# NSF Workshop: Systems and Control Theory for Synthetic Biology

Domitilla Del Vecchio

November 4-5, 2021

## 1 Introduction

The engineering of biological systems, whether within the cell (bacterial or mammalian) or in cell-free extracts, can be applied to a number of areas, ranging from diagnostic, biosafety (i.e. biothreat detection), and materials in the short term, to targeted drug delivery and the engineered microbiome in the medium term, to engineered living materials and reprogrammed patient-specific cells to repair wounds and cure diseases in the long term [1–4]. Modular and hierarchical design has been a highly convenient and successful approach to design complex systems in electrical and computer engineering, wherein a system is described as the composition of simpler subsystems whose properties are fully understood and maintained after composition. As they move from "lower" to "higher" layers, such as from the physical layer of parts, to logic gates, to computers, to networks, system designers typically ignore details concerning systems at the lower layers while designing any given layer, focusing only on suitable input/output connectivity. This is possible because subsystems often contain compensation mechanisms (such as through feedback) that maintain desired input/output properties, thus providing convenient and simplified abstractions for design.

Since the creation of the first two synthetic genetic circuits in the year 2000 [4], the field of synthetic biology has been strongly influenced by the electrical and computer engineering community, and these modular and hierarchical design abstractions have been proposed as a starting premise for design [5, 6]. The input/output view of genetic modules, often as static and digital gates, has permeated the field and has provided convenient abstractions to reason about composition, as demonstrated by design tools such as Cello and COMET [6, 7]. On the other hand, there is clear evidence today that the input/output properties of defined genetic modules are not maintained upon composition. Instead, they are highly dependent in rather surprising ways on the genetic, intra-cellular, extra-cellular, and environmental context through a variety of interactions, such as DNA supercoiling, retroactivity, resource sharing, and growth rate feedback to name a few [8–19]. This is not surprising, given that there is nothing within the genetic module's design that should guarantee "robustness to context," which is instead taken care of in electrical and computer engineering through suitable internal compensation mechanisms. Indeed, many of the failures that we experience when building synthetic genetic circuits can be often traced back to the fact that a working system may not work anymore as intended once a new component is added to it. This forces one to redesign existing parts each time a new component is added, and this requirement

leads to a largely monolithic and hence not scalable design process. There is no design formalism today that explicitly manages context-dependence of genetic modules. Thus, the questions remain: how can we move forward and reach a robust and scalable design approach that guarantees quantitatively predictable outcomes? To what extent can we adapt the design abstraction hierarchy used in computer engineering to biological systems? These and related questions have been lingering in the community in the past few years, and are at the core of this workshop.

Systems and control theory can offer a powerful formalism to reason about what is today vaguely referred to as context-dependence. As an example, the moment that we consider resource loading, one of the contributors to intra-cellular context dependence, as a disturbance input on a module of interest, then we can analyze and mitigate it just like we analyze and design robustness to disturbance inputs and uncertainty in any engineered system [20, 21]. Of course, the details of the physical domain lead to mathematical model descriptions that carry substantial differences from what we have in more traditional engineering domains and often require new theory and deeper understanding of biomolecular interactions [22, 23]. Additionally, in many engineering systems, a module interacts with its surroundings through well-defined media, so the connectivity with the context is relatively simpler to capture and isolation can be often aided by using physical compartments. None of this is the case for biomolecular circuits running inside cells or in cell-free extracts, and achieving full insulation may even be suboptimal. A natural question is, thus: to what extent existing formalism and mathematical frameworks in systems and control theory are applicable to engineering biology or need to be reinvented?

**Workshop structure.** To address these questions, this two-day workshop brought together leading researchers worldwide, who work at the intersection of synthetic biology and systems and controls. The workshop structure and schedule is reported in Figure 1. Specifically, the morning of the first day featured longer, introductory, and more provocative talks from a subset of the participants, meant to stimulate discussion. During the following working lunch, a different subset of participants presented more technical shorter talks focused on specific problems and questions. After this, the participants were organized into fours subgroups, each associated with a theme highly represented in both the Q&A and the on-line collaboration document, which was shared with all the participants and available for editing throughout the whole workshop. Then, participants regrouped, and each subgroup discussed the assigned theme and prepared a presentation, for which a template was given and included: (i) challenges discussed; (ii) approaches that may be used from systems and control; (iii) where available approaches need re-thinking; and (iv) how other disciplines may complement systems and control. Day 2 started with morning short talks, again with an aim of triggering discussion, then additional (or same) themes were assigned to each of four groups for discussion in the break-out session. At the end of Day 2, for each of the eight groups over the two days, a lead person had 10 minutes to go over the presentation of the discussion outcomes, including Q&A, with the exception of the group discussing the "Community Building" theme, which had 20 minutes allocated.

This report is the summary of the group presentations' content, the collaboration document, and the Q&A during the day 2 final discussion. This summary is focusing for the most part on the following highly debated topics: modularity and scalability, robustness, pedictability, multicellular systems, applications, and community building.

**Workshop participants.** Neda Bagheri (U. Washington, Seattle); Enoch Yeung (UCSB); Eduardo Sontag (NEU); Hana El Samad (UCSF); Murat Arcak (UC Berkeley); Brian Munsky (Colorado State U.); Jeff Hasty (UCSD); Howard Salis (Penn. State U.); Leonidas Bleris (UT Dallas);

