RESEARCH STATEMENT

Tracey Oellerich (toelleri@gmu.edu)

My research focuses on mathematical biology, particularly the modeling of biological structures and diseases. What I find fascinating about this work is how it utilizes techniques from different disciplines of mathematics, including network science, data science, linear algebra, and more. In addition, my work presents a scope of adaptation into a graph theory and an algebraic geometry framework, making it accessible and useful to a broader audience. I am interested in uncovering why biological networks behave the way they do and how this can be used to help inform other sciences. These results have applications in disease research, such as for cancers and drug therapies. Furthermore, while I am primarily studying applications to biological systems, the results are not limited to this field. This work can be applied to various fields, such as control theory and recovering dynamics from any form of data. This research has lead to two publications so far, one focusing on extending adaptation criteria [16] and the other on inferring dynamics from biological data [17].

Past Research

During my undergraduate years at Wilkes University, I was introduced to the idea of utilizing mathematics in biology applications. I began in the field of computational biophysics where I worked on "Enhanced protein folding through confinement inside a hydrophilic nanopore." For this work, we investigated the complexity of protein folding inside a cell, in particular, under confinement. Proteins are large molecules composed of one or more long chains of amino acids. Protein folding is the process in which a protein reaches its functional shape. In this heterogeneous environment, proteins can fold inside of special nanocages called chaperones, which are cone-shaped proteins that assist other proteins in folding or unfolding. Through my work, I have developed a model describing how these cages can accelerate protein folding by modifying the thermodynamics of confined water. I tested these models through distributed computing and molecular dynamics simulation and used state of the art master equation approaches to analyze the results and verify my model.

During a study abroad in Padua, Italy, I had the opportunity to participate in the "Bench to Bedside: Translational Molecular Research" course. This course focused on the practical use of genomics, proteomics and bioinformatics for the diagnosis, prevention and treatment of disease. It was co-taught by faculty from US and Italy including experts in the field of medical research on topics such as leukemia, cancer research, infectious diseases, etc. The course utilized hands-on technical workshops and touched on issues related to implementation of new platforms and assays. Through both presentations and experiments, I was given the opportunity to see where the data I work with is collected, the challenges involved in this process, and a different prospective on how it is used in research.

In the summer of 2021, I was given the opportunity to participate as a Data Science intern at the National Center for Advancing Translational Sciences (NCATS) at the National Institutes of Health (NIH). There, I worked on using deep learning to prioritize disease targets for protein-protein interaction networks. We worked on utilizing the data in the Pharos library of known diseases and protein targets to create a ranked list of protein targets by modifying the GuiltyTargets algorithm, a machine learning algorithm that ranks protein targets using gene expression and positive-unlabeled machine learning. This work helps to shave years off the drug development process by eliminating unsuitable targets without the need for testing.

Through my previous work and experience, I was able to discover how fascinating I find applying mathematics to biological systems. One of the things that drew me to this field is being able to understand how complex biological systems work from a mathematical perspective. This can then be

used in such a way that we can understand and model systems which are either too small to study efficiently in a lab, such as protein folding, require expensive lab equipment to obtain data or would require years to validate experimentally.

Current Research

My PhD research is carried out under the supervision of Prof. Maria Emelianenko at George Mason University (GMU). I am working on interdisciplinary research at the intersection of Mathematics and Biology. I am primarily interested in data science and network science, which can be used to model biological structures and analyze them. In particular, I work with biological networks which are systems formed by the interactions of a finite number of molecules such as proteins, RNA transcripts, genes, or various combinations. This work is helpful in identifying effective drug treatments for adapting diseases, such as cancer, in which a drug (external input) being applied to the system is being neutralized within the network. My work incorporates both theoretical and experimental mathematics in order to identify these features. Utilizing aspects of linear algebra and graph theory, I work to understand how and why a biological network will adapt to external inputs. Then, using experimental data, I employ novel data science and machine learning techniques to infer dynamics and create a dynamical system model.

