DEVELOPMENT OF A MOLECULAR EXCHANGE MECHANISM-BASED BIOMOLECULAR NEURAL NETWORK

Muhtasim Ishmum Khan*, **Moshiur Rahman*** **& Md. Shahriar Karim**†
Department of Electrical and Computer Engineering
North South University, Dhaka, Bangladesh
{muhtasim.ishmum,moshiur.rahman21,shahriar.karim}@northsouth.edu

ABSTRACT

Molecular computing applied in disease diagnosis, personalized medicine, therapeutics, and other applications often relies on picomolar (pM) to nanomolar (nM) range concentration of interacting species, making noise an inherent part of the computation. While the presence of noise in the inputs may work favorably in an artificial neural network (ANN), in its molecular counterpart, namely the Biomolecular Neural Network (BNN), uncontrolled noise may be detrimental as the decision threshold of a molecular perceptron can deviate from the threshold and result in erroneous classifications. To improve the noise controllability in BNN applications, we develop a multi-layer molecular perceptron network that relies on a molecular exchange-based (MEM) perceptron as its fundamental building block. The underlying Chemical Reaction Network (CRN) module used in BNN here also includes negative feedback of the repressor form for additional control over the threshold dynamics. In addition to ReLU behavior, the proposed MEM-based BNN realizes the XOR operation, demonstrating its potential for linear and nonlinear classification in molecular computation.

1 INTRODUCTION

The biomolecular neural network (BNN) performs as a feature classifier in biophysical and aqueous environments where the classical artificial neural network (ANN) is undeployable. This ability makes BNN a potential solution to designing smart therapeutics, diagnosing different classes of pathogens, personalized medicine, pollutant detection, etc. The versatility of such BNN applications (Fig. 1a) is evident from recent works encompassing applications such as *in vitro* image classification (Arcadia et al., 2021), point-of-care disease diagnosis (Lopez et al., 2018; Zhang et al., 2020), etiological diagnostics (Ma et al., 2022), etc. Generally, a BNN takes concentration $(x_1, x_2, \dots x_n)$ as inputs and performs computation through a chemical reaction network (CRN) module (see Fig. 1b) of kinetic rates arising from the weighted $(w_1, w_2, \dots w_n)$ inputs and the chosen kinetic rate constants for other interaction steps. In BNN, the steady state concentration of species forms a threshold that demarcates between classes. Often, the concentrations are as low as nano and picomolar levels (Plesa et al., 2018), introducing noise to the system and making stochastic analysis of the CRN module of the BNN a critical step. Precisely, fluctuations in CRN arise for numerous reasons that can deviate the steady state concentration from the actual threshold long enough to misclassify the data points. Also, as the molecular perceptron dynamically varies because of parametric dependency, it appears highly sensitive, as evidenced by about 3% error in input reflected by a 10% error in bias (Okumura et al., 2022). These affect faithful classification, making a perceptron unreliable. Thus, designing a molecular perceptron and extending it in a feedforward cascade to form the BNN must address such issues to improve the reliability and reproducibility of classification accuracy for applications demanding high precision.

*Equal contribution
†Corresponding author

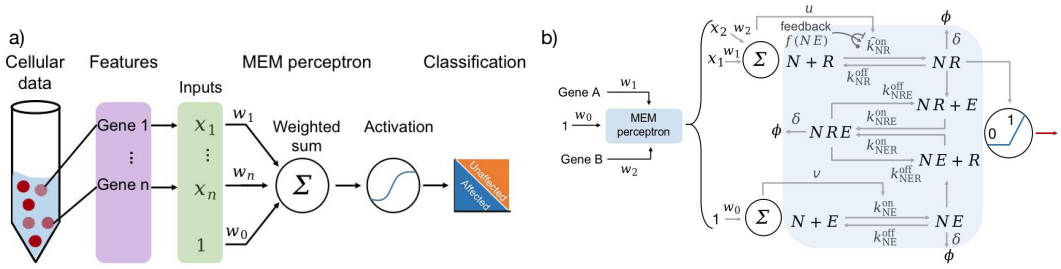

Figure 1: a) Molecular computing applications in diagnostics: Features are fed as inputs to BNN to classify them. b) An elaborated view of two input features with MEM-perceptron design details. NE formation imposes positive (arrowhead), negative (vertical bar) and no (blunt arrow) regulation on the formation rate of NR.

Mostly, the design of a molecular perceptron adopts deterministic approaches that model the reaction dynamics of the interacting species using mass-action kinetics dependent Ordinary Differential Equations (ODE) and form a solid theoretical cue for general computing (Bournez et al., 2017). Following the ODE framework, a few earlier studies used enzymatic interactions, namely molecular sequestration (Moorman et al., 2019) and phosphorylation-dephosphorylation (Samaniego et al., 2021), to design molecular perceptron and extended beyond a single layer design to a multi-layer perceptron (MLP) for nonlinear classification. Other works demonstrating perceptron-like behavior (Banda et al., 2013; Anderson et al., 2021), signal processing filters (Zhang et al., 2022), etc., by molecular computation also adopt a deterministic approach, leaving the role of noise in the context of low-concentration molecular interaction unanalyzed. In addition, perturbations due to ambient temperature variations or input concentration errors that affect the system's decision-making threshold are yet to be studied. As a step to reduce the gap, we developed a molecular exchange mechanism (MEM)-based (Karim et al., 2012) perceptron and extended it to a multilayer BNN achieving nonlinear (Exclusive OR) feature classification. The MEM-based CRN module is relatable to the evolutionarily preserved biophysical interaction steps commonly seen in cellular differentiation and neuronal interactions between neurotransmitters, receptors, and regulators (Papouin et al., 2012; Umulis et al., 2009). Interestingly, the proposed design keeps the threshold-forming species under negative control by the repressor role of the regulator molecule. Overall, the specific contributions we have made in this study are

- Developed a multilayer molecular perceptron network (defined as BNN) performing nonlinear classification as manifested through implementing an Exclusive OR (XOR).
- Improved the MEM-perceptron developed in our previous study and used it as the fundamental building block of the designed BNN.
- *In silico* assessment of the stochastic deviation of the decision boundary between competitive design choices.

