

Computation of Natural Logarithm Using Abstract Chemical Reaction Networks

Iuliia Zarubiieva, Joyun Tseng, Vishwesh Kulkarni

Abstract—Recent researches has focused on nucleic acids as a substrate for designing biomolecular circuits for *in situ* monitoring and control. A common approach is to express them by a set of idealised abstract chemical reaction networks (ACRNs). Here, we present new results on how abstract chemical reactions, viz., catalysis, annihilation and degradation, can be used to implement circuit that accurately computes logarithm function using the method of Arithmetic-Geometric Mean (AGM), which has not been previously used in conjunction with ACRNs.

Keywords—Abstract chemical reaction network, DNA strand displacement, natural logarithm.

I. INTRODUCTION

AN objective of synthetic biology is to design biomolecular circuits for *in situ* monitoring and control. Recently, nucleic acid reactions have been proposed as a potential solution for these purposes [1]-[4]. A key advantage of nucleic acid reactions consists in the ease and precision with which these can be implemented, as their design relies essentially on the well-known Watson-Crick base-pairing mechanism (i.e. adenine-thymine and guanine-cytosine pairing), which enables precise programming and timing of molecular interactions simply by the choice of relevant sequences. This approach has allowed the implementation of a number of complex circuits based on DNA strand displacement [5], DNA enzyme [6] and RNA enzyme [7] reactions, and has been used for the modelling and implementation of various nucleic-acids-based circuits such as feedback controllers [8] and predator-prey systems [9]. Recently, it has been shown that any *chemical reaction network* can be closely approximated by a set of suitably designed DNA strand displacement reactions [10]. This logic can be extended to approximate a set of linear ordinary differential equations (ODEs) by a set of idealised *abstract chemical reaction networks* (ACRNs) which can then be approximated by a set of suitably designed DNA strand displacement reactions [4].

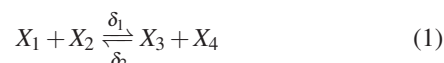
In order to exchange information with environment and make decisions on their behaviour, living cells use chemical reactions as a mean of communication. It was shown that logarithmic sensing is present in various signal transduction mechanisms of a cell and is related to the concept of fold-change detection. Hence, in order to decode the signals the cell is sending, it is necessary to compute natural logarithm [11]-[13].

Iuliia Zarubiieva, Joyun Tseng, and Vishwesh Kulkarni are with the School of Engineering, University of Warwick, Coventry, CV4 7AL, UK. (e-mail: I.Zarubiieva@warwick.ac.uk, J.Tseng.1@warwick.ac.uk, V.Kulkarni@warwick.ac.uk).

In this paper, we present a circuit for computing natural logarithm using the method of Arithmetic-Geometric Mean. We are also comparing the results obtained by AGM method with those received from the implementation of Newton-Raphson method and the method proposed by Chou [14].

II. NOTATION AND BACKGROUND RESULTS

To ensure consistency, the notation used in [15] and [4] is used throughout in this paper. For example, a bidirectional (i.e., a reversible bimolecular chemical reaction) is represented as



where X_i are chemical species with X_1 and X_2 being the reactants and X_3 and X_4 being the products. Here, δ_1 and δ_2 denote the forward and backward reaction rates, respectively. A unimolecular reaction features only one reactant whereas a multimolecular reaction features two or more reactants. Degradation of a chemical species X at rate K (or conversion of X into an inert form at a rate K) is denoted by $X \xrightarrow{K} \emptyset$.

A. Representing Signals Using Differences of Concentrations

Whereas signals in systems theory can take both positive and negative values, biomolecular concentrations (with Molar (M) as unit) can only take non-negative values. Thus, following the same approach suggested in [15] and [4], we represent a signal, x as the difference in concentration of two chemical species, x^+ and x^- . Here, x^+ and x^- are respectively the positive and negative components of x such that $x = x^+ - x^-$. The consequence of adopting this scheme is that there is no unique representation for a particular signal. As an example, $x = 20$ M can be represented by both $x^+ = 50$ M and $x^- = 30$ M or equivalently, $x^+ = 20$ M and $x^- = 0$ M. In practice, x^+ and x^- can be realised as single strand DNA molecules, as illustrated in [4] where these complementary positive and negative components would annihilate each other at reaction rate η (i.e. $x^+ + x^- \xrightarrow{\eta} \emptyset$). A key advantage of using this scheme is that it allows the realisation of the ‘subtraction’ operation, as discussed further below.