| **Day 1: Thursday November 4** | |
|---|---|
| **Morning: Introductory talks and themes discussion** | |
| 8:00AM-9:00AM | Breakfast pick-up |
| 9:00AM- 9:05AM | *Domitilla Del Vecchio* (MIT): Opening remarks |
| 9:05AM- 9:15AM | *Walter Rance Cleaveland II* (NSF, CCF Division Director): Welcome |
| 9:15AM- 9:30AM | *Domitilla Del Vecchio* (MIT): Workshop introduction and Q&A |
| 9:30AM-10:00AM | *Hana El Samad* (UCSF): Build to understand |
| 10AM-10:30AM | *Mustafa Khammash* (ETH): Nothing so practical as a good theory: Theory and methods for practical synthetic controller design |
| 10:30AM-11:00AM | Coffee break |
| 11:00AM-11:30AM | *Eduardo Sontag* (NEU): Systems and control for composition |
| 11:30AM-12:00PM | *Jeff Hasty* (UCSD): Dynamics and control in small ecologies |
| 12:00PM-12:30PM | *Ophelia Venturelli* (U. of Wisconsin): Dynamics and control in engineered microbial communities |
| **Afternoon: Short talks and themes discussion** | |
| 12:30PM-12:45PM | **Lunch box pick-up** |
| 12:45PM-2:20PM | **Working lunch: Short talks from participants (10 minutes each)** - *Diego Oyarzun* (U. of Edinburgh (UK)), *Neda Bagheri* (U. Washington, Seattle), *David Ross* (NIST), *Elisa Franco* (UCLA), *Brian Munsky* (Colorado State U.), *Enoch Yeung* (UCSB), *Andras Gyorgy* (NYU, Abu-Dhabi) |
| 2:20PM-3:00PM | **Organization of break-out sessions:** extraction of main research topics emerging from talks, starting with robustness, scalability, modularity, and systematic design - grouping into four teams - electing a team leader for each team - break into groups |
| 3:00PM-5:00PM | **Groups discussions on the main themes** - each team leader will collect ideas into a power point presentation |
| 5:00PM-5:30PM | Summary and adjourn for the day |
| **Day 2: Friday November 5** | |
| 8:00AM-8:45AM | Breakfast pick-up |
| 8:45AM-9:00AM | *Domitilla Del Vecchio* (MIT): Opening remarks for day two |
| 9:00AM-11:40AM | **Short talks from participants (10 minutes each)** - *Josh Leonard* (North Western U.), *Murat Arcak* (UC Berkeley), *Howard Salis* (Penn State U.), *Leo Bleris* (UT Dallas), *Marcella Gomez* (UC Santa Cruz), *Xiaojun Tian* (Arizona State U.), *Chris Myers* (U Colorado, Boulder), *Xiao Wang* (Arizona State U.) |
| 11:40AM-12:30AM | **Organization of break-out sessions:** shuffling groups - create four new teams - assign new leader for each team - break out into four new groups - lunch pick up |
| 12:30PM-2:00PM | **Working lunch:** Groups discussions on the main themes - each team leader will collect ideas into a power point presentation |
| 2:00PM-3:00PM | Team leaders finalize presentations |
| 3:00PM-4:30PM | All team leaders (~8 teams) each give a 10-minute presentation on the group discussion outcomes |
| 4:30PM-5:00PM | Summary, follow-ups, and meeting adjourn |

Figure 1: **Workshop schedule.**

Ophelia Venturelli (U. of Wisconsin, Madison); Chris Myers (U Colorado, Boulder); Marcella Gomez (UC Santa Cruz); Andras Gyorgy (NYU, Abu Dhabi); Diego Oyarzun (U. of Edinburgh (UK)); Xiaojun Tian (Arizona State U.); Xiao Wang (Arizona State U.); Josh Leonard (Northwestern U.); Elisa Franco (UCLA); Mustafa Khammash (ETH); Aura Gimm (OSD); Mitra Basu (NSF); Sankar Basu (NSF); David Ross (NIST); Ron Weiss (MIT); Domitilla Del Vecchio (MIT).

# 2 Modularity and scalability

Modularity is generally intended as the property by which the input/output behavior of a system does not change when this is composed with other systems. Modularity therefore allows to predict the behavior of a complicated system by that of subsystems through well defined input/output

connections. The extent to which natural cellular networks are modularly organized has been one of the most vexing questions in systems biology and has been investigated for decades by leading researchers in the field [24–27]. For engineering, modularity is highly convenient as it allows to design larger systems by incrementally adding new components without ever worrying about whether the functions of the pre-existing subsystems are impacted by the addition of new ones. It is thus a critical requirement for design processes whose complexity scales with the number of subsystems. Here, we report the main discussion points raised during the workshop and the break-out group discussions.

**Challenges discussed.** Leading questions were whether biology is modular at all, what is a definition of modularity, and whether for design we really need strict modularity and if not what would be a suitable form of modularity that could facilitate scaling up design practices. Related questions include whether genetic parts themselves have any hope to be modular and if not how to handle this [18]. Indeed, strict modularity enables scalable and rational design, so it is highly convenient, but it can also be highly inefficient. In fact, designing for insulation has been shown to be possible in engineering biology but also costly in terms of parts, energy, and cellular resources [28, 29]. We, in engineering, are perhaps trying to enforce modularity too strictly; yet, it may not be needed in order to obtain robust and quantitatively predictable functions. After all, biology has reached highly sophisticated designs that are robust, reliable, and predictable, while also being highly efficient, without strict modularity. Although, nature took the course of evolution to reach today's robustness and reliability; by contrast, we are trying to rationally engineer robust, predictable, and reliable biological systems within months or years.

**Where available approaches need re-thinking.** There was an overall consensus among workshop participants that synthetic biology does not need to look like electronics (VLSI), even though much of its inspiration may be derived from it, at least conceptually. Clearly, new theory and formalisms are required to help us design functions without strict modularity, in a way that allows to design systems "embracing" contextual dependencies instead of fighting against them to reach full module insulation. To this end, though, it is required that we become more proficient in quantitatively predicting system's behavior while accounting for contextual interactions.

**Rationale for an alternative approach.** While at the functional level, some form of modularity is desirable and convenient, this may not be the case at the level of physical parts and their interactions, as multiple functions can be realized by the same parts. Evidence is emerging that a systems design theory can be obtained where function is achieved without imposing strict modularity. This is consistent with a one-part/multiple-roles paradigm, such as within the notions of conserved core processes and weak regulatory linkages [27], with the approximate internal model principle from control theory [30], and with celebrated results from electrical network analysis such as Thévenin's theorem. Within a new paradigm of composability, the notion of constrained structure, as opposed to strict modularity, could be the driving principle. More generally, an approach that combines functional modularity, convenient for design, with the more lenient structural design paradigm is needed. Specifically, since design is performed through several iterations, at the first iteration, a designer could assume a modular design framework where each module ideally achieves a strictly delineated subfunction and modules have strictly unidirectional interactions. At a second iteration, the designer could collapse the strict boundary between modules, allowing parts to be shared and information to flow in both directions such as with retroactivity [31]. Within this framework, the new looser interactions between modules can be determined based on yet to be developed theory that preserves the function while blurring the boundary between modules. Such

a theory could be based on optimization, network theory, stochastic analysis such as a stochastic behavioral approach for biomolecular systems à la Willems [32]. In this framework, retroactivity, loading, parts-sharing, may become features that enable efficiencies, not deleterious effects to be designed away.

