Future Systems and Control Research in Synthetic Biology

Domitilla Del Vecchio[1*], Yili Qian[1], Richard M. Murray[2], and Eduardo D. Sontag[3]

[1] Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

[2] California Institute of Technology, Pasadena, CA 91125, USA

[3] Department of Electrical and Computer Engineering and Department of Bioengineering, Northeastern University, Boston, MA 02115, USA

Abstract

Synthetic biology is the application of engineering principles to the fundamental components of biology, with the aim of creating systems with novel functionalities that can be used for energy, environment, and medical applications. While the potential impact of this new technology is enormous, there are challenges that we need to overcome before the impact of synthetic biology can be fully realized. A large number of these challenges fall beyond the scope of molecular biology and are indeed "system-level" problems, where very little research is being performed. This paper identifies pressing challenges in synthetic biology that can be formulated as systems and control theoretic problems and outlines potentially new systems and control theory/tools that are required to tackle such problems. The aim is to attract more systems and control theorists to synthetic biology to help it reach its promises. At the same time, engaging the systems and control community more broadly into the rich research opportunities and life-changing applications of synthetic biology may provide added visibility to the field of systems and controls.

1 Introduction

Synthetic biology is an emergent interdisciplinary field of research, whose aim is to engineer biological systems to achieve useful functionalities. Synthetic biology provides powerful tools to address many pressing societal needs. For example, in the past decade, researchers in synthetic biology have created engineered bacteria that can produce biofuel [1] and sense heavy metals [2]; genetic circuits that can reprogram cell identity to become beta cells to treat diabetes have been designed [3]; and engineered immune cells that can track and kill cancer cells have been proposed [4]. While these efforts, among many others, demonstrate the great impact that synthetic biology can have on society, they also currently remain mostly at the laboratory stage. These efforts, in fact, rely on lengthy and *ad-hoc* design processes that do not yet give predictable outcomes in less controlled environmental conditions. Overall, poor robustness, lack of reliability, and the current inability to predict the emergent behavior of many interacting genetic components are hampering progress in the field.

The origins of these problems can to some extent be traced back to molecular biology issues relating to the way the DNA is assembled, such as relating to the reliability and orthogonality of genetic parts, and intense research efforts are underway in this direction (see [5], for example). To a large extent, however, issues of robustness, reliability, and predictability, are due to the complex dynamic interactions among system components and can be classified as "system-level" problems that fall well beyond the scope of molecular biology. Comparatively, in these problems, very little research is being performed [6–8]. The aim of this report is to provide a perspective on future systems and control research that can help solve a wide range of these system-level problems in synthetic biology, with the hope to attract more systems and control engineers to the many interesting open questions in synthetic biology that may have life-changing applications.

After a brief introduction to synthetic biology, we identify a few pressing system-level challenges that are hampering the development of synthetic biology in Section 3. In particular, the problems of

compositionality, stochasticity and spatial heterogeneity largely limit the scalability and complexity of synthetic biological systems that we can build today. In Section 4, we highlight future research opportunities that can potentially benefit the characterization, design, verification, implementation and re-design of synthetic biological systems, which can help this nascent field move forward. Some problems may involve adopting existing systems and control theoretic tools to entirely new contexts, while many others require creating novel theories and mathematical frameworks that are complementary to existing ones.

This report is largely based on the outcome of an AFOSR-funded workshop titled "The Compositionality Problem in Synthetic Biology: New Directions for Control Theory" held on June 26-27, 2017 at MIT. The workshop was co-organized by D. Del Vecchio, R. M. Murray and E. D. Sontag and was attended by the participants listed in the acknowledgements at the end of this report.

2 A Glimpse into Synthetic Biology

The ability of all living organisms to sense, communicate, and make decisions relies on a handful of highly conserved core biological processes such as gene regulation and protein-protein interactions. These, among many others, are used as functional building blocks in the *de-novo* creation of genetic circuits (Figure 1).