B. Realising Elementary Linear System Theoretic Operators

In [15], results on how to represent linear system theoretic operations such as gain, summation and integration using

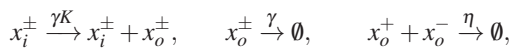
idealised abstract chemical reactions are presented and it is shown that only three types of elementary chemical reactions, namely, catalysis, annihilation and degradation are needed for such representations. In [4], this set of elementary chemical reactions is further reduced to only two. We here summarise their main results and refer the interested reader to [15] and [4] for details.

Throughout the rest of the paper, equations with superscript \pm and \mp are used as shorthand notations that represent the ‘+’ and ‘-’ individual reactions — for example, $x_i^\pm \xrightarrow{K} x_i^\pm + x_o^\pm$ should be understood as the the set of two reactions: $x_i^+ \xrightarrow{K} x_i^+ + x_o^+$ and $x_i^- \xrightarrow{K} x_i^- + x_o^-$. Likewise, the notation $x_i^\pm \xrightarrow{K} x_i^\pm + x_o^\mp$ is used to represent the set of two reactions: $x_i^+ \xrightarrow{K} x_i^+ + x_o^-$ and $x_i^- \xrightarrow{K} x_i^- + x_o^+$. For brevity and following [15], we will represent such a set of reactions compactly as $x_i^\pm \xrightarrow{K} x_i^\pm + x_o^\pm$ and $x_i^\pm \xrightarrow{K} x_i^\pm + x_o^\mp$.

As noted in [15], one limitation of representing signals as the difference of concentrations is that the requirement of having the same reaction rate, K , for both positive and negative components may not be easy to implement experimentally. However, as shown in [15], this requirement can be relaxed if the annihilation rate, η in the annihilation reaction, $x_o^+ + x_o^- \xrightarrow{\eta} \emptyset$ is chosen to be sufficiently large. Hence, we assume this condition of $\eta \gg K$ throughout the rest of this paper.

Lemma 1. [Scalar gain K]

Let $x_o = Kx_i$ where x_i is the input, x_o is the output and K is the gain. This operation is implemented using the following set of abstract chemical reactions:



where γK , γ and η are the kinetic rates associated with catalysis, degradation and annihilation respectively. □

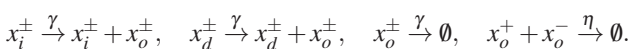
Proof. Using generalised mass-action kinetics, it follows that the gain operator realised in this manner is described using the following ODE,

$$\begin{aligned} \frac{dx_o^+}{dt} &= \gamma(Kx_i^+ - x_o^+) - \eta x_o^+ x_o^- \\ \frac{dx_o^-}{dt} &= \gamma(Kx_i^- - x_o^-) - \eta x_o^+ x_o^- \\ \frac{dx_o}{dt} &= \frac{dx_o^+}{dt} - \frac{dx_o^-}{dt} = \gamma(Kx_i - x_o) \end{aligned} \quad (2)$$

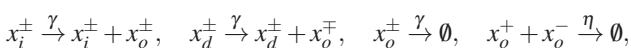
Using the final value theorem, the steady state value of x_o for constant input x_i is given by $\lim_{t \rightarrow \infty} x_o(t) = Kx_i(t)$. □

Lemma 2. [Summation]

Consider the summation operation $x_o = x_i + x_d$, where x_i and x_d are the inputs and x_o is the output. This operation is implemented using the following set of abstract chemical reactions:



Using the following set of abstract chemical reactions:



the subtraction $x_o = x_i - x_d$ is implemented. □

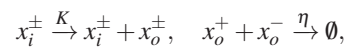
TABLE I
STEPS FOR COMPUTING $\ln(x)$ USING AGM

Computation steps:	
1) Initialization	$w(0) = 4/x$ $g(0) = 1$
2) Iteration	$w_{n+1} = (w_n + g_n)/2$ $g_{n+1} = \sqrt{w_n g_n}$ = AGM(1, 4/x)
3) Compute $\ln(x)$	$\ln(x) = \frac{\pi/2}{AGM(1, 4/x)}$

Remark 1. Scaled summation $x_o = K(x_i + x_d)$, and scaled subtraction, can be implemented by choosing the catalysis rates in the construct of Lemma 2 to be $K\gamma$. □

Lemma 3. [Scaled Integration]