2 Models and Results

A generic molecular classifier takes n features to produce a signal as a weighted sum of the gene expression, as schematically shown in Fig. 1a. Each gene expression resembles an input for the BNN, and the binary classification requires weight adjustment for threshold-dependent decision-making. An example of a molecular classifier would be a two-input (x_1, x_2) perceptron shown in Fig. 1b that requires tuning of the three weights (w_0, w_1, w_2) to achieve a perceptron-like behavior, and may be helpful for applications in diagnostics. For instance, the hTERT/GAPDH classifier (Lopez et al., 2018), a two-gene classifying model, which appears similar in summing and activation of the proposed MEM-perceptron shown in Fig. 1b, is applied in cancer diagnostic. In this model, a few reactants and other molecular components are in the order of pM to nM range (1 nM to 30 nM) that in a rectangular volume of dimension $(5 \times 5 \times 0.5\ \mu m^3)$ transforms to 8 to 226 molecular count, making the role of stochastic fluctuations inherent on the reaction dynamics. So, the perceptron design for precision applications, such as disease diagnostics, needs controllability of noise and has been a focus of the proposed MEM-based BNN design.

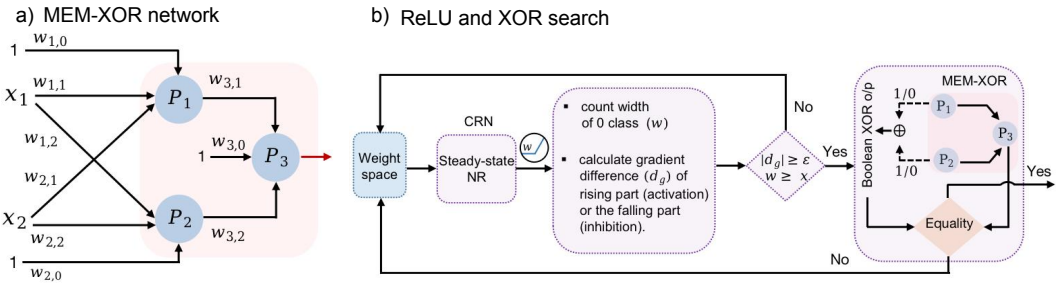

Figure 2: a) Schematic diagram of the MEM-XOR network with P_1, P_2, P_3 representing individual MEM-perceptrons. b) ReLU and XOR search: For an input range $x_1 = [0\ 1]$, ODE simulation of the CRN module generates steady state NR, which produces 0 and 1 classes against a threshold. Identification of ReLU is done through checking the width of the class 0 region and the linearity of class 1 region in perceptron P_i using gradient comparison (d_g). The XOR implementation uses three nodes, P_1, P_2, and P_3, with P_1 and P_2 forming the input layer. A computational screen searches the appropriate weights, ensuring each node P_i is a perceptron for their respective inputs. The comparison between $P_1 \oplus P_2$ and output obtained from P_3 in MEM-XOR completes the automated XOR search process.

2.1 Improved MEM-Perceptron

In our previous work (Rahman et al., 2024), we designed a MEM-based biomolecular perceptron. To identify potential improvement steps, initially we perform an exhaustive computational screen of feedback mechanisms encompassing activation, repressor and none role of the exchanger molecule E (Fig 1b). Specifically, a MEM-perceptron allows regulation of the formation rate (k^{on}_{NR}) of the species NR that acts as the decision threshold. Among the alternatives, a negative control on the forward rate $k^{\text{on}}_{\text{NR}}$ by the exchanger molecule E produces perceptron-like behavior with striking similarities to a smooth-ReLU (Fig. 6b) as defined in Anderson et al. (2021). We also investigate whether feedback alone without forming the NRE complex exhibits ReLU behavior and found NRE formation to be indispensable (see Appendix A.3). Weights (w_0, w_1, w_2) assessment demonstrates that w_0 (bias) variation offsets the simulated decision boundary from the true classification boundary by a traceable deviation Δ. The deviation Δ is dependent on the parameters w_0, n, R_{TOT}, and E through the term $\Omega = [N]/(\beta r[N] + \gamma r[NE])$ with $\beta = f(K, NE, n)$ as in Eq. 3, that defines the threshold condition necessary for perceptron-like behavior (see Appendix A.1, A.2).

2.1.1 Deviation adjustment of the classification boundary

The deviation Δ remains nearly identical regardless of the input combinations for the MEM-perceptron. Such deviation also appeared in molecular sequestration-based perceptron in Moorman et al. (2019) and is amenable in the MEM-perceptron (Fig. 8). Here, we identify ways to compensate the Δ for aligning the *in silico* boundary to the analytically derived one. Precisely, by varying one or a combination of the factors affecting Δ, we identify a functional dependency between the bias w_0 and the ratio $\hat{r} = R_{\text{TOT}}/E$ using a polynomial relation of the following form

$$\hat{r} = 0.2818x^5 - 1.1639x^4 + 1.5789x^3 - 0.6818x^2 + 0.8106x + 1.1743 \tag{1}$$

where, x denotes the bias w_0 and $\hat{r}$ is the ratio R_{TOT}/E. Using Eq. 1, a perceptron with $\Delta = 0$ is achievable as shown in Fig. 8b, given that other parameters confer perceptron behavior. We also relate the range of K and $\hat{r}$ through functional approximation of the dependency relation of both with Δ as in Fig. 4b. The traceability of the deviation Δ is elaborated further in later sections.

2.2 BNN design for XOR operation

A perceptron is impertinent in more complex applications where a nonlinear classifier (for instance, an XOR) is necessary and is a well-known nonlinear classification problem in machine learning (Michalski et al., 2013). Extending on the MEM-perceptron, we design a BNN (the MEM-XOR in Fig. 2a) that performs the logical XOR function (Fig. 3b).

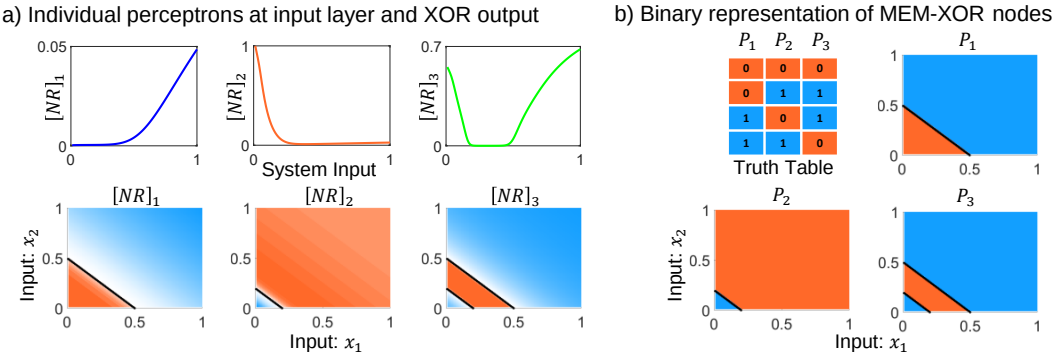

Figure 3: a) Top: Input-output relation for perceptron 1 (P_1), perceptron 2 (P_2), and perceptron 3 (P_3) in an XOR network. P_3 output represents overall XOR concentration output w.r.t the system input. Bottom: Decision boundary formation in exact concentration term for the given feature inputs to be classified. b) The concentration of each perceptron is mapped to class 0 or 1, and the final output (P_3) behaves as a logical XOR between P_1 and P_2 (Algorithm 1).