Robust Perfect Adaptation

The first part of my work focuses on understanding and extending the notion of Robust Perfect Adaptation (RPA). Adaptation in the sense of asymptotic tracking of a 'set-point,' has been widely explored in the literature [4, 11] at various levels ranging from an individual cell to the whole-organism level in mammals. At the cellular level, several types of adaptation have been studied in previous works, including perfect adaptation [9], fold-change detection (FCD) [20], absolute concentration robustness [19], homeostasis [21], and robust perfect adaptation [3]. All of these concepts share adapting behavior, although they highlight certain specific types of adaptive behavior. My focus will be on understanding and extending the criteria under which a biological network will exhibit RPA. *Robust perfect adaptation (RPA)* refers to the property of a biological system to return to the same activity level following any persistent change to the incoming signal received at the input node, without the need for fine-tuning of parameters [3]. There are several reported instances of RPA occurring in biological systems, such as in bacterial chemotaxis [23, 5, 2], EGFR-regulated signaling pathways ([9],[18]), and transcription networks [10]. This notion is closely related to the phenomenon called *homeostasis*, which is the biological property of maintaining a fixed state despite external perturbations, generally achieved by coupling a biological sensing mechanism to a feedback loop, so when perturbations attempt to change the state of the system, the feedback mechanism acts to resist these changes and restore the system to its default state. Under this definition, one may distinguish between single-input/single-output and more general multi-input/multi-output systems.

Consider a system with state space denoted as $\mathcal{P} \subset \mathbb{R}^N$. The state space consists of N nodes, say $\mathbf{P}(t) = [P_1(t), \ldots, P_N(t)]^T \in \mathcal{P}$, representing the interacting molecules of interest, such as proteins, RNA transcripts, genes.

Let $\mathcal{U} \subset \mathbb{R}^M$ be the *input* space and $\mathbf{U}(t) = [U_1(t), U_2(t), \ldots, U_M(t)]^T \in \mathcal{U}$ be the time dependent input to the system. Consider the nonlinear dynamical system (1) for this system.

$$\frac{dP_i}{dt} = f_i(\mathbf{U}, \mathbf{P}) \qquad i = 1, ..., N \tag{1}$$

For a system to exhibit RPA, we have the following two conditions:

$$\det(M_{IO}) = 0 \tag{2}$$

$$\det(D_{\mathbf{P}}\mathbf{F}) \neq 0 \tag{3}$$

where (2-3) is precisely the *RPA equation* and *RPA constraint* as defined in [3].

The RPA criteria defined in Araujo et al. does not account for systems in which the RPA constraint ($\det(D_{\mathbf{P}}\mathbf{F}) \neq 0$) is not met. To this end, we have developed criteria which extends the notion of RPA to account for a system in which $\det(D_{\mathbf{P}}\mathbf{F}) = 0$.

Proposition 1 (Extension of the RPA Criteria)**.** *A biochemical system* (1) *with inputs* $\boldsymbol{U} \in \mathbb{R}^M$ *for which* $\det(D_{\boldsymbol{P}}\boldsymbol{F})(P^*) = 0$ *exhibits adaptation if*

$$hV_1\Sigma_1^{-1}W_1^T D_{\boldsymbol{U}}\boldsymbol{F} = 0 \tag{4}$$

at the system's steady state P^**, where* V_1, Σ_1*, and* W_1 *are the components of the reduced singular value decomposition of* $D_{\boldsymbol{P}}\boldsymbol{F}$*.*

The proof for Proposition 1 can be found in Oellerich et al. [16].

Sources of Singular Jacobians

The question then becomes what conditions cause a biological network to exhibit a singular jacobian at the network steady state? In Oellerich et al., we prove the instance of a singular Jacobian for several cases.