Consider the integrator $x_o = K \int x_i dt$ where x_i is the input, x_o is the output, and K is the DC gain. Using the following set of abstract chemical reactions:



such an integrator is implemented. □

Proof. Using generalised mass-action kinetics the ODEs for the summation and integrator operations are given by $\frac{dx_o}{dt} = \gamma(x_i + x_d - x_o)$ and $\frac{dx_o}{dt} = Kx_i$, respectively. Then, the proof for Lemmas 2 and 3 can be trivially obtained following the same logic as for the proof of Lemma 1. □

III. MAIN RESULT

Here, we are presenting block diagram and computation details of the circuit for computation of $\ln(x)$. It is an important function for biological systems as it can be related to the cell signalling mechanisms. To design the circuit, we used the principle of Arithmetic-Geometric Mean. Also, we recreated the circuits based on Newton-Raphson method and by Chou, and compared the obtained results.

AGM is a hybrid quantity which is defined by combining the arithmetic and geometric means of two positive numbers. The arithmetic mean of two numbers w and g is defined as $w_{n+1} = (w_n + g_n)/2$; the geometric mean of same two numbers is $g_{n+1} = \sqrt{w_n g_n}$. Further, the process is iterative and converges to a number that is between the arithmetic mean and the geometric mean.

It has been shown by [16] that AGM can be used for approximating the value of natural logarithm:

$$\ln(x) \approx \frac{\pi/2}{AGM(1, \frac{4}{x})} \quad (3)$$

The steps for implementing the computation are listed in Table I. Firstly, we set out the initial values of signals w and g , where one signal is set as “ $\frac{4}{x}$ ” and the other is “1”. Then we find the Arithmetic-Geometric Mean of two inputs, after which we can approximate the $\ln(x)$.

Fig. 1 illustrates a block diagram for computing AGM (steps 1 and 2 from Table I). In Fig. 2 the output of AGM block (Fig. 1) serves as one input, the second input is a signal of the value $\frac{\pi}{2}$. This circuit performs accurate division of two signals and was presented by us earlier [17]. Since all the blocks can be derived using abstract chemical reactions, we note down all ACRNs and ODEs in Table III.

TABLE II
 COMPUTATION OF $\ln(x)$ USING AGM: DNA IMPLEMENTATION, THE ACRNs AND THE ODES

DNA Implementation	Formal CRNs	ODEs
Block 1		
$g^\pm + G_1^\pm \xrightarrow{q_1} \emptyset + O_1^\pm$	$g^\pm \xrightarrow{\frac{1}{2}\gamma_1} g^\pm + z^\pm$	$z = \gamma_1(\frac{1}{2}(g-w) - z)$
$O_1^\pm + T_1^\pm \xrightarrow{q_{max}} g^\pm + z^\pm$		
$w^\mp + G_2^\pm \xrightarrow{q_2} \emptyset + O_2^\pm$		
$O_2^\pm + T_2^\pm \xrightarrow{q_{max}} w^\mp + z^\pm$	$w^\mp \xrightarrow{\frac{1}{2}\gamma_1} w^\mp + z^\pm$	
$z^\pm + G_3^\pm \xrightarrow{q_3} \emptyset$	$z^\pm \xrightarrow{\gamma_1} \emptyset$	
$z^+ + L_{1z} \xrightarrow{q_{max}} H_{1z} + B_{1z}$	$z^+ + z^- \xrightarrow{\eta} \emptyset$	
$z^- + L_{S_{1z}} \xrightarrow{q_{max}} H_{S_{1z}} + B_{S_{1z}}$		
$z^- + H_{1z} \xrightarrow{q_{max}} \emptyset$		
Block 2		
$z^\pm + G_4^\pm \xrightarrow{q_4} \emptyset + O_4^\pm$	$z^\pm \xrightarrow{\gamma_2} z^\pm + w^\pm$	$\dot{w} = \gamma_2 z$
$O_4^\pm + T_4^\pm \xrightarrow{q_{max}} z^\pm + w^\pm$		
$w^+ + L_{1w} \xrightarrow{q_{max}} H_{1w} + B_{1w}$		
$w^- + L_{S_{1w}} \xrightarrow{q_{max}} H_{S_{1w}} + B_{S_{1w}}$	$w^+ + w^- \xrightarrow{\eta} \emptyset$	
$w^- + H_{1w} \xrightarrow{q_{max}} \emptyset$		
Block 3		
$w^\pm + g^\pm + G_5^\pm \xrightarrow{q_5} \emptyset + O_5^\pm$	$w^\pm + g^\pm \xrightarrow{\gamma_3} w^\pm + g^\pm + u^\pm$	$\dot{u} = \gamma_3(w * g - u)$
$O_5^\pm + T_5^\pm \xrightarrow{q_{max}} w^\pm + g^\pm + u^\pm$		
$u^\pm + G_6^\pm \xrightarrow{q_6} \emptyset$		
$u^+ + L_{3u} \xrightarrow{q_{max}} H_{3u} + B_{3u}$	$u^\pm \xrightarrow{\gamma_3} \emptyset$	
$u^- + L_{S_{3u}} \xrightarrow{q_{max}} H_{S_{3u}} + B_{S_{3u}}$	$u^+ + u^- \xrightarrow{\eta} \emptyset$	
$u^- + H_{3u} \xrightarrow{q_{max}} \emptyset$		
Block 4		
$\frac{1}{2}u^\pm + G_7^\pm \xrightarrow{q_7} \emptyset + O_7^\pm$	$\frac{1}{2}u^\pm \xrightarrow{\gamma_4} \frac{1}{2}u^\pm + v^\pm$	$\dot{v} = \gamma_4(u^{\frac{1}{2}} - v)$
$O_7^\pm + T_7^\pm \xrightarrow{q_{max}} \frac{1}{2}u^\pm + v^\pm$		
$v^\pm + G_8^\pm \xrightarrow{q_8} \emptyset$		
$v^+ + L_{4v} \xrightarrow{q_{max}} H_{4v} + B_{4v}$	$v^\pm \xrightarrow{\gamma_4} \emptyset$	
$v^- + L_{S_{4v}} \xrightarrow{q_{max}} H_{S_{4v}} + B_{S_{4v}}$	$v^+ + v^- \xrightarrow{\eta} \emptyset$	
$v^- + H_{4v} \xrightarrow{q_{max}} \emptyset$		

