian matrix evaluated at $\bar{x}^{(i)}$ is nonzero: $\det(J_f(\bar{x}^{(i)})) \neq 0 \ \forall i$. Then, $\sum_{i=1}^{N} = \text{sign}[\det(-J_f(\bar{x}^{(i)}))] = 1$.
For the Jacobian in Equation (31), $\det(-J) > 0$ structurally; therefore, the equilibrium must be unique.

IV. Tunable Molecular Linear Classifiers

In Section III, we have shown that a reaction network based on post-translational modification by phosphorylation and dephosphorylation may work as a tunable thresholding mechanism (see Fig. 1C). In particular, the threshold can be adjusted by the production of kinase and phosphatase species by positive and negative inputs, respectively, per equation (26). Altogether, this circuit can be viewed as a molecular perceptron with a sigmoidal activation function that incorporates positively- and negatively-weighted inputs.

In this section, we consider two case studies to show how the phosphorylation-dephosphorylation cycle, and by extension any other post-translational modification cycle, may be engineered to build bio-molecular linear classifiers in the presence of multiple inputs, and we discuss how adjusting the weights associated with kinetic rates can predictably program the linear classification.

A. Network with two positive and one negative weights

We start with the design of a linear classifier with three inputs, as shown in Fig. 2A. We consider two inputs X_1 and X_2, which produce the kinase Z with rates w_1 and w_2, as well as a constant production rate w_0 for the phosphatase Y. All other chemical reactions are taken as in Section III, as shown in Fig. 2B. Recall that, in our scheme, input species that produce Z are considered positively-weighted, while input species that produce Y are negatively-weighted. Thus, this reaction network represents a single perceptron with two positive inputs and a single, fixed negative input, or bias.

Using analogous assumptions as in Section III to simplify the model leads to the ODE system:

$$\begin{aligned} \dot{z}^{tot} &= w_1 x_1 + w_2 x_2 - \delta z^{tot} \\ \dot{y}^{tot} &= w_0 - \delta y^{tot} \\ \dot{p}^{tot} &= \theta - \delta p^{tot} \\ \dot{p} &= k_1 \frac{p^{tot} - p}{p^{tot} - p + K_1} z^{tot} - k_2 \frac{p}{p + K_2} y^{tot}. \end{aligned}$$

At the steady state, $\bar{z}^{tot} = \frac{w_1}{\delta} x_1 + \frac{w_2}{\delta} x_2$, $\bar{y}^{tot} = \frac{w_0}{\delta}$, and $\bar{p}^{tot} = \frac{\theta}{\delta}$. We can find $\bar{p}$ as a function of the system inputs by following the same procedure as in Section III, which eventually yields (25) where

$$k_1 \bar{z}^{tot} - k_2 \bar{y}^{tot} = \left(\frac{w_1 k_1}{\delta}\right) x_1 + \left(\frac{w_2 k_1}{\delta}\right) x_2 - \left(\frac{w_0 k_2}{\delta}\right).$$

The terms inside the parenthesis are the effective weights of the system. Fig. 2C-D shows how the linear classification of the system changes when we change the production rates w_1 and w_2: Varying w_1 rescales the x-axis, varying w_2 rescales the y-axis, and varying w_0 rigidly shifts the white region (not shown here).

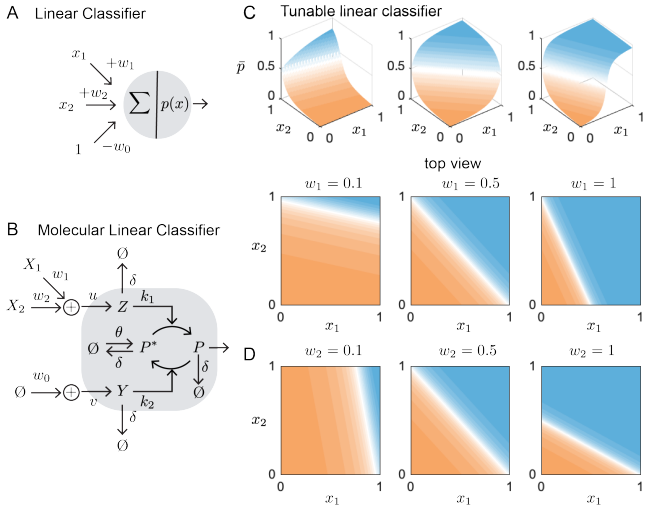

Fig. 2. **Three-input phosphorylation-based linear classifier with two positive weights and one negative weight** A) Design schematic of a single perceptron with three inputs presenting two positive and a single negative weight. B) Phosphorylation-based realization of the perceptron. C) Steady state solutions are shown as heat maps for a range of inputs x_1 and x_2. A top-view of the heat-maps for different values of weights w_1, demonstrating the tunability of the decision boundary (white region). D) The effect of varying w_2 changes the y-axis scale on the decision boundary (top-view heat-maps). Parameter for simulations $w_0 = w_1 = w_2 = 0.5$, $\theta = \delta = 1$, $k_1 = k_2 = 10$, and $K_1 = K_2 = 0.05$.