Proposition 2 (Linear Conservation Law)**.** *Consider a biological system with* $N \geq 2$ *nodes which contains a linear conservation law. The system is defined as follows:*

$$\begin{aligned} \frac{d\boldsymbol{P}}{dt} &= \boldsymbol{F}(\boldsymbol{P}) \\ P_{tot} &= \boldsymbol{c}^T\boldsymbol{P}(t) \end{aligned} \tag{5}$$

where $\boldsymbol{c} \not\equiv \boldsymbol{0}$ *is a vector of coefficients in the conservation law and* P_{tot} *is the total concentration which remains constant. Then,* $\det(J(\boldsymbol{P}(t))) = 0, \forall t \geq 0$*.*

Proposition 3 (Nonlinear Conservation Law)**.** *Consider a biological system with* $N \geq 2$ *nodes which contains a nonlinear conservation law. Let* $M = \{i : P_i \text{ appears in the conservation law}\}$*,* $2 \leq m = |M| \leq N$*. The system is defined as follows:*

$$\begin{aligned} \frac{d\boldsymbol{P}}{dt} &= \boldsymbol{F}(\boldsymbol{P}) \\ P_{tot} &= v(P_{i_1}, \ldots, P_{i_m}), \qquad i_j \in M \text{ are distinct} \end{aligned} \tag{6}$$

where $v \in C^2(\mathcal{P}^m)$*,* $\dfrac{\partial v}{\partial P_{i_j}}(P^*_{i_1}, \ldots, P^*_{i_m}) \neq 0$ *for non-trivial steady state* $\boldsymbol{P}^*$*, and* P_{tot} *is the total concentration which remains constant. Then,* $\det(J(\boldsymbol{P}^*)) = 0$ *for this system.*

Proposition 4 (Special Network Structures)**.** *Consider an* N*-node network,* $N \geq 3$*, in which one node, say* P_N*, is a function of the other nodes within the network. Let the system be defined as:*

$$\begin{cases} \dfrac{d\boldsymbol{P}}{dt} = \boldsymbol{F}(\boldsymbol{P}) \\ P_N = v(P_1, \ldots, P_{N-1}) \end{cases} \tag{7}$$

where $v \in C^2(\mathcal{P}^{N-1})$ *and at least one* $\dfrac{dP_N}{dP_i} \neq 0$ *for* $i = 1, \ldots, N-1$*. Then,* $\det(J(\boldsymbol{P}^*)) = 0$ *for this system.*

Inferring Dynamics

For my work in adaptation, I am using theoretical models to explore the requirements for adaptation. Now I will discuss my efforts to recover the dynamics for biological networks from data [17], thus helping to eliminate utilizing assumptions to construct the model as has been done in past works. My inspiration for this work began with the Sparse Identification of Nonlinear Dynamics (SINDy) method proposed by Brunton et al. [6]. The SINDy algorithm presents a method for automating the discovery of the governing equations. SINDy works under the assumption that many dynamical systems, $\frac{d}{dt}\mathbf{x} = \mathbf{f}(\mathbf{x})$, have dynamics with only a few active terms in the space of all possible right-hand side functions. This enables SINDy to bypass the combinatorial search through all possible model structures. SINDy begins with time-series data $\mathbf{X} = [\mathbf{X}_1(t), \ldots, \mathbf{X}_n(t)]$ harvested from experiments and assuming the structure of the dynamical system is a generalized linear model:

$$\mathbf{f}_k(\mathbf{x}) \approx \boldsymbol{\Theta}(\mathbf{x})\boldsymbol{\xi}_k \qquad k = 1, \ldots, n \tag{8}$$

where $\boldsymbol{\Theta}(\mathbf{x}) \in \mathbb{R}^{1\times l}, \boldsymbol{\xi}_k \in \mathbb{R}^{l\times 1}$, l is the number of candidate nonlinear functions in $\boldsymbol{\Theta}(\mathbf{x})$, and $\boldsymbol{\xi}_k$ contains the fewest nonzero terms as possible. Nonzero entries of the sparse vector $\boldsymbol{\xi}_k$ correspond to the active terms in the resulting dynamical system. Here, $\boldsymbol{\Theta}(\mathbf{x})$ refers to the library of candidate nonlinear functions constructed from the data:

$$\boldsymbol{\Theta}(\mathbf{X}) = \begin{bmatrix}\mathbf{1} & \mathbf{X} & \mathbf{X}^2 & \ldots & \mathbf{X}^d & \ldots & \sin(\mathbf{X}) & \ldots\end{bmatrix} \tag{9}$$