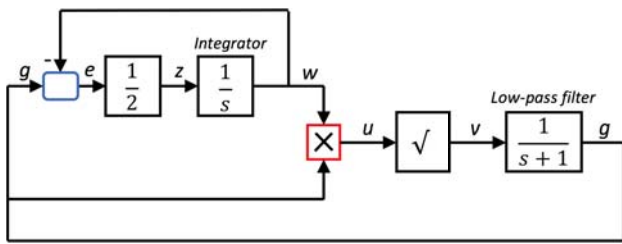


Fig. 1 Block diagram of the AGM block

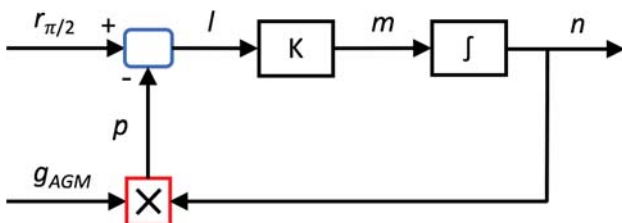


Fig. 2 Block diagram of the circuit for obtaining $\ln(x)$ by AGM method

Fig. 3 provides the overview of the computation results obtained by 3 different methods: AGM, Chou and Newton-Raphson. More detailed investigation showed that the system by our method has the lowest percentage error of $3,09 \times 10^{-4}\%$ for the chosen example. The system tends to have smaller error (less than 8%) for x between 50 and 300, and the error increases as x moves further away from this interval. So, the system is to be improved to overcome this disadvantage.

IV. CONCLUSIONS

We have presented a circuit for computing natural logarithm using the method of Arithmetic-Geometric Mean, and its abstract chemical reaction network representation. We have compared the obtained results with those received from Newton-Raphson and the method by Chou [14]. Our design shows the lowest percentage error, however, the improvements can be made to reduce it even more for chosen parameter range.

In Table III, we note down the ODE's, ACRNs, and DNA implementation details for our circuit. The annihilation reaction rate η is to be chosen arbitrarily large.