**Additional thoughts.** What is then modularity? The answer may depend on the level of abstraction. For example, systems that are modular at the function level may not be so at the parts level. Indeed, modularity in biology seems to depend on scale: modularity is seen at the large scale (organs), while less modularity is observed at the intra-cellular level but it emerges again at the level of protein domains, for example. Many times, what is at a first sight considered modular, it is often not so after a closer inspection (example: Syn-notch antigen/receptor/TFs are billed as modular but they are not quite so). Strict modularity is expensive, while weak regulatory linkages are more efficient. Yet, to start a design process, it is still convenient to assume strict modularity and then iterate on the strict boundaries and allow weak interactions. Rational design performed this way can bring us close to the final design objective, but complementary approaches, such as directed evolution [33], may help refine the design and reach the optimum. Overall, modularity may just be a convenient assumption made at the early stage of design at the functional level, but not enforced strictly at the implementation level. This way, modularity becomes more of a convenient tool as opposed to a property that needs to be enforced on the systems that we are engineering. In circuit design, the notion of modularity may be much stronger than what is required or even possible in engineering biology, that is, infinite input impedance and zero output impedance. After all, Thévenin's theorem tells us how to compose complex electrical circuits without infinite input impedance or zero output impedance. Another important aspect to consider is temporal dynamics. Although sometimes subsystems may look modular by their steady state behavior, they may actually not behave modularly when considering temporal behavior. Overall, structural modules could be entities that are reused in different contexts, but that have a different quantitative behavior depending on the context in which they are inserted. We therefore need to determine what an acceptable notion of modularity is for engineering biology and a formalism for design that is based on it.

# 3 Robustness

In general, robustness of a system to some form of perturbation is the property that allows the system to maintain a desired behavior despite having unexpected influences (i.e., perturbations or disturbances) act on the system itself. Every engineering system incorporates mechanisms that ensure robustness to perturbations. For example, the thermostat controls the heating of a room such that the room remains at a desired temperature, independent of the outside temperature and of whether it snows or is a sunny day. The cruise control system allows a vehicle to travel at a user-set speed, independent of how many passengers the vehicle carries and of the slope of the road. Indeed, control system design is largely centered around the problem of making quantitative performance specifications robust to uncertainty in parameter values, unmodeled dynamics, and perturbing inputs that cannot be controlled nor measured [34]. Biological systems also have natural robustness mechanisms built-in, which are critical for healthy physiology. For example, the human body's temperature is regulated to stay between 97 F and 99 F, independent of whether we are running outdoor on a sunny day or are standing at the bus stop in a cold winter night. In this sense,

robustness generally allows a system to perform as specified despite significant uncertainty around it.

Therefore, to some extent, the problem of enforcing some form of modularity when engineering biology is closely related to the problem of making modules somewhat robust to the influences of their environment (context). More generally, genetic circuits built today only function as predicted in narrowly controlled laboratory conditions, and even then, if the cellular stresses or the genetic background of the cell change due to insertion of other modules, the circuit will most likely cease to function as desired. With this fragility, it is difficult to envision a future where engineered organisms will be deployed in the field, whether for environmental biosensing, in-gut applications, or within patients' bodies to cure illnesses.

**Challenges discussed.** In the context of engineering biology, the main challenges include carefully determining *which* system properties need to be robust and *to what* perturbations these properties need to be robust. Also, robustness needs to be considered not only for a given physical system's property but also for the design process itself – if every biological solution requires its own a unique one-off set of tools, then progress in biotechnology will be unbearably slow. At a high level, there is a general need for robust design methods that incorporate or mitigate the effect of context on biomolecular systems. It may be more important and efficient to focus on just a handful of specific key system variables that we want to make robust (e.g., uniformly throughout entire cells or in critical spatially resolved compartments), since all the other variables may be mostly dependent on them and less directly dependent on the perturbation/environment. This may be similar as to what already occurs in natural systems: the body temperature is tightly regulated because there are a large number of physiological responses that depend on it and in fact temperature disregulation (hyperthermia or hypothermia) can lead to significant health threats. A specific standing grand-challenge is how to engineer robustness to extra-cellular factors, such as nutrients and environmental conditions (temperature and pH). In fact, today's engineered cells often can perform as predicted only when in narrowly maintained laboratory conditions and completely fail if these vary slightly. How to engineer robustness to differential growth rates for genetic circuits that are distributed across engineered strains, or perhaps even between different spatial compartments within individual cells, remains also a significant challenge. In general, there appear to be a lack of a unifying theory of robustness that spans across scales of organizations (from intra-cellular to inter-cellular and extra-cellular) and accounts for weak regulatory linkages to determine the critical knobs that should be kept homeostatic.

At a higher level, there is a compelling need for engineering processes that are robust to lack of information and even wrong, or non-reproducible, information. Many of the interactions involving engineered modules in the cell will be hard or impossible to completely untangle, so how to establish a design process that withstands this lack of information? We may need new modeling frameworks that are robust to uncertainty especially in the connectivity among subsystems, stepping away from more convenient, dogmatic, modeling frameworks and instead embrace such an uncertainty in the design phase. We need descriptors and checkable quantitative standards for identifying where and why a specific design has failed and under what conditions could its performance be rescued. We need more accessible and inter-operable descriptions of robustness of measurements, laboratory outcomes, and move from qualitative descriptions to precise quantitative characterizations. To this end, it would be useful to develop calibration approaches that allow quantitative comparisons of heterogeneous data sets (e.g., from different labs or using different measurement devices) that are quantitatively different (even under specific controls), but that can

reliably identify qualitative and quantitative trends. It may be helpful for community to reconsider carefully and quantify past failures and successes to ask if there are precise principles that improve (or common mistakes that impede) performance of genetically engineered processes. As an example, if we were to re-design a genetic toggle switch, what would be the design process to augment the original, more fragile, toggle switch design, and enable a module that, on the first try, would immediately be robust to temperature changes, to growth conditions variability, and to the variable presence of other modules in the cell (genetic background)?

An additional challenge discussed is the tradeoff between robustness and malleability. Specifically, we need to balance between responding to inputs of relevance while maintaining low sensitivity to other stimuli that may harm function. How to establish frameworks to design this selective robustness within systems? Furthermore, robustness is often only considered for steady state level, and referred to as homeostasis. It is yet unclear how to design robust dynamic behaviors such as oscillators, signal processing units, and multi-stable systems. Stochastic approaches are also required if we want to engineer circuits in low copy to ensure robustness of overall system function to resource burden.