The designed MEM-XOR network has three nodes (P_1, P_2, P_3) arranged in two layers, as shown in Fig. 2a. All the weights must be tuned together to ensure each node is a perceptron in this feed-forward cascade. Precisely, the first and second nodes of the MEM-XOR network receive system inputs and produce individual decision boundaries for their respective weights. The concentration output of P_1 and P_2 are the inputs for P_3, against which P_3 demonstrates perceptron behavior, and the relevant steps are summarized in Fig. 3. The visualization of the perceptron-like behavior of each node is presented in Fig. 3a (upper panel). The deviation Δ, as observed in each node and the XOR boundary, depends on the bias w_0 and the amount of R_{TOT} and E, and can be eliminated following the dependency relation given in Eq. 1. All the steps summarizing the search of an XOR output are shown in Fig. 2b.

Our analysis identifies Ω as a source of the boundary deviation Δ and depends on multiple parameters (as in Eq. 3) affecting the magnitude of Δ. While the polynomial approximation eliminates Δ for bias w_0 (Fig. 4a) and ratio $\hat{r} = R_{\text{TOT}}/E$ for a fixed K (Eq. 1), we extend further to establish dependence between Δ, K and $\hat{r}$. Through extensive simulation, we develop a range-specific dependency between Δ, K, and $\hat{r}$, capable of eliminating Δ. Interestingly, saturation of R also relates to the deviation Δ by affecting the linearity of the class 1 region. The Δ elimination process requires additional exploration of the Ω dependency to make the BNN design of MEM-perceptrons more scalable and falls under our ongoing extension and analysis.

2.2.1 ODE Model of MEM-XOR Network

Applying the mass-action law for the interaction steps (shown in Fig. 1b) and extending the process for a three-node XOR, we obtain

$$\begin{aligned}
\dot{[NR]}_i &= \hat{k}^{\text{on}}_{\text{NR}_i}[N][R]_i + k^{\text{off}}_{\text{NRE}_i}[NRE]_i - k^{\text{off}}_{\text{NR}_i}[NR]_i - k^{\text{on}}_{\text{NRE}_i}[NR]_i[E]_i - \delta[NR]_i \\
\dot{[NE]}_i &= k^{\text{on}}_{\text{NE}_i}[N][E]_i + k^{\text{off}}_{\text{NER}_i}[NRE]_i - k^{\text{off}}_{\text{NE}_i}[NE]_i - k^{\text{on}}_{\text{NER}_i}[NE]_i[R]_i - \delta[NE]_i \\
\dot{[NRE]}_i &= k^{\text{on}}_{\text{NRE}_i}[NR]_i[E]_i + k^{\text{on}}_{\text{NER}_i}[NE]_i[R]_i - (k^{\text{off}}_{\text{NRE}_i} + k^{\text{off}}_{\text{NER}_i} + \delta)[NRE]_i
\end{aligned} \tag{2}$$

Here, $i = 1, 2, 3$ for the three-node MEM-XOR network, $\dot{[NR]}_i$ denotes $d[NR]_i/dt$ and $[R_{\text{TOT}}]_i = [R]_i + [NR]_i + [NRE]_i$ is the conservation condition of species R for perceptron P_i. The conservation condition mimics many *in vivo* systems where a signaling molecule similar to N binds to a conserved amount of R. However, the MEM-perceptron design is amenable to no conservation on R, maintaining its qualitative role as the ReLU. The threshold conditions of species NR for P_1 and P_2 is expressed as

$$[NR]_i \approx \begin{cases} \dfrac{(\beta_i k^{\text{on}}_{\text{NR}_i} r_i[N] - k^{\text{on}}_{\text{NE}_i}[N] + [NE]_i(\gamma r_i k^{\text{on}}_{\text{NR}_i}))}{\gamma k^{\text{on}}_{\text{NE}_i}} & \text{if } w_{i,1}x_1 + w_{i,2}x_2 \geq \Omega_i w_{i,0} \\ 0 & \text{if } w_{i,1}x_1 + w_{i,2}x_2 < \Omega_i w_{i,0} \end{cases} \tag{3}$$

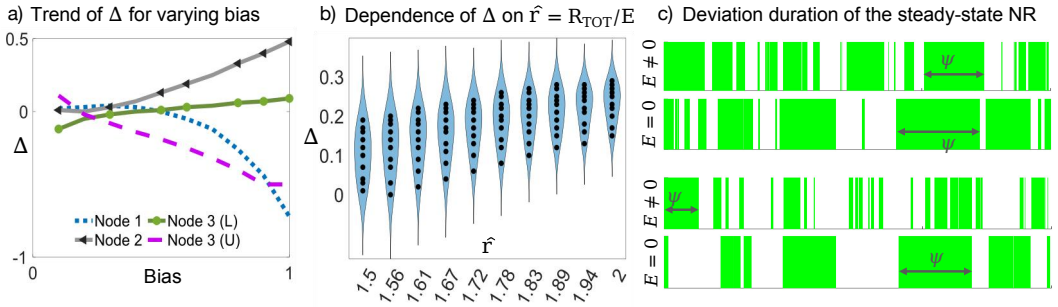

Figure 4: a) The change in Δ for boundaries of each node in the MEM-XOR network, given a fixed parameter set. The comparison is between the logical and the theoretical boundaries, with U and L denoting Δ for the upper and lower boundaries, respectively, for the MEM-XOR. b) Varying K for a fixed ratio $\hat{r} = R_{\text{TOT}}/E$ on the context-dependent Δ elimination. A perceptron-like behavior for different kinetics requires a similar search. c) Deviation assessment of the BNN's CRN module, which is the MEM-perceptron, in the presence and absence of E for percentage deviation greater than 10% (upper panel) and 15% (lower panel). Here, lines in green represent deviation greater than 10%, whereas the white region represents threshold concentration is less than 10%.