B. A network with one positive and two negative weights

As a second case study, we consider the design of a linear perceptron with one positive and two negative weights, as shown in Fig. 3A top. Hence, this reaction network features a single input X_2, which produces Z with rate w_2, and two inputs which produce Y – a constant production w_0 and an input X_1 associated with rate w_1 – as shown in Fig. 3A (bottom). This leads to the ODE system:

$$\dot{z}^{tot} = w_2 x_2 - \delta z^{tot} \quad (32)$$
$$\dot{y}^{tot} = w_0 + w_1 x_1 - \delta y^{tot} \quad (33)$$
$$\dot{p}^{tot} = \theta - \delta p^{tot} \quad (34)$$
$$\dot{p} = k_1 \frac{p^{tot} - p}{p^{tot} - p + K_1} z^{tot} - k_2 \frac{p}{p + K_2} y^{tot}. \quad (35)$$

At steady state, $\bar{z}^{tot} = \frac{w_2}{\delta} x_2$, $\bar{y}^{tot} = \frac{w_0}{\delta} + \frac{w_1}{\delta} x_1$, and $\bar{p}^{tot} = \frac{\theta}{\delta}$, while $\bar{p}$ is expressed by (25) with

$$k_1 \bar{z}^{tot} - k_2 \bar{y}^{tot} = \left(\frac{w_2 k_1}{\delta}\right) x_2 - \left(\frac{w_0 k_2}{\delta}\right) - \left(\frac{w_1 k_2}{\delta}\right) x_1.$$

Again, varying the production rates of the system has an equivalent effect on the network's linear classification: w_1 and w_2 rescale the x- and y-axes, respectively (Fig. 3B-C) and w_0 rigidly shifts the white region (not shown here).

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.10.377077; this version posted November 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

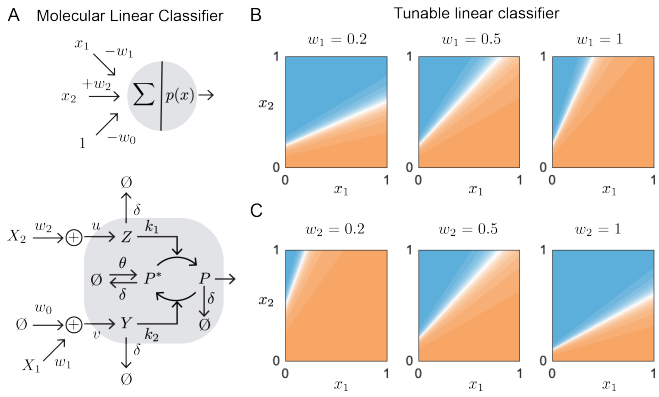

Fig. 3. **Three-input phosphorylation-based linear classifier with one positive and two negative weights** A) Design specification and molecular implementation of a single perceptron with three inputs, one positive and two negative weights. C) Steady state solutions for different values of weights w_1, it results in rescaling the y-axis. D) Varying the production rate w_2 rescale the x-axis. Parameter for simulations $w_0 = w_1 = w_2 = 0.5$, $\theta = \delta = 1$, $k_1 = k_2 = 10$, and $K_1 = K_2 = 0.05$.

In this example, we can observe the importance of incorporating negatively-weighted input values: Changing the sign of w_1 from positive to negative reorients the decision boundary and affords this perceptron greater representation power than the previous example. Specifically, a linear classifier with only positive weights (excluding the bias term), like the perceptron in Fig. 2, is incapable of producing the classification in Fig. 3 without negative weights. In the following sections, we will also demonstrate how tuning positive and negative weights in molecular networks can enable us to engineer more complex decision boundaries when multiple nodes are interconnected.