**Tools that may be used**. Uncertainty quantification (UQ) approaches from statistical analysis could be useful to characterize robustness to parameter variations, since parameters are largely unknown and highly variable. Accordingly, UQ-informed experiment design approaches are needed to reduce uncertainty, if not with respect to parameters then with respect to predictions that may depend on those parameters. Generally, exploiting notions of entrainment seems promising, wherein constituting agents or subsystems may each be allowed to be poorly robust and highly variable on their own, but overall combine together to produce an emergent (e.g., population-level) behavior that is robust and predictable. Related to this, multi-agent control for either reasoning about multi-cellular systems (distribution of function, metabolic burden) or multi-module systems are mostly available but underutilized (e.g., decentralized control). Adapting them to the nonlinear and highly heterogenous dynamics of the composing agents (or subsystems) will require significant effort. Stochastic control approaches may also be promising, where the focus is to control a distribution rather than the means, and can handle systems functioning at low-copy numbers, which could be advantageous for alleviating cellular burden. With respect to cellular burden specifically, we mainly have approaches today to re-distribute resources but do not have approaches for modulating resource production to match demands. This, in turn, is the general approach in other engineering fields such as in the control of the power grid.

**Where current approaches need rethinking.** At a high level, there is a concern that pursuing research on robustness may not be well aligned with current trends within research groups because current focus is often placed on short-term rewards, and more narrowly focused projects appear to be better able to reach a goal of publication in high-profile journals (e.g., Nature, Science, or Cell) that may give preference to catchy ideas rather than to thorough studies of long-term unresolved problems. Careful consideration of robustness will likely require substantially longer time investments and more technically-oriented research. Although this research may lead to high impact outcomes that the community will utilize at large, such long-term studies may require larger collaborative teams and may be more difficult to publish in high-profile journals. On the technical side, classical robust control methodologies [35] are focused on exact parameterization of control synthesis, which is not useful for biologists. Therefore, new methods will be needed that lead to robust designs, without the requirement for explicit parameterizations of uncertainty and with the acceptance that the control components that we can design are also subject to large parameter

uncertainty. Remarkably, natural systems that implement (integral) feedback control (chemotaxis) are extremely accurate, precise, and repeatable. Yet, if one looks at their mathematical models at the scale of molecular reactions, every part of the system is subject to parameter uncertainty and noise. Learning from how biology can design precise and accurate functions out of uncertain and variable components would be highly valuable for engineering biology.

# 4 Predictability

Overall, our current ability to predict system's behavior in a quantitative way is still severely limited by poor understanding of how the properties of parts, modules, and systems, are affected by their genetic context [18,19], intra-cellular conditions and connectivity [8–10,15,17], and extra cellular environment [36]. To handle this uncertainty on interactions with surrounding systems, two orthogonal, yet not mutually exclusive, approaches have been discussed: (a) engineering insulation from surrounding context so that prediction of behavior becomes simpler or (b) substantially improving the modeling framework to enable prediction of complicated interactions with surrounding elements. Approach (a) has been employed especially for predicting the behavior of modules, that is, including negative feedback compensation mechanisms [10,13,20,21,28], and approaches closer to (b), for predicting behavior of parts, are starting to arise [18]. As outlined in Section 2, approach (a) alone may not be the most effective or optimal way of engineering biology and a mix of the two approaches may be required in the end. Here, we outline some of the details discussed with this respect during the workshop.

**Challenges discussed.** Although quantitative prediction is still a challenge, predicting the qualitative behavior (trend) of a system is today largely achievable through suitable mathematical models. A key question then is when is quantitative predictive accuracy needed versus qualitative model prediction. This issue is tightly linked to the definition of system's performance that we take, which should be re-usable and operationally relevant. With respect to quantitative prediction accuracy, one question regards the role of machine learning (ML) to reduce the uncertainty of biophysical models, in a way that allows system composition and prediction in different contexts. Reversely, within a ML model, biophysical constraints could be included as a prior in order to bias emerging models towards those that are compatible with known physics. How to use ML or biophysics models to merge high throughput and low fidelity experiments (from random construct design, *in vitro* experiments) with low throughput high-fidelity experiments (from rational design *in vivo*) is also still a major question. Tightly linked to the issue of quantitative prediction accuracy is the issue of measurements: we often have inaccurate, indirect, and sparse measurements, wherein only a subset of the system's state can be measured, through some proxy, and also with population-level resolution and less commonly with single-cell resolution. With this respect, there is also a challenge about how to integrate heterogeneous data streams (e.g., RNAseq bulk/single-cell, flow cytometry, and smFISH). Overall, a central challenge remains how to incorporate a model of the environment of a circuit of interest in a way that is generalizable to new designs and situations and that can allow quantitative prediction accuracy. With this respect, redundancy and feedback compensation may be employed also as a way to make predictability simpler, especially on longer time scales of operation.

Furthermore, although the emphasis in the field of synthetic biology has mostly been on predicting steady state and deterministic behavior in the form of input/output characteristics, an in-

creasingly important challenge is to predict temporal, dynamic, and stochastic behavior, as these are critical in systems where temporal specifications are important (see Section 6) or where the behavior of the system should be multi-modal at the cell population level. Specifically, how to use cellular noise and temporal dynamics as a phenotypic observation (fluctuation fingerprints) that can discriminate between available hypotheses is still unclear. Noise, especially, is typically treated as something to defeat, but in reality it could be exploited for inferring critical interactions and also for design. For example, how to use data from heterogeneous cell populations (individual cells within a population or spatial heterogeneity within single cells) to infer temporal dynamics of circuits is still unclear. Finally, there is a major issue about the fact that often distinct models have equivalent descriptive and predictive power but have entirely different mechanistic interpretation. How to reconcile this within a systematic and rigorous predictive modeling framework? Would there be an issue of model discrimination that needs to be addressed? With this respect, it is still not clear how to choose the right scale and granularity of a model and required data to validate it for a given application at hand. There is currently no systematic approach to expand a model to include more mechanisms and interactions in such a way that the behavior remains compatible with the specifications as one moves through the different stages of design.

**Tools that may be used.** ML and biophysical models could be merged together. Specifically, ML could offer improved learning across biological parts, systems, and domains; however, improved interpretability of ML model constants in terms of physical parameters would be required to be useful for design. Most importantly, we need ML models that address synthetic biology design and prediction questions and allow composition. On the other hand, biophysics models can explicitly incorporate mechanisms from molecular interactions (DNA, RNA, and protein), to gene expression (transcription, translation, and decay), to metabolism (kinetics & fluxes). Agent-based models could then be better leveraged for design of multi-cellular systems, although analysis and interpretation approaches that make these models suitable for forward engineering, as opposed to only simulation, would be required. Methods such as massively parallel reporter assays, which characterize 1000s to millions of genetic system variants quickly, could be better used to identify interactions among subsystems and to create a predictive modeling framework that scales with system size. Other tools that could be better leveraged to identify input/output responses is optogenetic control and, more generally, *in silico* control of genetic circuits [37].