P_3 takes the outputs of P_1 and P_2 as input, and its threshold condition is expressed as

$$[NR]_3 \approx \begin{cases} \dfrac{(\beta_3 k^{\text{on}}_{\text{NR}_3} r_3[N] - k^{\text{on}}_{\text{NE}_3}[N] + [NE]_3(\gamma r_3 k^{\text{on}}_{\text{NR}_3}))}{\gamma k^{\text{on}}_{\text{NE}_3}} & \text{if } w_{3,1}[NR]_1 + w_{3,2}[NR]_2 \geq \Omega_3 w_{3,0} \\ 0 & \text{if } w_{3,1}[NR]_1 + w_{3,2}[NR]_2 < \Omega_3 w_{3,0} \end{cases} \tag{4}$$

2.3 Deviation reduction of the threshold in MEM-perceptron

Maintaining the steady state NR concentration around the threshold is essential for reliable ReLU behavior. For instance, erroneous classification due to stochastic fluctuations is evident in Fig. 5a for a molecular sequestration-based ReLU (Moorman et al., 2019). How the MEM-perceptron responds to stochastic fluctuations is assessed by investigating its ability to reduce the deviation duration from the mean of steady state NR. We use the Chemical Master Equation (CME) (Van Kampen, 1992) of the CRN module of the MEM-perceptron and approximate it using the Gillespie's SSA (Gillespie, 1976) available in the GillesPy2 (Matthew et al., 2023) to study the stochastic variability of steady state NR. The CME describes the time evolution of the probability that the system state is m at time t, where m represents the molecular count of species $s = \{s_1, s_2, \ldots s_N\}$ with $m = \{m_1, m_2, \ldots m_N\}$ for R reactions $r = \{r_1, r_2 \ldots r_R\}$

$$\frac{d}{dt}p(m,t) = \sum_i^R a_i(m - \nu_i, t)p(m - \nu_i, t) - \sum_i^R a_i(m,t)p(m,t) \tag{5}$$

where ν_i is the stoichiometric vector that tracks molecular count change of species in every reaction and forms the stoichiometric matrix $\nu = [\nu_1, \nu_2 \ldots \nu_R]^T$. The fidelity quantification of the target threshold NR in the presence ($E \neq 0$) and absence ($E = 0$) of exchanger is defined as

$$\text{percentage-deviation} = \left| \frac{[NR]_{\text{sample}} - \text{mean level of } [NR]}{\text{mean level of } [NR]} \right| \times 100 \tag{6}$$

To assess the MEM's role in noise reduction, in Fig. 4c, we analyze how lengthy the deviation duration is in the absence of the exchanger-based mechanism. Interestingly, the width (ψ) of the most extended deviation duration is comparatively large in the absence of E, which holds equally for the percentage-deviation = 10 (Fig. 4c, upper panel) and percentage-deviation = 15 (Fig. 4c, lower panel), respectively. This suggests that MEM can keep the decision threshold around the mean and may eventually increase the fidelity of a threshold-dependent decision boundary in stochastic fluctuations. Besides, in a preliminary comparison of MEM to alternative perceptron design methods

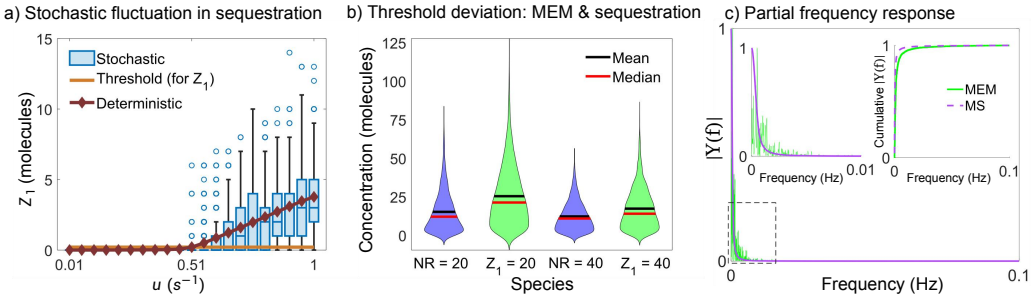

Figure 5: a) Impact of stochasticity on the ReLU behavior, marked by class 1 turning to class 0 due to fluctuating Z_1 count. Here, the yellow line is the threshold producing the binary classes. b) Preliminary comparison on noise controlling ability of the MEM-based approach; for two target thresholds, 20 and 40, a large deviation (for 6 hours of steady state data) from the mean is less probable for the MEM-based approach (blue) in comparison to the molecular sequestration approach (green). c) Fourier Transform of steady state NR and Z_1 molecular count (= 40) of the MEM-perceptron and molecular sequestration, respectively. The presence of $|Y(f)|$ in high-frequency regions is more in the MEM-perceptron (inset, left, green). The cumulative progression of $|Y(f)|$ also approaches gradually to 1, indicating the presence of Fourier spectra in the high-frequency region. Here, $|Y(f)|$ is normalized by their respective highest magnitude.

such as the molecular sequestration (Moorman et al., 2019), the proposed MEM-based approach exhibits a comparatively lower percentage-deviation (Fig. 5b).

Further comparison using the Fourier Transform of the steady state threshold suggests that noise in MEM-perceptron includes more high-frequency components than molecular sequestration. Precisely, the magnitude of the spectrum amplitude ($|Y(f)|$) (as in Fig. 5c) and the cumulative progression of $|Y(f)|$ (inset, right of Fig. 5c) both represent a more high-frequency component in the fluctuations of the threshold species in MEM-perceptron. Intriguingly, in biological neurons (Stein et al., 2005), a signal contaminated by low-frequency noise may lack sufficient information to retrieve the input signal. In contrast, noise becomes beneficial when the noise frequency in the signal is high. Increasing the frequency of fluctuations of steady state concentration, as observed for the MEM-perceptron, may transform noise to be partially beneficial. Our ongoing work capitalizes on these preliminary findings and intends for an exhaustive study of the kinetic dependence of the noise controllability of the MEM-perceptron.