V. Multi-Layer Biomolecular Neural Networks

Having shown how post-translational modification cycles can operate as a single biomolecular perceptron, we now study networks obtained by cascading many such perceptrons into a multi-layer system. In the following, we denote as *Biomolecular Neural Network* (BNN) a network including multiple perceptrons, that are organized into layers linked only by feedforward connections (cf. Fig. 4). Each perceptron in a layer processes input species $x_{i,j}$ ($i \in \{1, \dots, W\}$, $j \in \{1, \dots, D\}$) to produce output species $y_{i,j}$, where W is the width and D is the depth of the network.

BNNs including layered perceptrons based on molecular sequestration have been studied in [4]. Here, we analyze the stability of the equilibria of a BNN composed of layered phosphorylation-dephosphorylation modules.

Proposition 2: Consider a BNN with depth D and width w_j for each layer j (hence, $W = \sum_j w_j$). Every equilibrium point is locally asymptotically stable.

Proof: The general Jacobian matrix for a feedforward BNN network takes the lower block-triangular form:

$$J_{BNN} = \left[\begin{array}{ccccc|c} L_1 & 0 & \cdots & 0 & \cdots & 0 \\ C_{1,2} & L_2 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ C_{1,j} & C_{2,j} & \cdots & L_j & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \hline C_{1,D} & C_{2,D} & \cdots & C_{j,D} & \cdots & L_D \end{array}\right],$$

where

- matrix $C_{j,k}$, with $j, k \in \{1, \dots, D\}$ and $j < k$, describes how the variables in layer j affect the dynamics of the variables in layer k;
- matrix L_j represents the autonomous dynamics of layer j and has the block-diagonal form

$$L_j = \text{blockdiag}\{P_{j,1}, P_{j,2}, \dots, P_{j,m_j}\},$$

where $P_{i,j}$ denotes the Jacobian matrix of the biomolecular perceptron in position i, j.

Since the Jacobian J_{BNN} is a block-triangular matrix, its characteristic polynomial $\psi_{BNN}(s)$ is the product of the characteristic polynomials of its diagonal blocks. Also, since each diagonal block L_j is block-diagonal, its characteristic polynomial is in turn the product of the characteristic polynomials $\psi_{i,j}(s)$ of the perceptrons $P_{i,j}$ included in the corresponding layer. Therefore,

$$\psi_{BNN}(s) = \prod_{j=1}^{D} \left(\prod_{i=1}^{m_j} \psi_{i,j}(s) \right).$$

Since the generic phosphorylation-dephosphorylation perceptron is structurally asymptotically stable, as shown in Theorem 2, all the eigenvalues of matrix J_{BNN} have negative real part. Therefore, the BNN is asymptotically stable around any equilibrium. ■

VI. Molecular non-linear classifiers

In the previous section we demonstrated the stability of any feedforward BNN that includes a cascade of perceptrons based on post-translational modification cycles. In this section, we use three simulation examples to illustrate

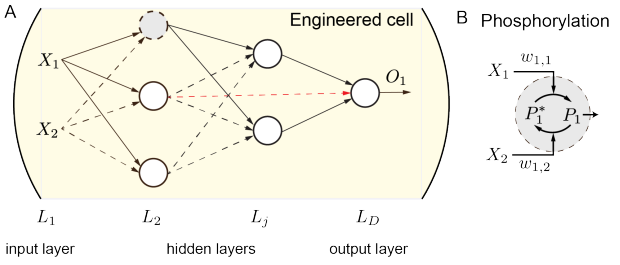

Fig. 4. **Biomolecular Neural Network** A) The schematic diagram for a BNN of depth $D = 4$ with two input species, X_1 and X_2, six biomolecular perceptrons in cicles, and one output species O_1. Black arrows represent feedforward connections between neighboring layers, and red lines, feedforward connections spanning more than one layer. B) A simple biological implementation of each node based on phosphorylation reactions.

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.10.377077; this version posted November 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

how feedforward BNNs may be used to build non-linear classifiers, highlighting their versatility and computational power.

A. XNOR and XOR bio-molecular networks

Exclusive logical or (XOR) classification is a quintessential problem in the field of machine learning. It is the only logical operation which is not linearly separable, and has consequently motivated the use of multi-layer networks: Unlike single perceptrons, networks with two or more layers can perform non-linear classifications (e.g. XOR) provided appropriate weights.