Tools from the field of uncertainty quantification could be employed to determine the extent of parameter variation that is consistent with experimental data. Specific tools include Bayesian analysis (including approximate Bayesian computing), Monte Carlo sampling, and bootstrapping/cross validation. Model-driven experiment design is also a concept that is currently under-utilized. Specific tools that could be adapted include Fisher information matrices, Bayesian experimental design, variance minimization, sensitivity analysis methods, and Latin hypercube sampling methods. Analytical tools such as bifurcation and stability analysis as well as frequency domain noise reduction analysis to explore parameter space could be better utilized and adapted. A popular concept in learning and adaptive control is persistence of excitation, which ensures that the input/output data is sufficient to identify the required model parameters [38]. How to generalize similar design approaches to allow for on-line identification and control of complex biological processes would be highly valuable. Methods to extend work on dynamics of distributed systems, decentralized and cooperative control, to multi-module genetic circuits could also be useful. Similarly, fluctuation-dissipation theorems, modeling based on the chemical master equation and the stochastic simulation algorithms could be better leveraged to understand the noise in the context of control

of interaction between synthetic gene circuits and host cells.

**Where current approaches need rethinking.** Although ML has a plethora of potential models and learning algorithms for parameters, ML approaches are currently not well suited to address design questions in engineering biology. They would need to be customized to recapitulate and engineer biological functions, resulting in human-understandable, interpretable, models with high accuracy and high generalizability. These models should further allow some form of composability to enable scalable, as opposed to monolithic, design of biological systems. Specific examples include ML frameworks that allow to map sequence to function in a composable way, that direct experimental design, and select among design options for high-success outcomes. Although neural networks (NNET) may not be the best modeling framework for engineering biology, the concept of physics-informed ML, developed for other engineering problems, may be leveraged [39]. In general, we need to improve model development & predictable design by storing experimental protocols, sequences, measurements using minimum and domain-adaptive information encoding. Methods are needed to improve the design of "open-loop" systems with well-controlled behaviors and to design optimal controllers to correct system's behavior.

**How tools from other disciplines can complement systems and controls.** In general, we should think of revisiting the foundations of genetics and gene regulation altogether to ensure that whatever modeling framework we adopt is really compatible with the critical physical constraints that we know of. This could lead to create and test better models of gene regulation to carry out improved design of genetic systems. Accordingly, building databases of mechanistic features (DNA sequence to DNA rigidity) that can be re-used to predict biological functions across different systems and domains (as input features into machine learning models) could be valuable. Furthermore, text-mining for automating the reading of papers could be highly useful to ensure sufficient biological knowledge when creating a model.

# 5 Multi-cellular systems

Although much of the field has focused on engineering genetic circuits within single cell strains, the design of engineered multi-cellular systems has been a growing field in the community. The basic problem is to engineer genetically different cellular strains that, in some form, interact with one another to obtain a population-level behavior. This is particularly relevant in applications such as for the engineered microbiome [40] and in future applications of mammalian synthetic biology to *in vivo* directed differentiation. A major challenge in designing multi-cellular systems where two or more cell strains co-exist is the issue of differential growth, which causes the faster growing strains to take over the population and hence harm the desired overall function [36]. Here, we summarize some of the discussion topics in this area.

**Challenges discussed.** Today, a grand challenge still remains to determine the most effective way to coordinate, control, and engineer more than 2 species such that they can co-exist together. One reason why we may want to have more than one strain in the population is to divide the labor across multiple hosts. However, there is still a question of whether the task of ensuring co-existence is actually harder than figuring out how to manage the cellular load of having one species do all the tasks. How to distribute the load optimally across multiple strains or cell types is also still unanswered. Accordingly, what the relative advantages are of having two (or more) different genotypes of cells versus just one genotype that can dynamically adopt more than one

state are still unclear. Indeed, direct ratiometric control of population sizes versus intrinsic control of growth rate are two different approaches to manage metabolic burden and the advantages of one over the other are not very clear, and ultimately a combination of the two may be needed. How to engineer a single (or limited number) of cells to regulate/control the behavior of larger, diverse populations of un-engineered cells is also a major question. For example, one could introduce an engineered immune cell to control the coordinated behavior of a non-engineered population of immune cells; similarly, one could introduce engineered microbes in the gut to control the un-engineered diverse microbiome. Related to these, the following questions arise: (a) What data/understanding/characterization is required to enable this goal? (b) What mathematical or computational models could predict and characterize emergent population dynamics in a way that is useful for design? (c) What methods would enable control of such emergent dynamics? (d) How to identify/validate control handles? These would be high dimensional systems with relatively lower number of observables. How do we reason about controllability, observability, and reachability in these systems [34]? Related to this, how do we better understand the impact of cell-cell communication on these properties? Ultimately, a grand challenge remains finding ways to design robust systems that operate in a highly heterogeneous and multiclonal population. Finally, functionality in heterogeneous environments is a general issue difficult to handle, whether in bacteria or mammalian cells, wherein the cell itself often evolves, through genetic mutations (frequent in bacteria) or through cell-state changes such as in cellular differentiation.

**Tools that may be employed from systems and control theory.** Among existing tools that could provide a conceptual basis for reasoning about the above questions, there are those established by the field of decentralized and cooperative control [41]. These could be leveraged to coordinate the dynamics of independent individual agents in a population that is largely not controllable. This is similar to what found in coordination of multiple drones, multi-vehicle systems, and swarms of robots. There is also the general pairing problem where one needs to identify control handles in multi-cellular systems, wherein adding a resource (input) to a microbial community shifts the relative competition for resources and niche availability to control the frequency of the species (output). Accordingly, as done for Boolean node control, one can design a system to have specific attractors. That is, one could define a "base case" for frequencies of different species of microbes in a population, such that the population might shift to respond to a perturbation, but it, by definition, relaxes back to the base case distribution of species frequencies.

**Where available approaches need rethinking.** Coordinating the behavior of spatially distant cells will likely need completely different approaches and could be enabled by using computer-in-the-loop methods. For example, optical sensing and optical control can reduce the timescales required for distant cells to communicate and coordinate. But formulating a control objective/strategy to align with an application is more complicated and needs rethinking. Controlling complex networks of cells might require defining and implementing multi-parametric control objectives and could benefit from optimal control, but needs adaptation. Finally, tools from controlling complex networks (economics and logistics) can complement approaches currently used, which are more based on dynamical and control systems, using other mathematical frameworks for describing the systems.