Brief History

The root of synthetic biology can be traced back to the Nobel winning discovery of the lac operon's regulation in bacteria *E. coli* by Jacob and Monod in the early 1960s [9]. The fact that a protein (called a transcription factor) can bind the promoter region of the gene of another protein to regulate (i.e., either activate or repress) its rate of synthesis allows us to view the gene expression process as a dynamical system with input and output (Figure 1-A), with an hope that these input/output (I/O) systems can be composed together to build more sophisticated functionalities. The advancement of biotechnology since the late 1960s has enabled time and cost-efficient technological tools to extract, sequence, amplify, and insert foreign DNA elements into cells [10]. In the year 2000, the first two synthetic genetic circuits were constructed: an oscillator [11] and a toggle switch [12]. Although these circuits were built with the aim to understand natural systems, they clearly demonstrated our technological capabilities to create *de novo* functional dynamics through model-based design of gene regulation. In the early 2000s, a number of small-scale synthetic genetic circuits, or functional modules, were constructed (Figure 1-B), including various forms of logic gates, cell-cell communication modules, cascades, feedback loops, and feedforward motifs (see [6–8] for more details). The successful assembly of biological parts into functional modules triggered the first wave of applications of synthetic biology, a few noticeable examples include environmental biosensors, *ex vivo* cell type classifiers [13], and biofuel production pathways [1] (see more examples in [14, 15]).

In the past decade, research efforts can be roughly categorized as moving along two orthogonal directions. In one direction, efforts concentrated on discovering, creating, and characterizing new biological parts and tools (e.g., CRISPR-based regulators [16]). In the other direction, efforts focused on increasing the complexity of circuits by establishing general approaches to combine available parts and modules into larger systems [17] (Figure 1-C). This is motivated by the need for sophisticated circuit functionalities in most emerging applications of synthetic biology such as those in the health industry. For example, in cancer immunotherapy, T cells need to be engineered to sense, track, and attack cancer cells while avoiding side effects to normal cells [4]; when using cell-fate reprogramming to produce insulin-secreting beta cells, the level and timing of transcription factor production must be tightly controlled [3]; in regenerative medicine, synthetic gene circuits need to accomplish multicellular coordination to form spatial patterns [18]; and for smart drug delivery, cells producing the therapeutic protein need to coordinate the timing of their lysis to release drugs periodically [19].

While these applications are exciting, their success today often relies on a lengthy trial-and-error process. In fact, when parts, modules, and systems are combined together, they work unpre-

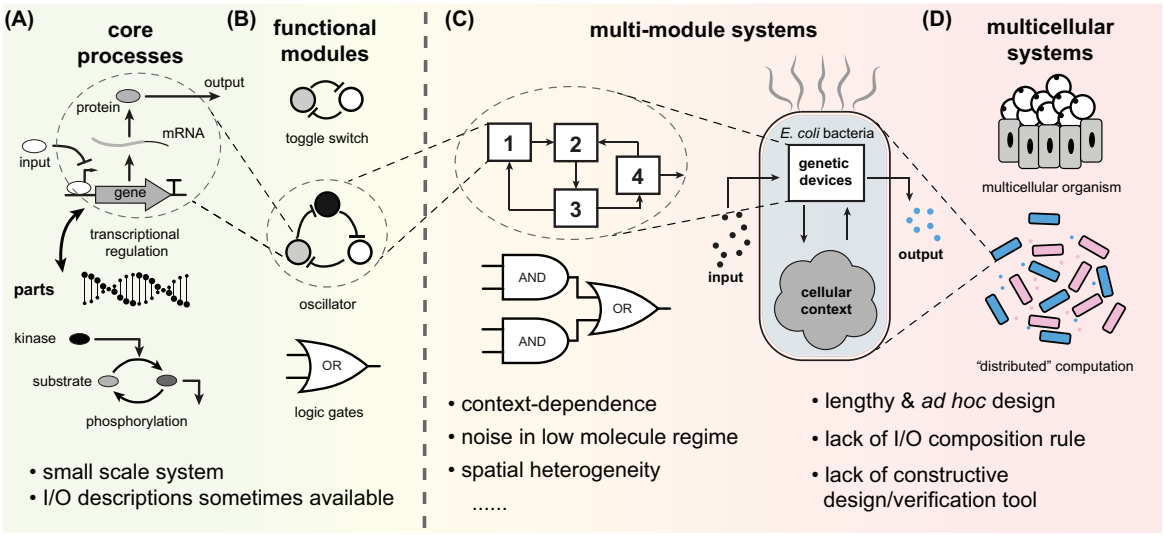

Figure 1: **Synthetic biology as the bottom-up/layered design of biological systems**. Core biological processes are encoded in DNA (panel A). These processes can be engineered into functional modules (panel B). Largely due to context-dependence in biological systems, composing functional modules together to build complex systems in a single cell (panel C) and/or in multicellular systems (panel D) is still a challenging task.