For example, by adjusting its parameter values, the two-layer, three-node network shown in Fig. 5A can implement either XNOR or XOR logical functions. To conserve space, we do not show the detailed chemical reactions for this network, but describe the system with the following ODEs, following steps as given in previous sections:

$$\dot{z}_1^{tot} = w_{1,1}x_1 + w_{1,2}x_2 - \delta z_1^{tot} \quad (36)$$
$$\dot{y}_1^{tot} = w_{1,0} - \delta y_1^{tot} \quad (37)$$
$$\dot{z}_2^{tot} = w_{2,0} - \delta z_2^{tot} \quad (38)$$
$$\dot{y}_2^{tot} = w_{2,1}x_1 + w_{2,2}x_2 - \delta y_2^{tot} \quad (39)$$
$$\dot{z}_3^{tot} = w_{3,1}p_1 + w_{3,2}p_2 - \delta z_3^{tot} \quad (40)$$
$$\dot{y}_3^{tot} = w_{3,0} - \delta y_3^{tot} \quad (41)$$
$$\dot{p}_i^{tot} = \theta - \delta p_i^{tot}, \quad (42)$$
$$\dot{p}_i = k_1 \frac{p_i^{tot} - p_i}{p_i^{tot} - p_i + K} z_i^{tot} - k_2 \frac{p_i}{p_i + K} y_i^{tot}, \quad (43)$$

with $i = 1, 2, 3$.

Notably, the network topology above may carry out either classification by only modifying the bias weights. In Fig. 5, we display the intermediate transfer functions given by perceptrons p_1 and p_2 (Fig. 5B and C, Left and Middle) for each non-linear classifier. To obtain an XNOR classification, we require p_1 and p_2 to have oppositely signed weights applied to x_1 and x_2 – positive weights for p_1 and negative for p_2 – and a larger magnitude of bias for $w_{1,0}$ than $w_{2,0}$ (see Fig. 5B). Under these conditions, the entire network represents an XNOR function; the output of p_3 is maximal when x_1 and x_2 both take on low or high values. In contrast, if we decrease the magnitude of bias $w_{1,0}$ and increase biases $w_{2,0}$ and $w_{3,0}$, the same network topology can perform an XOR operation, which has the opposite output (see Fig. 5C).

This case study illustrates an important general property of BNNs: By changing the weights of each node, the classification boundary, represented by the white area of each plot, can be adjusted in a predictable way and to represent highly dissimilar decision-making functions. Biologically, such changes are tantamount to adjustments of the rate parameters of the network's constituent reactions; no modifications to its underlying topology are required.

B. More complex decision-making

Although the previous classifier is able to deliver a non-linear function of its inputs, its network topology is incapable

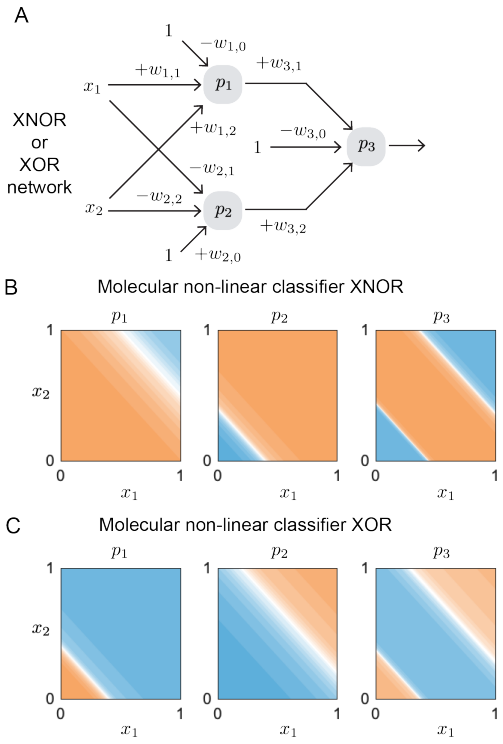

Fig. 5. **Phosphorylation-based XNOR and XOR classifiers** A) Network design. B) Steady state solution for an XNOR with optimized weights ($w_{1,0:2} = [1.5, 1, 1]$, $w_{2,0:2} = [0.4, 1, 1]$, $w_{3,0:2} = [0.3, 1, 1]$). C) Steady state solution for an XOR with optimized weights ($w_{1,0:2} = [0.4, 1, 1]$, $w_{2,0:2} = [1.2, 1, 1]$, $w_{3,0:2} = [1.3, 1, 1]$). The rest of the parameter for both simulations are $\theta = \delta = 1$, $k_1 = k_2 = 10$, and $K = 0.05$.