# 6 Applications

The discussion on applications focused on highlighting challenges that are common to different application domains and setups (bacterial, mammalian, or cell-free systems), but used specific applications as running examples to keep arguments anchored to practically relevant problems. Motivating applications with cross-cutting challenges include, but are not limited to, cell-based and gene-based therapy, tissue engineering and regenerative medicine, diagnostic, environmental sensing and agriculture, the engineered microbiome (sentinels and sense-and-respond systems in the gut), engineered living materials, and biomanufacturing [1–4].

**Challenges discussed.** Most of these applications present unique challenges for design and control often driven by safety and performance requirements that should be robust despite environmental uncertainty. Achieving these goals often require consensus and/or engineered diversity in the cell population, which necessitates new control algorithms. Furthermore, cellular applications often require sophisticated multi-cellular sense-and-respond mechanisms (for controlling population size and growth and to avoid aberrant differentiation in the case of mammalian cells). In what follows, we expand on specific requirements motivated by the concrete application of these technologies. *Safe operation* in variable and potentially unpredictable environments must be ensured for practical application of these new technologies. This requires confidence in the design, which should be associated with safety guarantees and fail-safe operation modes. For example, if cells were to start doubling too quickly while in an internal state abiding the specification, then a safety mechanism could be triggered to detect the failure mode and kill cells that are proliferating while in the incorrect state. In the specific case of environmental applications, cells could commit suicide once the task is completed for a full cleanup. A plausible approach for guaranteed safety would be to limit the environment configuration to a set of a few situations that are reasonably well characterized up to some uncertainty, which could be taken care of by a robust control mechanism. A way to sense when environment conditions deviate from standard ones could allow engineered organism to commit suicide to avoid operating in a poorly characterized scenario. An interesting challenge with this respect is to device control solutions that adapt to changing environments through some form of adaptive control. Furthermore, generating population diversity with desired features, consensus and spatio-temporal coordination among diverse populations, can overall improve performance and robustness. Here, local sensing could be used to reach global consensus and spatio-temporal averaging.

**Tools that may be used from systems and control theory.** At a high level, control approaches for evolving systems should be considered, which conceptually fall in the class of adaptive and learning control systems but with notable differences on the level and form of uncertainty and on the extent to which the "plant" to be controlled and the environment are varying in time. Multimodal sensing, where information from different sensors and parts of a system are merged, could leverage methods developed for sensor fusion and could be used to reach consensus among different cell populations. At a high level, methods to determine system structure that allows consensus from decentralized control could be used, but again with substantial differences in the level of complexity of the dynamics of single agents (cells) or of circuit components. Also, methods for generating population diversity to allow exploration of different solutions could be used, similarly to, for example methods based on stochastic gradient descent that allow some level of exploration to reach an optimal performance.

# 7 Community Building

This topic was discussed with respect to two main aspects: data sharing and education.

**Data sharing.** Among the data and material that could be shared, we have DNA editors, DNA repositories (*addgene*), software tools to predict I/O response of parts and modules (RBS calculator [42]), design processes, experimental protocols, biophysical and simulation models along with the parameter values for clearly defined cellular and genetic backgrounds. In particular, *addgene* (`http://www.addgene.org/`) constitutes a successful example of a curated DNA repository, which could be followed to create and maintain repositories of other materials, software, and data. Data sharing protocols must allow researchers to encode data in a way that is easy to re-use. For example, starting with data (as opposed to DNA parts), such as the input/output transfer curve of a promoter-RBS-reporter system, a data sharing system would have a set of customized Q&A that allow to input all the relevant information on the conditions under which the experiments were carries, both intra-cellular and extra-cellular. Indeed, the data that people want to access must be detailed enough to allow understanding of the context of the experiment (the system's genome and other DNA background, i.e., what other plasmids are present), which is critical to re-usability. Further, a database must be easily *searchable*, by organism, by part, and by function, and all data must be *comparable*, i.e., we need an internal reference for the database such as the one we currently use for promoter strengths (reference promoter). Data representation should be human-understandable and complete, easily usable by an experimentalist. We further need an electronic methods section that should be easily filled with an interactive Q&A process. Ultimately, the way to digitally encode this information should be easy and fast. *Curators* are needed for maintaining the repository and cleaning it up, just like it is done by addgene for genetic constructs and by the ENCODE project.

One critical topic that was discussed pertains the *incentive* of researchers to contribute to the database. One potential incentive is that researchers typically want their results and data to be re-used by others and a database could provide the means to do so. Further internal incentive could be that in order for a researcher to use the information on the database they also need to contribute. A later method could be an external incentive such as a journal requiring that the data is submitted to the repository before a paper is published. For example, some journals require submission of DNA constructs to addgene before publication. An alternate or complementary approach could provide an external positive incentive, such as a journal providing a "blue ribbon" to the abstract of a paper upon publication if data was submitted to the database. This could exert some peer pressure on the authors to submit to the database. So, how do we start? We need to encourage everyone to digitize and share their data (broadly defined) to maximize their re-use in future applications. Thus, there should be an easy and fast way to make data public. For example, an interactive web-based software could be created to assist with the encoding, upload, and re-sharing of data with data formats that are adaptive to experimentalist needs. We need to lower friction/frustration and create incentives, preferably using positive incentive first, i.e., we want our own information to be re-used and also be able to access others' information and maybe obtain a blue ribbon upon publication. These could also be ways to encourage researchers to publish detailed datasets in Supplementary Data in a pre-set tabular format. Then, we could build on initial successes to keep the momentum going and possibly reach the stage where detailed tabular data formats and upload become a standard requirement for final publication.

**Education.** A typical question that we, as researchers working at the crossroads of engineering

biology and mathematical modeling/control design, are often asked by (perspective) students is "What courses do I need to take to get your expertise and skills and ultimately do the research that you do?" Today, it is still challenging to address this question and the answer varies greatly, depending on the institution. For example, some institutions have established curricula and courses that students can take both at the undergraduate and graduate levels, but other institutions have more fragmented teaching in this area. We therefore discussed what general means could be to educate students and young researchers in this area. Clearly, establishing institution-based courses and curricula is an effective way to do so, which requires resources in terms of time of faculty, teaching assistants, and teaching materials for laboratory courses. On-line courses, tutorials, and edX are also possibilities, however resources like edX need curation. On-line accredited courses, in which enrolled students obtain a degree at the end could also be an effective mechanism. One specific model to implement this could be to have students at one institution remotely take a course for credit offered at another institution, where such curricula are available. This would broaden the reach of the education curriculum created at one institution to other places where such curricula are not yet established. Summer boot camps, especially for experimentalists, could be an effctive way to start someone's education, wherein undergraduate students could use an REU program just for synthetic biology and graduate students could receive travel grants to attend the same. There is an issue, however, of knowledge retention after these short fast-paced courses. So, there should be some detailed plan for right after the camp for applying the acquired knowledge within a research project at the home institution. Overall, these structured pedagogical approaches should improve students abilities not just to perform experiments in the lab, but also and especially to create and run models, and interpret their outputs, to design systems with given specifications in mind, to analyze and interpret experimental data, and to store data in effective formats.