3 Discussion

Nonlinear classification requires a multilayer perceptron implementation, which is often challenging due to molecular systems' high sensitivity to the species concentration and the underlying kinetics. We successfully implemented a MEM-XOR by extending an improved version of our previously designed MEM-perceptron into feed-forward cascades. The decision boundary deviation in MEM-XOR is removable, often following a range-dependent relation among the parameters. The exact role of the MEM in the perceptron is also substantiated through an ablation study. However, the efficacy of the MEM-perceptron for traditional training data (Deng, 2012), or of other forms (Banda et al., 2013), still needs to be tested. Another potential advantage of MEM-based design is its ability to curb noise to increase the fidelity of the classification threshold. While noise is detrimental to maintaining a threshold at a fixed value, on the contrary, noise injection in neural network training regularizes its performance (Bishop, 1995). So, in an inherently noisy system, tuning noise to a level beneficial for molecular computation and imposing greater control over it, thus, becomes necessary. The introduced MEM is known for reducing stochastic noise (Karim et al., 2012) and achieving greater noise control. Also, as added in the perceptron design for BNN, a negative feedback loop is expected to work against such perturbations (Dublanche et al., 2006; Singh, 2011). Together, the developed BNN achieves both linear and nonlinear classification, suggesting it as a potential alternative for the design of low-concentration molecular classifiers. Besides our ongoing analysis, we also consider extrinsic noise analysis (for instance, because of temperature fluctuations) to identify the necessary adjustment in the circuitry that absorbs such extrinsic variations (Appendix A.6).

REFERENCES

David F Anderson, Badal Joshi, and Abhishek Deshpande. On reaction network implementations of neural networks. *Journal of the Royal Society Interface*, 18(177):20210031, 2021.

Christopher E Arcadia, Amanda Dombroski, Kady Oakley, Shui Ling Chen, Hokchhay Tann, Christopher Rose, Eunsuk Kim, Sherief Reda, Brenda M Rubenstein, and Jacob K Rosenstein. Leveraging autocatalytic reactions for chemical domain image classification. *Chemical Science*, 12(15):5464–5472, 2021.

Peter Banda, Christof Teuscher, and Matthew R Lakin. Online learning in a chemical perceptron. *Artificial life*, 19(2):195–219, 2013.

Chris M Bishop. Training with noise is equivalent to tikhonov regularization. *Neural computation*, 7(1):108–116, 1995.

Olivier Bournez, Daniel S Graça, and Amaury Pouly. Polynomial time corresponds to solutions of polynomial ordinary differential equations of polynomial length. *Journal of the ACM (JACM)*, 64 (6):1–76, 2017.

Li Deng. The mnist database of handwritten digit images for machine learning research [best of the web]. *IEEE signal processing magazine*, 29(6):141–142, 2012.

Yann Dublanche, Konstantinos Michalodimitrakis, Nico Kümmerer, Mathilde Foglierini, and Luis Serrano. Noise in transcription negative feedback loops: simulation and experimental analysis. *Molecular systems biology*, 2(1):41, 2006.

Daniel T Gillespie. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *Journal of computational physics*, 22(4):403–434, 1976.

Mohammad Shahriar Karim, Gregery T Buzzard, and David M Umulis. Secreted, receptor-associated bone morphogenetic protein regulators reduce stochastic noise intrinsic to many extracellular morphogen distributions. *Journal of The Royal Society Interface*, 9(70):1073–1083, 2012.

Randolph Lopez, Ruofan Wang, and Georg Seelig. A molecular multi-gene classifier for disease diagnostics. *Nature chemistry*, 10(7):746–754, 2018.

Qian Ma, Mingzhi Zhang, Chao Zhang, Xiaoyan Teng, Linlin Yang, Yuan Tian, Junyan Wang, Da Han, and Weihong Tan. An automated dna computing platform for rapid etiological diagnostics. *Science Advances*, 8(47):eade0453, 2022.

Sean Matthew, Fin Carter, Joshua Cooper, Matthew Dippel, Ethan Green, Samuel Hodges, Mason Kidwell, Dalton Nickerson, Bryan Rumsey, Jesse Reeve, et al. Gillespy2: A biochemical modeling framework for simulation driven biological discovery. *Letters in biomathematics*, 10(1):87, 2023.

Ryszard Stanislaw Michalski, Jaime Guillermo Carbonell, and Tom M Mitchell. *Machine learning: An artificial intelligence approach*. Springer Science & Business Media, 2013.

Andrew Moorman, Christian Cuba Samaniego, Carlo Maley, and Ron Weiss. A dynamical biomolecular neural network. In *2019 IEEE 58th conference on decision and control (CDC)*, pp. 1797–1802. IEEE, 2019.

S Okumura, G Gines, N Lobato-Dauzier, A Baccouche, R Deteix, T Fujii, Y Rondelez, and AJ Genot. Nonlinear decision-making with enzymatic neural networks. *Nature*, 610(7932):496–501, 2022.

Thomas Papouin, Laurent Ladépêche, Jérôme Ruel, Silvia Sacchi, Marilyne Labasque, Marwa Hanini, Laurent Groc, Loredano Pollegioni, Jean-Pierre Mothet, and Stéphane HR Oliet. Synaptic and extrasynaptic nmda receptors are gated by different endogenous coagonists. *Cell*, 150(3): 633–646, 2012.

Tomislav Plesa, Konstantinos C Zygalakis, David F Anderson, and Radek Erban. Noise control for molecular computing. *Journal of the Royal Society Interface*, 15(144):20180199, 2018.

Moshiur Rahman, Muhtasim Ishmum Khan, and Md. Shahriar Karim. Design of a molecular exchange-based robust perceptron for biomolecular neural network. In *The Second Tiny Papers Track at ICLR 2024*, 2024. URL https://openreview.net/forum?id=AXwGBliKOV.

Christian Cuba Samaniego, Andrew Moorman, Giulia Giordano, and Elisa Franco. Signaling-based neural networks for cellular computation. In *2021 American Control Conference (ACC)*, pp. 1883–1890. IEEE, 2021.

Abhyudai Singh. Negative feedback through mrna provides the best control of gene-expression noise. *IEEE transactions on nanobioscience*, 10(3):194–200, 2011.

Richard B Stein, E Roderich Gossen, and Kelvin E Jones. Neuronal variability: noise or part of the signal? *Nature Reviews Neuroscience*, 6(5):389–397, 2005.

David Umulis, Michael B O'Connor, and Seth S Blair. The extracellular regulation of bone morphogenetic protein signaling. *Development*, 136(22):3715–3728, 2009.

Nicolaas Godfried Van Kampen. *Stochastic processes in physics and chemistry*, volume 1. Elsevier, 1992.