The choice of the $\boldsymbol{\Theta}$ library is up to the user's discretion and is only limited by their imagination. For example, if one were attempting to recover a second-order mass-action system, the $\boldsymbol{\Theta}$ library may take the form: $\boldsymbol{\Theta}(\mathbf{X}) = [\mathbf{1}, \mathbf{X}, \mathbf{X}^2]$. Once the library matrix has been defined, the algorithm utilizes l_1 normalized sparse regression on $\boldsymbol{\xi}_k = \arg\min_{\boldsymbol{\xi}_k'} \|\dot{\mathbf{X}}_k - \boldsymbol{\Theta}(\mathbf{X})\boldsymbol{\xi}_k'\|_2 + \lambda\|\boldsymbol{\xi}_k'\|_1$, where λ is used to enforce sparsity and determined by a Pareto front [15], in order to solve for optimal coefficients.

One limitation of the SINDy algorithm is that it cannot be used to identify dynamics which contains functions such as rational functions, such as Michaelis–Menten kinetics and Hill equations. In Mangan et al. [15] and Kaheman et al. [12], they propose a solution to this by modifying the SINDy algorithm to allow for the identification of implicit dynamics, called implicit SINDy (implicit SINDy) and Robust Parallel Identification of Implicit Dynamics (SINDy-PI), respectfully. For implicit SINDy, the identification, and main source of computational complexity, lies in the identification of the sparsest vector $\boldsymbol{\xi}$ for the subspace $S = Null(\boldsymbol{\Theta}) \in \mathbb{R}^p$ with $dim(S) = n < p$. SINDy-PI tries to alleviate the need for finding the null space by transforming the problem into an iterative method which re-formats the problem. During the iterations, the method assumes a coefficient in the recovered model is known and then solves for the known coefficient and associated function; however, it is important to note that this method incorporates derivative information into the recovery. The problem is then in a form which can be solved using the optimization routine in SINDy. After iterating, the best model is chosen as the solution.

My work utilizes the same ideas presented in SINDy (and implicit SINDy) to infer dynamics, however, the optimization routine utilized differs. In our efforts to find a numerical method which responds well to noise and low amounts of data, we looked for alternatives to SINDy. Our solution was to use the Non-negative Least Squares (NNLS)[13], which is a variation of the Linear Least Squares (LLS) algorithm with linear inequality constraints. This algorithm as we will show can perform well with relatively little data and responds well even in the presence of extreme noise.

As with SINDy, consider an approximation of $\mathbf{f}(\mathbf{x})$ using a generalized linear model:

$$\mathbf{f}_k(\mathbf{x}) \approx \begin{bmatrix}\boldsymbol{\Theta}(\mathbf{x}) \\ -\boldsymbol{\Theta}(\mathbf{x})\end{bmatrix}\boldsymbol{\omega}_k \tag{10}$$

where $\boldsymbol{\Theta}$ is a library of candidate nonlinear functions constructed from the data and $\boldsymbol{\omega}_k$ contains the fewest terms ≥ 0 as possible. Now the goal is to find the optimal $\boldsymbol{\omega}_k$ which correspond to the recovered coefficients. Utilizing the time series data and the **Non-negative Least Squares (NNLS)** algorithm [13], consider the following minimization problem:

$$\boldsymbol{\omega}_k = \underset{\boldsymbol{\omega}_k' \geq 0}{\arg\min} \left\| \begin{bmatrix} \boldsymbol{\Theta}(\mathbf{X}) \\ -\boldsymbol{\Theta}(\mathbf{X}) \end{bmatrix} \boldsymbol{\omega}_k' - \dot{\mathbf{X}}_k \right\|_2 \tag{11}$$

where the top entries of $\boldsymbol{\omega}_k$ will correspond to positive coefficients in the recovered dynamics and the bottom entries are the negative.