of creating a closed decision boundary. To achieve this form of classification, we add a perceptron to the first layer of the network, creating a four-node network (Fig. 6). We then optimize the weights of $\bar{p}_1, \bar{p}_2$ and $\bar{p}_3$ to enclose a central region of the classification space, as shown in Fig. 6B. Only where their maximal values (i.e. blue regions) overlap is the threshold value of the fourth node exceeded. This "activated" region forms a closed boundary where the output of the entire network is high. Moreover, by modifying the weights of each node it is possible to shift, scale, and rotate this closed boundary and enclose alternative regions (not shown here). The ODEs describing this network are given below:

$$\dot{z}_1^{tot} = w_{1,1}x_1 + w_{1,2}x_2 - \delta z_1^{tot} \quad (44)$$
$$\dot{y}_1^{tot} = w_{1,0} - \delta y_1^{tot} \quad (45)$$
$$\dot{z}_2^{tot} = w_{2,1}x_1 + w_{2,0} - \delta z_2^{tot} \quad (46)$$
$$\dot{y}_2^{tot} = w_{2,2}x_2 - \delta y_2^{tot} \quad (47)$$
$$\dot{z}_3^{tot} = w_{3,0} + w_{3,2}x_2 - \delta z_3^{tot} \quad (48)$$
$$\dot{y}_3^{tot} = w_{3,1}x_1 - \delta y_3^{tot} \quad (49)$$
$$\dot{z}_4^{tot} = w_{4,1}p_1 + w_{4,2}p_2 + w_{4,3}p_3 - \delta z_4^{tot} \quad (50)$$
$$\dot{y}_4^{tot} = w_{4,0} - \delta y_4^{tot} \quad (51)$$
$$\dot{p}_i^{tot} = \theta - \delta p_i^{tot} \quad (52)$$
$$\dot{p}_i = k_1 \frac{p_i^{tot} - p_i}{p_i^{tot} - p_i + K} z_i^{tot} - k_2 \frac{p_i}{p_i + K} y_i^{tot}, \quad (53)$$

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.10.377077; this version posted November 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

for $i = 1, 2, 3, 4$.

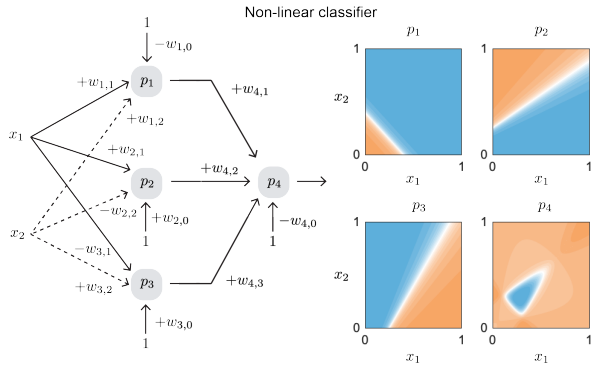

Fig. 6. **Phosphorylation-based non-linear classifiers** Network design for a closed-type decision making with four nodes. Parameter for simulations $w_{1,0:2} = [0.2, 0.5, 0.5]$, $w_{2,0:2} = [0.2, 0.5, 0.8]$, $w_{3,0:2} = [0.2, 0.8, 0.5]$, $w_{4,0:3} = [1.3, 0.5, 0.5, 0.5]$ $\theta = \delta = 1$, $k_1 = k_2 = 10$, and $K = 0.05$.

VII. Discussion

We have discussed how, under some assumptions, a broad class of biomolecular signaling networks have the capacity to operate as molecular perceptrons, due to their sigmoidal input-output behavior with a tunable threshold, that can integrate additively the contribution of multiple inputs. In particular we have shown that the perceptron model arising from this type of network includes inputs that are weighted both positively and negatively. A great advantage of realizing and tuning negative weights is that they make it possible to optimize networks for classification. We illustrated this concept by simulating linear and non-linear classifiers based on phosphorylation/dephosphorylation cycles. This manuscript extends to signaling networks our earlier work that demonstrated how molecular sequestration motifs have the capacity to implement biomolecular neural networks [4]. While here we focus exclusively on demonstrating the computational and classification power of signaling networks, our goal for future work is to build on this approach to examine the contribution of cross-talk to the decision boundary of post-translational modification cycles.

Our analysis complements theoretical and experimental efforts toward building complex molecular networks such as chemical Boltzmann machines [7] and DNA-based classifiers [13], [11]. As molecular neural networks are becoming relevant for applications that include *in vitro* diagnostics [12] and *in vivo* cellular classifiers[13], we expect that methods to harness the computational power and implementation options of BNNs will make it possible to expand the ways we can engineer living cells for advanced decision making.