Chao Zhang, Yumeng Zhao, Xuemei Xu, Rui Xu, Haowen Li, Xiaoyan Teng, Yuzhen Du, Yanyan Miao, Hsiao-chu Lin, and Da Han. Cancer diagnosis with dna molecular computation. *Nature nanotechnology*, 15(8):709–715, 2020.

Yichi Zhang, Christian Cuba Samaniego, Katelyn Carleton, Yili Qian, Giulia Giordano, and Elisa Franco. Building molecular band-pass filters via molecular sequestration. In *2022 IEEE 61st Conference on Decision and Control (CDC)*, pp. 3890–3895. IEEE, 2022.

A Appendix

A.1 MEM Chemical Reaction Network

There are three chemical species working in the MEM: N (signaling agent, which could be a neurotransmitter), R (receptor), and E (exchanger molecule). As molecular production adds noises, the designed dynamics take N and E as constant sources to establish better control over noise. A conservation condition for R is maintained in the continuum. The chemical reaction network (CRN) of MEM is as follows:

$$\text{r1: } N + R \underset{k^{\text{off}}_{\text{NR}}}{\overset{\hat{k}^{\text{on}}_{\text{NR}}}{\rightleftharpoons}} NR, \quad \text{r2: } N + E \underset{k^{\text{off}}_{\text{NE}}}{\overset{k^{\text{on}}_{\text{NE}}}{\rightleftharpoons}} NE$$

$$\text{r3: } NR + E \underset{k^{\text{off}}_{\text{NRE}}}{\overset{k^{\text{on}}_{\text{NRE}}}{\rightleftharpoons}} NRE, \quad \text{r4: } NE + R \underset{k^{\text{off}}_{\text{NER}}}{\overset{k^{\text{on}}_{\text{NER}}}{\rightleftharpoons}} NRE$$

$$\text{r5: } NR, NE, NRE \xrightarrow{\delta} \emptyset$$

The following conservation condition is maintained on the receptor R:

$$[R_{\text{TOT}}] = [R] + [NR] + [NRE] \tag{7}$$

Upon formation, NE negatively regulates the formation rate of NR ($k^{\text{on}}_{\text{NR}}$) by the Hill Equation

$$\hat{k}^{\text{on}}_{\text{NR}} = k^{\text{on}}_{\text{NR}} \left(\frac{K^{\text{n}}}{K^{\text{n}} + [NE]^{\text{n}}} \right) = f\left(k^{\text{on}}_{\text{NR}}\right) \tag{8}$$

where $k^{\text{on}}_{\text{NR}}$ is the basal NR production rate.

The CRN is modeled using mass-action-based ODEs as follows:

$$\begin{aligned}
[\dot{NR}] &= \hat{k}^{\text{on}}_{\text{NR}}[N][R] + k^{\text{off}}_{\text{NRE}}[NRE] - k^{\text{off}}_{\text{NR}}[NR] - k^{\text{on}}_{\text{NRE}}[NR][E] - \delta[NR] \\
[\dot{NE}] &= k^{\text{on}}_{\text{NE}}[N][E] + k^{\text{off}}_{\text{NER}}[NRE] - k^{\text{off}}_{\text{NE}}[NE] - k^{\text{on}}_{\text{NER}}[NE][R] - \delta[NE] \\
[\dot{NRE}] &= k^{\text{on}}_{\text{NRE}}[NR][E] + k^{\text{on}}_{\text{NER}}[NE][R] - k^{\text{off}}_{\text{NRE}}[NRE] - k^{\text{off}}_{\text{NER}}[NRE] - \delta[NRE]
\end{aligned} \tag{9}$$

A.2 PERCEPTRON-LIKE BEHAVIOR OF THE MEM

At steady state, $\dot{[NR]} = \dot{[NE]} = 0$. We can ignore the term δ in steady state. We assume that $k^{\text{off}}_{\text{NRE}} = k^{\text{off}}_{\text{NER}}$, $k^{\text{on}}_{\text{NRE}} = \gamma k^{\text{on}}_{\text{NE}}$, $k^{\text{on}}_{\text{NER}} = \gamma k^{\text{on}}_{\text{NR}}$, and $r = [R]/[E]$. From Eq. 9,

$$\hat{k}^{\text{on}}_{\text{NR}}[N][R] + k^{\text{off}}_{\text{NRE}}[NRE] - k^{\text{off}}_{\text{NR}}[NR] - k^{\text{on}}_{\text{NRE}}[NR][E] = k^{\text{on}}_{\text{NE}}[N][E] + k^{\text{off}}_{\text{NER}}[NRE] - k^{\text{off}}_{\text{NE}}[NE] - k^{\text{on}}_{\text{NER}}[NE][R]$$

Eliminating $k^{\text{off}}_{\text{NRE}}[NRE]$ and $k^{\text{off}}_{\text{NER}}[NRE]$, and applying feedback expression of $\hat{k}^{\text{on}}_{\text{NR}}$, we get,

$$\begin{aligned}\frac{K^{\text{n}} k^{\text{on}}_{\text{NR}}}{K^{\text{n}} + [NE]^{\text{n}}}[N][R] - k^{\text{on}}_{\text{NE}}[N][E] &= k^{\text{off}}_{\text{NR}}[NR] + k^{\text{on}}_{\text{NRE}}[NR][E] - k^{\text{off}}_{\text{NE}}[NE] - k^{\text{on}}_{\text{NER}}[NE][R] \\ &= [NR](k^{\text{off}}_{\text{NR}} + \gamma k^{\text{on}}_{\text{NE}}[E]) - [NE](k^{\text{off}}_{\text{NE}} + \gamma k^{\text{on}}_{\text{NR}}[R])\end{aligned}$$

Making $[NR]$ the subject, and denoting the feedback expression as $K^{\text{n}}/(K^{\text{n}} + [NE]^{\text{n}}) = \beta$

$$\begin{aligned}[NR] &= \frac{\beta k^{\text{on}}_{\text{NR}}[N][R] - k^{\text{on}}_{\text{NE}}[N][E] + [NE](k^{\text{off}}_{\text{NE}} + \gamma k^{\text{on}}_{\text{NR}}[R])}{k^{\text{off}}_{\text{NR}} + \gamma k^{\text{on}}_{\text{NE}}[E]} \\ &= \frac{\beta k^{\text{on}}_{\text{NR}} r[N] - k^{\text{on}}_{\text{NE}}[N] + [NE](\frac{k^{\text{off}}_{\text{NE}}}{[E]} + \gamma r k^{\text{on}}_{\text{NR}})}{\frac{k^{\text{off}}_{\text{NR}}}{[E]} + \gamma k^{\text{on}}_{\text{NE}}}\end{aligned}$$