# 8 Conclusions and Outlook

A main focus of the workshop has been to revisit the tightness of the analogies between the design approach in synthetic biology and the modular and hierarchical approach used in electrical and computer engineering. These analogies have permeated the field of synthetic biology and a one-one mapping has often been used as a starting assumption for design. Although loose, conceptual, analogies can serve as means to reason about design problems at the initial stages in engineering biology, forcing a one-one mapping at the implementation level is, at this point, considered not beneficial for advancing the field and for overcoming current challenges in scalability, robustness, and predictability. Indeed, assumptions being made when engineering biological systems within a strict "logic gate" hierarchical design framework are clearly violated by the physics of biological systems. Examples include failure of strict modularity; genetic modules are often not digital and not even static, indeed temporal responses are substantial and neglecting them can lead to failures; and, finally, engineered genetic circuits are not robust to changes in surrounding intra-cellular and extra-cellular environment. Robustness to the environment (intra and extra-cellular) is a critical missing property of today's synthetic biology circuits that, if unresolved, will likely prevent the application of this technology in the real-world, where safety, accuracy, reliability, and performance requirements are needed while operating in uncertain and evolving environments. There has been almost no attention to problems of robustness of engineered genetic circuits so far as the focus has mostly been on creating proof-of-concept circuits that solve a specific task in specific

conditions.

In order to address the described challenges that the field is facing on issues discussed in this report, such as scalability, modularity, robustness, and predictability, it is time to deeply re-think these problems. We may want to investigate how tools from disciplines that have handled these problems in other domains, such as systems and control engineering, could be adapted or even re-invented and merged with other tools in order to be applicable in a new physical domain. Fundamental and longer term research would be needed that allows to investigate novel formulations and implementations of modularity, novel scalable approaches to design that embrace context-dependence instead of rejecting it, new mathematical formalisms for predicting systems' composition, and approaches to achieve emergent robustness from possibly non-robust composing units. These efforts should be accompanied by a community building effort that ensures the establishment of training and pedagogical resources at the intersection of synthetic biology and quantitative disciplines such as mathematical modeling, systems and control design, and dynamical systems, in a more systematic and broadly available fashion. Accordingly, a substantial effort should be placed to create shared and well curated databases for experimental protocols, systems characterization data, design processes, and mathematical and computational models. There are successful examples of these shared resources, and approaches could use these as inspiring examples to follow.

To support the above efforts, funding resources would be needed that give the sufficient temporal flexibility to explore fundamentally new approaches that can potentially be game changers in engineering biology. Within such opportunities it would be critical to keep theoretical and fundamental research grounded on concrete problems by having specific applications both as a motivation and as a test-bed used throughout the research program. It would be important that resources are allocated to broadening the reach of new or exiting pedagogical curricula beyond the home institution through a number of possible mechanisms such as, for example, discussed in this report. Tightly linked with the research project, resources could be allocated to contribute to a central database system, for which curators could be hired from the project itself to ensure the data is periodically cleaned up, internally compatible, and respects some pre-fixed metrics established by the overall research program.

# 9 Acknowledgements

We would like to thank the generous support of the National Science Foundation and Dr. Mitra Basu, who provided the funding for this event under NSF Award # 1941841.

# References

[1] D. Kang, M.and Choe, K. Kim, B-K. Cho, and S. Cho. Synthetic biology approaches in the development of engineered therapeutic microbes. *International Journal of Molecular Sciences*, 21(22):8744, 2020.

[2] W. C. Ruder, T. Lu, and J. J. Collins. Synthetic biology moving into the clinic. *Science*, 333(6047):1248, 09 2011.

[3] A. S. Khalil and J. J. Collins. Synthetic biology: applications come of age. *Nature reviews. Genetics*, 11(5):367–379, 05 2010.

[4] D. E. Cameron, C. J. Bashor, and J. J. Collins. A brief history of synthetic biology. *Nature Reviews Microbiology*, 12:381–390, 04 2014.

[5] P. E. M. Purnick and R. Weiss. The second wave of synthetic biology: from modules to systems. *Nature Reviews Molecular Cell Biology*, 10:410 EP –, 06 2009.

[6] A. A. K. Nielsen, B. S. Der, J. Shin, P. Vaidyanathan, V. Paralanov, E. A. Strychalski, D. Ross, D. Densmore, and C. A. Voigt. Genetic circuit design automation. *Science*, 352(6281):aac7341, 04 2016.

[7] P. S. Donahue, J. W. Draut, J. J. Muldoon, H. I. Edelstein, N. Bagheri, and J. N. Leonard. The COMET toolkit for composing customizable genetic programs in mammalian cells. *Nature Communications*, 11(1):1–19, 2020.

[8] R. Zhang, J. Li, J. Melendez-Alvarez, X. Chen, P. Sochor, H. Goetz, Q. Zhang, T. Ding, X. Wang, and X.-J. Tian. Topology-dependent interference of synthetic gene circuit function by growth feedback. *Nature Chemical Biology*, 16:695–701, 2020.

[9] R. Zhang, H. Goetz, J. Melendez-Alvarez, J. Li, T. Ding, X. Wang, and X.-J. Tian. Winner-takes-all resource competition redirects cascading cell fate transitions. *Nature communications*, 12(1):1–9, 2021.

[10] T. Frei, F. Cella, F. Tedeschi, J. Gutiérrez, G. Stan, M. Khammash, and V. Siciliano. Characterization and mitigation of gene expression burden in mammalian cells. *Nature communications*, 11:1–14, 2020.

[11] S. Jayanthi, K. S. Nilgiriwala, and D. Del Vecchio. Retroactivity controls the temporal dynamics of gene transcription. *ACS Synthetic Biology*, 2(8):431–441, 08 2013.