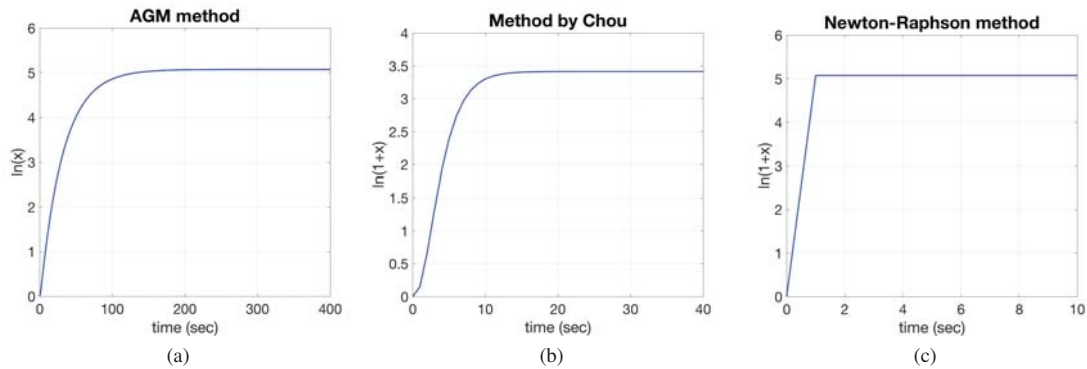


Fig. 3 Matlab simulation results for computing natural logarithm using a) AGM method; b) method by Chou [14] and c) Newton-Raphson method

TABLE III
 COMPUTATION OF $\ln(x)$ USING AGM: DNA IMPLEMENTATION, THE ACRNS AND THE ODES (CONTINUED)

DNA Implementation	Formal CRNs	ODEs
Block 5		
$v^\pm + G_9^\pm \xrightarrow{q_9} \emptyset + O_9^\pm$	$\left. \begin{array}{l} v^\pm \xrightarrow{\gamma_5} v^\pm + g^\pm \\ g^\pm \xrightarrow{\gamma_5} \emptyset \\ \dots \\ g^+ + g^- \xrightarrow{\eta} \emptyset \end{array} \right\}$	$\dot{g} = \gamma_5(v - g)$
$O_9^\pm + T_9^\pm \xrightarrow{q_{max}} v^\pm + g^\pm$		
$g^\pm + G_{10}^\pm \xrightarrow{q_{10}} \emptyset$		
$g^+ + L_{5g} \xrightarrow{q_{max}} H_{5g} + B_{5g}$		
$g^- + L_{5g} \xrightarrow{q_{max}} H_{5g} + B_{5g}$		
$g^- + H_{5g} \xrightarrow{q_{max}} \emptyset$		
Block 6		
$g^\pm + G_{11}^\pm \xrightarrow{q_{11}} \emptyset + O_{11}^\pm$	$\left. \begin{array}{l} g^\pm \xrightarrow{K\gamma_6} g^\pm + m^\pm \\ p^\mp \xrightarrow{K\gamma_6} p^\mp + m^\pm \\ \dots \\ m^\pm \xrightarrow{\gamma_6} \emptyset \\ \dots \\ m^+ + m^- \xrightarrow{\eta} \emptyset \end{array} \right\}$	$\dot{m} = \gamma_6(K(r - p) - m)$
$O_{11}^\pm + T_{11}^\pm \xrightarrow{q_{max}} g^\pm + m^\pm$		
$p^\mp + G_{12}^\pm \xrightarrow{q_{12}} \emptyset + O_{12}^\pm$		
$O_{12}^\pm + T_{12}^\pm \xrightarrow{q_{max}} p^\mp + m^\pm$		
$m^\pm + G_{13}^\pm \xrightarrow{q_{13}} \emptyset$		
$m^+ + L_{6m} \xrightarrow{q_{max}} H_{6m} + B_{6m}$		
$m^- + L_{6m} \xrightarrow{q_{max}} H_{6m} + B_{6m}$		
$m^- + H_{6m} \xrightarrow{q_{max}} \emptyset$		
Block 7		
$m^\pm + G_{14}^\pm \xrightarrow{q_{14}} \emptyset + O_{14}^\pm$	$\left. \begin{array}{l} m^\pm \xrightarrow{\gamma_7} m^\pm + n^\pm \\ \dots \\ n^+ + n^- \xrightarrow{\eta} \emptyset \end{array} \right\}$	$\dot{n} = \gamma_7 m$
$O_{14}^\pm + T_{14}^\pm \xrightarrow{q_{max}} m^\pm + n^\pm$		
$n^+ + L_{7n} \xrightarrow{q_{max}} H_{7n} + B_{7n}$		
$n^- + L_{7n} \xrightarrow{q_{max}} H_{7n} + B_{7n}$		
$n^- + H_{7n} \xrightarrow{q_{max}} \emptyset$		
Block 8		
$g^\pm + n^\pm + G_{15}^\pm \xrightarrow{q_{15}} \emptyset + O_{15}^\pm$	$\left. \begin{array}{l} g^\pm + n^\pm \xrightarrow{\gamma_8} g^\pm + n^\pm + p^\pm \\ p^\pm \xrightarrow{\gamma_8} \emptyset \\ \dots \\ p^+ + p^- \xrightarrow{\eta} \emptyset \end{array} \right\}$	$\dot{p} = \gamma_8(g * n - p)$
$O_{15}^\pm + T_{15}^\pm \xrightarrow{q_{max}} g^\pm + n^\pm + p^\pm$		
$p^\pm + G_{16}^\pm \xrightarrow{q_{16}} \emptyset$		
$p^+ + L_{8p} \xrightarrow{q_{max}} H_{8p} + B_{8p}$		
$p^- + L_{8p} \xrightarrow{q_{max}} H_{8p} + B_{8p}$		
$p^- + H_{8p} \xrightarrow{q_{max}} \emptyset$		