Assuming $\gamma k^{\text{on}}_{\text{NE}} \gg k^{\text{off}}_{\text{NR}}/[E]$ and $\gamma r k^{\text{on}}_{\text{NR}} \gg k^{\text{off}}_{\text{NE}}/[E]$, we finally get,

$$[NR] \approx \frac{\beta k^{\text{on}}_{\text{NR}} r[N] - k^{\text{on}}_{\text{NE}}[N] + [NE](\gamma r k^{\text{on}}_{\text{NR}})}{\gamma k^{\text{on}}_{\text{NE}}} \tag{10}$$

Here, $[NR] \geq 0$ as concentration is always non-negative. We can write,

$$\begin{aligned}\frac{\beta k^{\text{on}}_{\text{NR}} r[N] - k^{\text{on}}_{\text{NE}}[N] + [NE](\gamma r k^{\text{on}}_{\text{NR}})}{\gamma k^{\text{on}}_{\text{NE}}} &\geq 0 \\ k^{\text{on}}_{\text{NR}} &\geq \underbrace{\left(\frac{[N]}{\beta r[N] + \gamma r[NE]}\right)}_{\Omega} k^{\text{on}}_{\text{NE}} \\ k^{\text{on}}_{\text{NR}} &\geq \Omega k^{\text{on}}_{\text{NE}} \qquad (11)\end{aligned}$$

Hence, we get the threshold condition:

$$[NR] \approx \begin{cases} (\beta k^{\text{on}}_{\text{NR}} r[N] - k^{\text{on}}_{\text{NE}}[N] + [NE](\gamma r k^{\text{on}}_{\text{NR}}))/\gamma k^{\text{on}}_{\text{NE}} & \text{if } k^{\text{on}}_{\text{NR}} \geq \Omega k^{\text{on}}_{\text{NE}} \\ 0 & \text{if } k^{\text{on}}_{\text{NR}} < \Omega k^{\text{on}}_{\text{NE}} \end{cases} \tag{12}$$

We can rewrite the above condition in terms of perceptron inputs and weights as follows:

$$[NR] \approx \begin{cases} (\beta k^{\text{on}}_{\text{NR}} r[N] - k^{\text{on}}_{\text{NE}}[N] + [NE](\gamma r k^{\text{on}}_{\text{NR}}))/\gamma k^{\text{on}}_{\text{NE}} & \text{if } w_1 x_1 + w_2 x_2 \geq \Omega w_0 \\ 0 & \text{if } w_1 x_1 + w_2 x_2 < \Omega w_0 \end{cases} \tag{13}$$

Transformation of concentration values to binary:

$$[NR] = \begin{cases} 1 & \text{if } w_1 x_1 + w_2 x_2 \geq \Omega w_0 \\ 0 & \text{if } w_1 x_1 + w_2 x_2 < \Omega w_0 \end{cases} \tag{14}$$

A.3 RELU AND XOR SEARCH STEPS

The strategy to detect ReLU-like behavior in the species NR in steady state is as follows:

- The simulated steady state concentration values of NR are transformed to binary using Eq. 14 and the number of 0s are recorded.

- Linearity of the rising (for activation) or falling (for inhibition) region of the function is analyzed by calculating the difference between gradients calculated at two segments of the rising part and ensuring the absolute difference is within a threshold.
- For a number of 0s greater than the specific number determined by the experimenter and difference of gradients being within the fixed threshold, the steady state behavior of NR is flagged as a ReLU.

Next, the process to verify the implementation of the XOR function is given. We assume that the term Ω obtains such values at steady state that there is no deviation Δ for the output of each node.

- The steady state concentrations of NR of each perceptron node is converted to binary using Eq. 14.
- The output of a logical XOR operation between the binary values of P_1 and P_2 is calculated and saved.
- The XOR data and the binary values of steady state concentrations of NR of P_3 are compared, raising a positive flag for being identical and a negative flag otherwise.

The following algorithm is applied to every data point of the MEM-XOR network to verify XOR behavior. In our simulations, we generated 100×100 matrices of concentration data for each node and applied the algorithm to determine XOR behavior.

Algorithm 1 VerifyXOR ($k^{\text{on}}_{\text{NR}_1}$, $k^{\text{on}}_{\text{NR}_2}$, $k^{\text{on}}_{\text{NR}_3}$, $k^{\text{on}}_{\text{NE}_1}$, $k^{\text{on}}_{\text{NE}_2}$, $k^{\text{on}}_{\text{NE}_3}$, Ω_1, Ω_2, Ω_3)

if $k^{\text{on}}_{\text{NR}_1} \geq \Omega_1 k^{\text{on}}_{\text{NE}_1}$ **then**
  $n1 = 1$ ▷ Transform first node output to binary
end if
if $k^{\text{on}}_{\text{NR}_2} \geq \Omega_2 k^{\text{on}}_{\text{NE}_2}$ **then**
  $n2 = 1$ ▷ Transform second node output to binary
end if
if $k^{\text{on}}_{\text{NR}_3} \geq \Omega_3 k^{\text{on}}_{\text{NE}_3}$ **then**
  $n3 = 1$ ▷ Transform third node output to binary
end if
if $n1 \sim= n2$ **then**
  $xor = 1$ ▷ First node XOR second node
end if
if $xor == n3$ **then** ▷ Verify logical XOR output from the third node
  return true
else
  return false

A.4 Ablation study of the role of exchanger in perceptron-like behavior

The imposition of negative feedback on the formation rate of NR is essential for demonstrating perceptron-like behavior by the MEM, as we found from a parametric screen (Table. 1). In the exchange mechanism, NE negatively regulates the formation of NR. However, the species NRE has no direct participation in this feedback dynamic, and so, to gauge the exact importance of NRE, we design a two-node network with NR and NE. Each node regulates its formation and that of the other node, leading to 81 different topologies for exploration using the positive, negative, and no regulation impositions and a parameter screen of $186,624$ combinations for each topology. Interestingly, no such combination was observed that led to the formation of a perceptron, using NR as the decision-making species as in the MEM (Fig. 6). The evidence suggests that the role of NRE is essential for the MEM to work as a perceptron and that only negative feedback is likely to provide a perceptron for the current design of the MEM.

A.5 XOR network analysis

The decision boundaries generated to reflect nonlinear classification of inputs (XOR in this case) should resemble the boundary lines generated by the nodes which are given as input to the final node.