I was able to extend this method to allow for rational functions in the dynamics, such as Michaelis–Menten kinetics which are common in biology. Consider a dynamical system of the form:

$$\frac{d}{dt} x_k(t) = \frac{f_N(\mathbf{x})}{f_D(\mathbf{x})} \tag{12}$$

where $f_N(\mathbf{x})$ and $f_D(\mathbf{x})$ represent the numerator and denominator polynomials in the state variable $\mathbf{x} \in \mathbb{R}^n$ respectfully. Now approximate $f_{N,k}(\mathbf{x})$ and $f_{D,k}(\mathbf{x})$ by generalized linear models:

$$f_{N,k} \approx \Theta_N(\mathbf{x})\omega_{N,k} \tag{13}$$

$$f_{D,k} \approx \Theta_D(\mathbf{x})\omega_{D,k} \tag{14}$$

where $\Theta_N(\mathbf{x}), \Theta_D(\mathbf{x})$ are the candidate function libraries and $\boldsymbol{\omega}_{N,k}, \boldsymbol{\omega}_{D,k}$ are the corresponding coefficients.

We will consider the library of candidate linear functions constructed from the data:

$$\Theta_N(\mathbf{X}) = \Theta_D(\mathbf{X}) = \begin{bmatrix} \mathbf{1} & \mathbf{X} & \mathbf{X}^2 & \dots & \mathbf{X}^d & \dots \end{bmatrix} \tag{15}$$

Thus, we have:

$$\Theta_N(\mathbf{x})\omega_{N,k} - \Theta_D(\mathbf{x})\omega_{D,k}\dot{x}_k = 0 \tag{16}$$

In order to apply NNLS, we will assume the coefficient associated with the $x_k\dot{x}_k$ term is 1 and thus solve:

$$x_k\dot{x}_k = \Theta_N(\mathbf{x})\omega_{N,k} - \tilde{\Theta}_D(\mathbf{x})\omega_{D,k}\dot{x}_k \tag{17}$$

where $\tilde{\Theta}_D$ is the Θ_D matrix with the column corresponding to x_k removed. The optimal $\omega_{N,k}$ and $\omega_{D,k}$ can now be found using a similar method as before.

Future Directions

As I have worked on these problems, I have seen that there are many areas where the current work can be extended. Below I provide future research ideas I will explore.

Algebraic Geometry Approach

Currently, the criteria for adaptation is defined in linear algebra terms; however, there is room to extend these results using an algebraic geometry approach. Future work includes extending the results for the conservation laws and special network structures to an algebraic interpretation. In recent years, many mathematicians have approached systems biology using algebraic geometry. In her paper, Dickenstein [8] provides a survey of the recent applications of algebraic geometry in the understanding of systems biology. By extending these conditions to a more algebraic interpretation, we will have a connection between networks that exhibit RPA as presented in Araujo et al. [3] and those which require the generalized RPA condition.

Graph Theory

In [9] and [14], there is discussion of both perfect adaptation and near-perfect adaptation, however, they do not clarify what they mean by near-perfect adaptation. A clear understanding of what it means to be close to adaptation is useful for classifying networks and potentially being able to modify them to adapt.

Furthermore, I would like to examine how sensitive adaptation is to change. Of particular interest are the effects to the adaptation criteria when (1) the coefficients of the system are perturbed and (2) the network structure is perturbed. In the case of RPA, changing the network coefficients should have very little to no effect on the network's ability to adapt by the definition of RPA. For changes to the network structure, I expect to see networks potentially losing their adaptation properties by adding or removing connections. One way to potentially quantify the changes that will affect the adaptation properties is to look at the graph entropy [7], and in particular, the Kullback-Leiber (or Jensen-Shannon) divergence between the original network and that which was modified. In West et al. [22], network entropy is used to show that cancer is characterized by an increase in the network entropy. I conjecture that adapting behavior is a fragile property of a living system that may be easily lost by relatively simple network modifications.

Research Task 1. *I would like to explore the idea of a measure of adaptation and the limits of a biological network by answering the following questions:*

(a) Is there a way to measure a system's ability to adapt and how close it is to adaptation if it does not satisfy the adaptation conditions?

(b) How can we measure the sensitivity of adaptation to change? Possible sources of change include:

(i) perturbation of the coefficients

(ii) perturbations of the network structure, i.e. adding/removing network connections

(c) How well does the adaptation mechanism react to changes in the network structure, such as permuting the system topology as well as input/output structures, similar to the work performed in [14] and [21].