[12] Y. Qian, H-H. Huang, José. Jiménez, and D. Del Vecchio. Resource competition shapes the response of genetic circuits. *ACS Synthetic Biology*, 6(7):1263–1272, 07 2017.

[13] R. D. Jones, Y. Qian, V. Siciliano, B. DiAndreth, J. Huh, R. Weiss, and D. Del Vecchio. An endoribonuclease-based feedforward controller for decoupling resource-limited genetic modules in mammalian cells. *Nature Communications*, 11(1):5690, 2020.

[14] F. Ceroni, R. Algar, G. Stan, and T. Ellis. Quantifying cellular capacity identifies gene expression designs with reduced burden. *Nat Methods*, 12:415–418, 2015.

[15] A. Gyorgy, J. Jiménez, J. Yazbek, H-H. Huang, H. Chung, R. Weiss, and D. Del Vecchio. Isocost lines describe the cellular economy of genetic circuits. *Biophysical Journal*, 109(3):639–646, 2015.

[16] E-M. Nikolados, A. Y. Weiße, F. Ceroni, and D. A. Oyarzún. Growth defects and loss-of-function in synthetic gene circuits. *ACS Synthetic Biology*, 8(6):1231–1240, 06 2019.

[17] T. Grunberg and D. Del Vecchio. Modular analysis and design of biological circuits. *Current Opinion in Biotechnology*, 63:41–47, 2019.

[18] T. La Fleur, A. Hossain, and H. M. Salis. Automated model-predictive design of synthetic promoters to control transcriptional profiles in bacteria. *bioRxiv*, doi: https://doi.org/10.1101/2021.09.01.458561, 2021.

[19] E. Yeung, A. J. Dy, K. B. Martin, A. H. Ng, D. Del Vecchio, J. L. Beck, J. J. Collins, and R. M. Murray. Biophysical constraints arising from compositional context in synthetic gene networks. *Cell Systems*, 5(1):11–24.e12, 2017.

[20] H-H. Huang, Y. Qian, and D. Del Vecchio. A quasi-integral controller for adaptation of genetic modules to variable ribosome demand. *Nature Communications*, 9(1):5415, 2018.

[21] T. Shopera, L. He, T. Oyetunde, Y. J. Tang, and T. S. Moon. Decoupling resource-coupled gene expression in living cells. *ACS Synthetic Biology*, 6(8):1596–1604, 08 2017.

[22] S. K. Aoki, G. Lillacci, A. Gupta, A. Baumschlager, D. Schweingruber, and M. Khammash. A universal biomolecular integral feedback controller for robust perfect adaptation. *Nature*, 570(7762):533–537, 2019.

[23] A. H. Ng, T. H. Nguyen, M. Gómez-Schiavon, G. Dods, R. A. Langan, S. E. Boyken, J. A. Samson, L. M. Waldburger, J. E. Dueber, D. Baker, and H. El-Samad. Modular and tunable biological feedback control using a de novo protein switch. *Nature*, (572):265–269, 2019.

[24] U. Alon. Network motifs: theory and experimental approaches. *Nature*, 8:450–461, June 2007.

[25] L.H. Hartwell, J.J. Hopfield, S. Leibler, and A.W. Murray. From molecular to modular cell biology. *Nature*, 402:47–52, 1999.

[26] D. A. Lauffenburger. Cell signaling pathways as control modules: complexity for simplicity? *Proc. Natl. Acad. Sci.*, 97(10):5031–5033, May 2000.

[27] M. W. Kirschner and J. C. Gerhart. *The Plausibility of Life: Resolving Darwin's Dilemma*. Yale University Press, 2005.

[28] D. Mishra, P. M. Rivera-Ortiz, A. Lin, D. Del Vecchio, and R. Weiss. A load driver device to engineer modularity in biological circuits. *Nat. Biotechnology*, 32:1268–1275, 2014.

[29] J. Barton and E.D. Sontag. The energy costs of insulators in biochemical networks. *Biophysical Journal*, 104:1390–1380, 2013.

[30] B. W. Andrews, E.D. Sontag, and P. A. Iglesias. An approximate internal model principle: Applications to nonlinear models of biological systems. *Proc. of 17th IFAC Conference*, 2008.

[31] D. Del Vecchio, A. J. Ninfa, and E. D. Sontag. Modular cell biology: retroactivity and insulation. *Molecular Systems Biology*, 4(1):161, 01 2008.

[32] J. C. Willems and J. W. Polderman. *Introduction to mathematical systems theory: a behavioral approach*, volume 26. Springer Science & Business Media, 2013.

[33] J. W. Ellefson, A. J. Meyer, R. A. Hughes, J. R. Cannon, J. S. Brodbelt, and A. D. Ellington. Directed evolution of genetic parts and circuits by compartmentalized partnered replication. *Nat Biotechnol*, 32:97–101, 2013.

[34] K. J. Åström and R. M. Murray. *Feedback Systems: An Introduction for Scientists and Engineers*. Princeton University Press, 2008.

[35] K. Zhou, J. Doyle, and K. Glover. *Robust and Optimal Control*. Prentice Hall, 1996.

[36] MJ Liao, MO Din, L Tsimring, and J Hasty. Rock-paper-scissors: engineered population dynamics increase genetic stability. *BMC bioinformatics*, (365):1045–1049, 2019.

[37] G. Perrino, S. Napolitano, F. Galdi, A. La Regina, Davide Fiore, T. Giuliano, M. di Bernardo, and D. di Bernardo. Automatic synchronisation of the cell cycle in budding yeast through closed-loop feedback control. *Nature Communications*, 12:1–12, 2021.

[38] M. Krstic, I. Kanellakopoulos, and P. V. Kokotovic. *Nonlinear and Adaptive Control Design*. John Wiley & Sons, 1995.

[39] G. Em Karniadakis, I. G. Kevrekidis, L. L., P. Perdikaris, S. Wang, and L. Yang. Physics-informed machine learning. *Nature Reviews Physics*, (3):422–440, 2021.

[40] P. A. Leggieri, Y. Liu, M. Hayes, B. Connors, S. Seppälä, M. A. O'Malley, and O. S. Venturelli. Integrating systems and synthetic biology to understand and engineer microbiomes. *Annual Review of Biomedical Engineering*, (23), 2021.

[41] F. L. Lewis, H. Zhang, K. Hengster-Movric, and A. Das. *Cooperative Control of Multi-Agent Systems: Optimal and Adaptive Design Approaches*. Springer, 2014.

[42] H. M. Salis and C. Voigt. *Chapter two - The Ribosome Binding Site Calculator*, volume 498, pages 19–42. Academic Press, 2011.