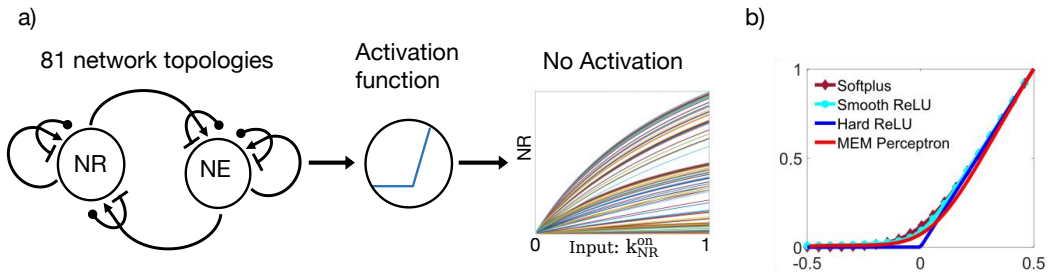

Figure 6: a) All regulations imposed in the two-node network and an example of the resultant functions obtained. b) Comparison of different ReLU functions with the smooth ReLU obtained by the MEM perceptron.

Table 1: ReLU search with positive, negative and no feedback

Feedback type	Parameter sets	Number of ReLU	Rate (%)
Negative	26244	75	0.28
Positive	26244	0	0
None	6561	0	0

For each node, changing the bias causes a deviation (Δ) in the decision boundary line. Similarly, the bias of the output node leads to a simultaneous deviation of both the decision boundaries generated by the output node. However, the XOR behavior is observed for specific combinations of bias inputs with the existence of the deviation Δ in many cases.

Therefore, a search of nodal biases was conducted by varying R_{TOT} and E for different values of the bias (w_0) of each node (Table. 2). A total of $15,625$ combinations were analyzed and among them, 10 combinations demonstrated XOR behavior with minimized deviation Δ. Furthermore, to verify that the designed MLP works as a nonlinear classifier for various parameter combinations, different variations were introduced in the parameters and the XOR classification was obtained for a variety of test cases. Afterwards, the Hill coefficient (n) and the NE concentration producing half occupation (K) were varied as well to study its impact on nonlinear decision-making i.e. implementation of the XOR function. A sample of MEM-XOR with different parameter combinations is shown in Fig. 7.

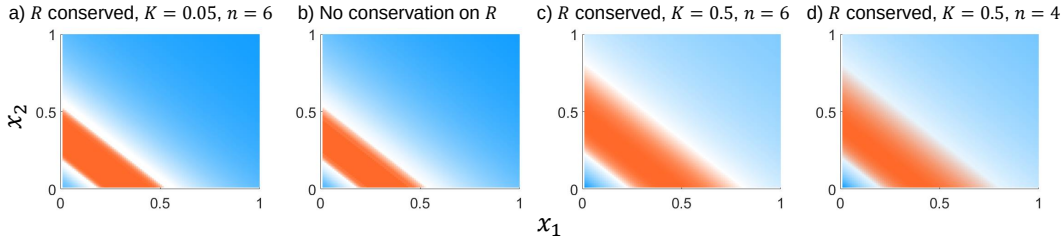

Figure 7: Surface plots showing XOR classification (without deviation Δ) for different parameter configurations. The weights for P_1, P_2 and P_3 are: a,b) $[0.5,\ 1,\ 1]$, $[0.2,\ 1,\ 1]$, $[0.1,\ 25,\ 1]$. c,d) $[0.8,\ 1,\ 1]$, $[0.2,\ 1,\ 1]$, $[0.52,\ 25,\ 0.8]$.

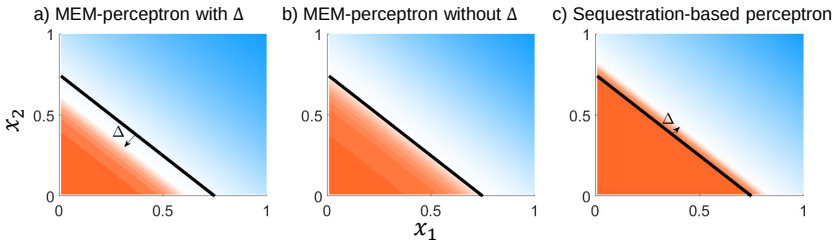

Figure 8: a-b) Removal of deviation Δ between true decision boundary (black) and produced decision boundary (white region) through functional approximation using Eq. 1. c) Presence of deviation Δ in the decision boundary of molecular sequestration-based perceptron.

Table 2: Context-dependency of XOR network search with minimized deviation Δ

Parameter	P1	P2	P3
w_0	0.1 to 1	0.1 to 1	0.01 to 1
R_{TOT}	5 to 20	5 to 20	5 to 20
E	5 to 20	5 to 20	5 to 20
Least Δ	0.1087	0.0075	0.1238 (U), 0.0412 (L)

A.6 Effect of temperature on the behavior of MEM

As part of analyzing the robustness of MEM against perturbations, and to identify additional interaction steps to absorb temperature variations, we apply the Arrhenius equation $k_{rate} = Ae^{-E_a/RT}$ to the forward rate constant of reaction r1 (Appendix A.1) at temperatures T_1 and T_2,

$$k_{NR}^{on}(T_1) = k_1 = Ae^{-E_a/RT_1}, \quad \ln k_1 = \ln A - \frac{E_a}{RT_1} \tag{15}$$

$$k_{NR}^{on}(T_2) = k_2 = Ae^{-E_a/RT_2}, \quad \ln k_2 = \ln A - \frac{E_a}{RT_2} \tag{16}$$

Eq. 16 - Eq. 15 removes the pre-exponential factor A, and we get,

$$\ln k_2 - \ln k_1 = -\frac{E_a}{RT_2} + \frac{E_a}{RT_1}, \quad \ln k_2 = \frac{E_a(T_2 - T_1)}{RT_1T_2} + \ln k_1 \tag{17}$$

Using Eq. 17, we can determine how the formation rate of the decision-making species NR, denoted as k_{NR}^{on} changes with variations in temperature. The MEM-perceptron relies on k_{NR}^{on} to produce a decision boundary, so it is important to absorb any fluctuations that temperature variation may cause in order to avoid erroneous classification. The related experimental studies are part of our ongoing explorations of the MEM-based biomolecular neural network.

A.7 Data and code availability

The simulation code and data used to generate the figures of the paper can be found at `https://github.com/theothermosh/MEM_BNN`.