VIII. Acknowledgments

EF and CCS are sponsored by NSF/BBSRC award 2020039. The content of the information does not necessarily reflect the position or the policy of the Government, and no official endorsement should be inferred.

References

[1] B. E. Housden and N. Perrimon, "Spatial and temporal organization of signaling pathways," *Trends in biochemical sciences*, vol. 39, no. 10, pp. 457–464, 2014.

[2] G. Manning, D. B. Whyte, R. Martinez, T. Hunter, and S. Sudarsanam, "The protein kinase complement of the human genome," *Science*, vol. 298, no. 5600, pp. 1912–1934, 2002.

[3] N. I. Markevich, J. B. Hoek, and B. N. Kholodenko, "Signaling switches and bistability arising from multisite phosphorylation in protein kinase cascades," *The Journal of cell biology*, vol. 164, no. 3, pp. 353–359, 2004.

[4] A. Moorman, C. C. Samaniego, C. Maley, and R. Weiss, "A dynamical biomolecular neural network," in *2019 IEEE 58th Conference on Decision and Control (CDC)*. IEEE, 2019, pp. 1797–1802.

[5] A. Hjelmfelt, E. D. Weinberger, and J. Ross, "Chemical implementation of neural networks and turing machines," *Proceedings of the National Academy of Sciences*, vol. 88, no. 24, pp. 10 983–10 987, 1991.

[6] D. Bray, "Protein molecules as computational elements in living cells," *Nature*, vol. 376, no. 6538, pp. 307–312, 1995.

[7] W. Poole, A. Ortiz-Munoz, A. Behera, N. S. Jones, T. E. Ouldridge, E. Winfree, and M. Gopalkrishnan, "Chemical boltzmann machines," in *International Conference on DNA-Based Computers*. Springer, 2017, pp. 210–231.

[8] N. E. Buchler, U. Gerland, and T. Hwa, "On schemes of combinatorial transcription logic," *Proceedings of the National Academy of Sciences*, vol. 100, no. 9, pp. 5136–5141, 2003.

[9] J. Kim, J. Hopfield, and E. Winfree, "Neural network computation by in vitro transcriptional circuits," in *Advances in neural information processing systems*, 2005, pp. 681–688.

[10] L. Qian, E. Winfree, and J. Bruck, "Neural network computation with dna strand displacement cascades," *Nature*, vol. 475, no. 7356, pp. 368–372, 2011.

[11] K. M. Cherry and L. Qian, "Scaling up molecular pattern recognition with dna-based winner-take-all neural networks," *Nature*, vol. 559, no. 7714, pp. 370–376, 2018.

[12] R. Lopez, R. Wang, and G. Seelig, "A molecular multi-gene classifier for disease diagnostics," *Nature chemistry*, vol. 10, no. 7, pp. 746–754, 2018.

[13] Z. Xie, L. Wroblewska, L. Prochazka, R. Weiss, and Y. Benenson, "Multi-input rnai-based logic circuit for identification of specific cancer cells," *Science*, vol. 333, no. 6047, pp. 1307–1311, 2011.

[14] G. Cybenko, "Approximation by superpositions of a sigmoidal function," *Mathematics of control, signals and systems*, vol. 2, no. 4, pp. 303–314, 1989.

[15] F. Blanchini and G. Giordano, "Piecewise-linear Lyapunov functions for structural stability of biochemical networks," *Automatica*, vol. 50, no. 10, pp. 2482 – 2493, 2014.

[16] F. Blanchini and S. Miani, *Set-theoretic methods in control*. Springer (Second Edition), 2015.

[17] F. Blanchini and G. Giordano, "Polyhedral Lyapunov functions structurally ensure global asymptotic stability of dynamical networks iff the Jacobian is non-singular," *Automatica*, vol. 86, pp. 183 – 191, 2017.

[18] M. Ali Al-Radhawi, D. Angeli, and E. D. Sontag, "A computational framework for a Lyapunov-enabled analysis of biochemical reaction networks," *PLOS Computational Biology*, vol. 16, no. 2, pp. 1–37, 2020.

[19] J. Hofbauer, "An index theorem for dissipative semiflows," *Rocky Mountain J. Math.*, vol. 20, pp. 1017–1031, 1990.

[20] N. G. Lloyd, *Degree Theory*. Cambridge University Press, 1978.