Comparison of Methods for Inferring Dynamics

In my work, I have both seen and employed various methods for inferring dynamics. One area of exploration is to consider the benefits and costs for each method and where one may prefer to use a specific method.

Research Task 2. *(a) What are the limitations for recovery algorithms?*

(b) How do the algorithms behave in the presence of conservation laws and special network structures? Are they able to infer the dynamics to include these results or will they produce a reduced system?

(c) What is the computational complexity of the problems?

Biological Modeling

The overall goal of this research is to apply it to biological data and test a protein-protein interaction network for adaptation. Thanks to a collaboration with Dr. Mariaelena Pierobon at the Center for

Applied Proteomics and Molecular Medicine (CAPMM), GMU, I have access to data for the mitogen-activated protein kinase (MAPK) pathway (see Figure 1). The MAPK signaling pathway is a key regulator of different cellular processes, such as gene expression and cellular growth. Furthermore, it is a sub-network for the Epidermal Growth Factor Receptor (EGFR) pathway, both of which have been shown to play a role in various cancers when dysregulated. Even after intense research scrutiny, deciphering mechanisms of action and regulation of the MAPK pathway still remains a challenge from a biological perspective.

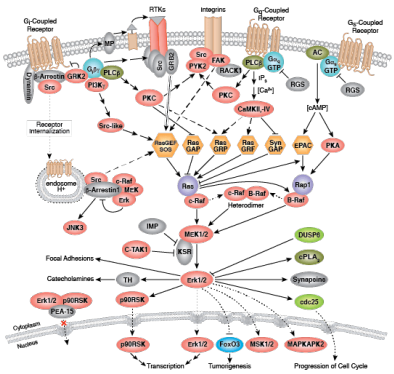

Figure 1: The full MAPK Pathway. [1]

Research Task 3. *Infer the dynamics of the MAPK pathway. Of particular interest are the following:*

(a) Infer the dynamics for the MAPK pathway in the presence of DMSO and Selumetinib.

(b) Infer dynamics for mutated cell lines and non-mutated cell lines.

In each model, we would like to:

(i) Identify any potential source of singularities within the system.

(ii) Test for adaptation.

While there is a growing interest in exploring these dynamic interactions from a biological prospective, modeling of systems such as the MAPK pathway presents multiple challenges due to a lack of reliable data and the combinatorial increase in complexity when considering the full-scale reaction network. This work proposes an innovative data-driven multidisciplinary approach that combines quantitative experimental measurements of network dynamics with mathematical modeling to devise novel multi-scale pattern-oriented methods for dissecting and understanding signal transduction-based mechanisms in complex biological samples. This will push the boundaries of sparse dynamics identification methods by enhancing them with mesoscale network characterization techniques. This novel framework has a potential to find broad applicability for illuminating basic molecular mechanisms associated with different pathological processes, uncovering target-able interactions within these networks, and predicting network adaptation mechanisms and response to perturbations with applications across different fields of biomedical disciplines and mathematics.

References

[1] G-protein-coupled receptors signaling to mapk/erk, 2002. URL `https://www.cellsignal.com/pathways/receptors-sign`

[2] U. Alon, M. Surette, N. Barkai, and S. Leibler. Robustness in bacterial chemotaxis. *Nature*, 397, 1999. doi: https://doi.org/10.1038/16483.

[3] R. P. Araujo and L. A. Liotta. The topological requirements for robust perfect adaptation in networks of any size. *Nature Communications*, 9(1), 2018. doi: https://doi.org/10.1038/s41467-018-04151-6.

[4] R. P. Araujo and L. A. Liotta. Design principles underlying robust adaptation of complex biochemical networks. *Methods in Molecular Biology*, in press, 2020. doi: https://doi.org/10.1101/2020.09.21.307140.

[5] N. Barkai and S. Leibler. Robustness in simple biochemical networks. *Letters to Nature*, 1997. doi: https://doi.org/10.1038/43199.

[6] S. L. Brunton, J. L. Proctor, and J. N. Kutz. Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proceedings of the National Academy of Sciences*, 113(15):3932–3937, 2016. ISSN 0027-8424. doi: 10.1073/pnas.1517384113. URL `https://doi.org/10.1073/pnas.1517384113`.