ACKNOWLEDGMENT

This research is supported, in parts, by the EPSRC INDUSTRIAL CASE AWARD (CASE Voucher 16000070), Microsoft Research, and the EPSRC/BBSRC grant BB/M017982/1 to the Warwick Integrative Synthetic Biology Centre.

REFERENCES

- [1] Seelig, G., Soloveichik, D., Zhang, D.Y., & Winfree, E. (2006). Enzyme-free nucleic acid logic circuits. *Science*, 314, 1585-1588.
- [2] Zhang, D.Y., Turberfield, A.J., Yurke, B. & Winfree, E. (2007). Engineering entropy-driven reactions and networks catalyzed by DNA. *Science*, 318, 1121-1125.
- [3] Padirac, A., Fujii, T., & Rondelez, Y. (2013). Nucleic acids for the rational design of reaction circuits. *Current Opinion of Biotechnology*, 24, 575-580.
- [4] Yordanov, B., Kim, J., Petersen, R.L., Shudy, A., Kulkarni, V.V., & Philips, A. (2014). Computational design of nucleic acid feedback control circuits. *ACS Synthetic Biology*, 3, 600-616.
- [5] Zhang, D.Y. (2011). Towards domain-based sequence design for DNA strand displacement reactions. *DNA Computing and Molecular Programming, Springer Berlin Heidelberg*, 162-175.
- [6] Montagne, K., Plasson, R., Sakai, Y., Fujii, T., & Rondelez, Y. (2011). Programming an *in vitro* DNA oscillator using a molecular networking strategy. *Molecular Systems Biology*, 7, 466.
- [7] Kim, J., & Winfree, E. (2011). Synthetic *in vitro* transcriptional oscillators. *Molecular Systems Biology*, 7, 465.
- [8] Chen, Y.-J., Dalchau, N., Srinivas, N., Philips, A., Cardelli, L., Soloveichik, D., & Seelig, G. (2013). Programmable chemical controllers made from DNA. *Nature Nanotechnology*, 8, 755-762.
- [9] Fujii, T., & Rondelez, Y. (2013). Predator-prey molecular ecosystems. *ACS Nano*, 7, 27-34.
- [10] Soloveichik, D., Seelig, G., Winfree, E. (2010). DNA as a universal substrate for chemical kinetics. *Proceedings of National Academy of Sciences, USA*, 12, 5393-5398.
- [11] Alberts, B. and Johnson, A. and Lewis, J. and Raff, M. and Roberts, K. and Walter, P. (2007). *Molecular Biology of the Cell* (5th Edition). Garland Science, New York, NY.
- [12] Lim, W. and Mayer, B. and Pawson, T. (2014). *Cell Signaling*. Garland Science, New York, NY.
- [13] Ma, K.C. and Perli, S.D. and Lu, T.K. (2016). Foundations and emerging paradigms for computing in living cells. *Journal of Molecular Biology*, 428, pp. 893-915.
- [14] Chou, C. T. (2017) Chemical reaction networks for computing logarithm. *Synthetic Biology*, 2(1), ysx002.
- [15] Oishi, K., & Klavins, E. (2011). Biomolecular implementation of linear I/O systems. *IET Systems Biology*, 5, 252-260.
- [16] Brent, R. P. (2018). Fast Algorithms for High-Precision Computation of Elementary Functions, 5.
- [17] Zarubiieva, I., Tseng, J.Y. and Kulkarni, V. (2018). Accurate Ratio Computation using Abstract Chemical Reaction Networks. *IAENG WCE 2018: International Association of Engineers World Congress on Engineering*.