[7] M. Dehmer, F. Emmert-Streib, Z. Chen, X. Li, and Y. Shi, editors. *Mathematical Foundations and Applications of Graph Entropy*. John Wiley Sons, Ltd, 2016. doi: https://doi.org/10.1002/9783527693245.

[8] A. Dickenstein. Biochemical reaction networks: an invitation for algebraic geometers. *Mathematical congress of the Americas, Contemporary Mathematics*, 656, 2016. doi: https://doi.org/10.1090/conm/656/13076.

[9] J. E. Ferrell. Perfect and near-perfect adaptation in cell signaling. *Cell Systems*, 2016. doi: https://doi.org/10.1016/j.cels.2016.02.006.

[10] L. Goentoro, O. Shoval, M. Kirschner, and U. Alon. The incoherent feedforward loop can provide fold-change detection in gene regulation. *Molecular cell*, 36, 2009. doi: 10.1016/j.molcel.2009.11.018.

[11] O. Hoeller, D. Gong, and O. D. Weiner. How to understand and outwit adaptation. *Developmental Cell*, 28, 2014. doi: 10.1016/j.devcel.2014.03.009.

[12] K. Kaheman, J. Kutz, and S. Brunton. Sindy-pi: a robust algorithm for parallel implicit sparse identification of nonlinear dynamics. *Proc. R. Soc.*, 476(2242), 2020. doi: http://doi.org/10.1098/rspa.2020.0279.

[13] C. L. Lawson and R. J. Hanson. *Solving Least Squares Problems*. SIAM, 1995. doi: 10.1137/1.9781611971217.ch23.

[14] W. Ma, A. Trusina, H. El-Samad, W. A. Lim, and C. Tang. Defining network topologies that can achieve biochemical adaptation. *Cell*, 2009. doi: https://doi.org/10.1016/j.cell.2009.06.013.

[15] N. Mangan, S. Brunton, J. Proctor, and J. Kutz. Inferring biological networks by sparse identification of nonlinear dynamics. *EEE Transactions on Molecular, Biological and Multi-Scale Communications*, 2, 2016. doi: https://doi.org/10.1109/TMBMC.2016.2633265.

[16] T. Oellerich, M. Emelianenko, L. A. Liotta, and R. P. Araujo. Biological networks with singular jacobians: their origins and adaptation criteria. *Submitted*. doi: https://doi.org/10.1101/2021.03.01.433197.

[17] T. Oellerich, M. Emelianenko, M. Pierobon, and E. Baldelli. Learning biological networks dynamic from data. *In preparation*, 2022.

[18] S. Sasagawa, Y. Ozaki, K. Fujita, and S. Kuroda. Prediction and validation of the distinct dynamics of transient and sustained ERK activation. *Nature Cell Biology*, 7, 2005. doi: https://doi.org/10.1038/ncb1233.

[19] G. Shinar and M. Feinberg. Structural sources of robustness in biochemical reaction networks. *Science*, 327, 2010. doi: 10.1126/science.1183372.

[20] O. Shoval, U. Alon, and E. Sontag. Symmetry invariance for adapting biological systems. *SIAM Journal of Applied Dynamical Systems*, 2011. doi: https://doi.org/10.1137/100818078.

[21] Z. F. Tang and D. R. McMillen. Design principles for the analysis and construction of robustly homeostatic biological networks. *Journal of Theoretical Biology*, 2016. doi: https://doi.org/10.1016/j.jtbi.2016.06.036.

[22] J. West, G. Bianconi, S. Severini, and A. Teschendorff. Differential network entropy reveals cancer system hallmarks. *Scientific reports*, 2, 2012. doi: 10.1038/srep00802.

[23] T. Yi, Y. Huang, M. Simon, and J. Doyle. Robust perfect adaptation in bacterial chemotaxis through integral feedback control. *Proceedings of the National Academy of Sciences*, 2000. doi: https://doi.org/10.1073/pnas.97.9.4649.

toelleri@gmu.edu