dictably [20]. Today, a designer is forced to re-characterize and re-tune the same circuit over and over again as new components are added to the system, as the host cell changes, as the growing media is slightly varied, and as temperature changes, to name a few. We analyze in more detail in the next several sections some of the roots of these problems, taking a systems and control engineering angle.

The Role of Systems and Control in Synthetic Biology

Systems-level concepts and analogies have permeated the field of synthetic biology since its interception. The abstraction hierarchy envisioned for the bottom-up design of a synthetic biomolecular system is constituted of different layers (Figure 1). This hierarchy starts with the lower abstraction layer constituted by "parts", which include DNA sequences corresponding to, for example, promoters, ribosome binding sites, and terminators. At the next layer up, we have "modules", which are pictured as I/O dynamical systems resulting from core biological processes, such as gene expression regulation, RNA-level regulation or protein modification, and (under certain assumptions) mathematical models for these processes are readily available [21, 22] (Figure 1-A,B). Next, we have "systems", which are obtained as the I/O interconnection of modules assembled in the cell, which is the circuit chassis [17, 23] (Figure 1-C). These systems could be implemented across multiple cells to lead to multicellular communities, tissues, and organs (Figure 1-D).

Context-dependence, the fact that behavior of a system changes with its environment (e.g., the chassis or the systems around it), affects all layers of the ideal abstraction hierarchy in synthetic biology. The behavior of a core process, such as gene regulation, is affected by direct connectivity to other processes, by resource competition with other system components, by the specific DNA layout around the parts that encode the process itself, by interactions with the cellular chassis (i.e., cell growth), by significant stochasticity in low molecular count regime, by spatial gradients of molecules and resources within the cell, and by spatial differences on the signaling molecule concentration in a cell population [6, 7, 20, 24].

For electronic systems, the ability to cope with uncertainty and noise, to maintain modularity

and enforce compositionality are largely based on systems and control conceptions. In fact, feedback control has been critical to *maintaining modularity* of systems, providing simplified abstractions of lower level layer functionalities for higher level layers. This is possible by conveniently describing any system through a composable I/O relationship, hence allowing to essentially "forget" about the specifics of a system's internal physical structure [25]. An example of how feedback design has enabled to hide details of dynamics and uncertainty is that of Black's feedback amplifier (1920s) [26]. The open loop amplifier device was plagued by distortion and fragility to temperature variations, making it unusable, while the feedback amplifier had an extremely robust and reliable I/O relationship independent of its context. More broadly, *managing uncertainty*, such as deriving from context, is a crucial ability in engineered systems. For example, feedback allows high performance in the presence of uncertainty by comparing actual and desired output values through accurate sensing (i.e., repeatable performance of amplifiers with 5x component variation). The use of similar design principles to robustify the behavior of genetic circuits and make it more modular has been considered in synthetic biology [27], but a unified design framework that makes circuits robust to all major sources of uncertainty from context is still largely missing.

3 Challenges and Open Questions in Synthetic Biology

Here, we identify and detail three key challenges that are impeding our current ability to perform robust and predictable design of synthetic biological systems.

3.1 Lack of Modularity and Composability

While modularity and composability are taken for granted and have enabled layered/hierarchical design in many engineering domains, in biological networks, these are still largely open questions. The reality is that genetic parts, functional modules, and systems are often influenced by their context [6, 20, 24]. As the complexity of genetic circuits has increased in recent years, various forms of context-dependence have been unveiled in experiments. Examples include, but are not limited to, loading effects resulting from direct connectivity [28–30], for which a mathematical analysis framework was developed [31]; competition for shared resources, which creates subtle coupling among otherwise unconnected modules [32, 33]; loading effects on the chassis (cell), which in turn impacts the functionality of modules and/or leads to mutations [34, 35]; and dependence of a circuit's function on the way DNA parts are assembled together [36].

Although a number of biophysical models have been established, they are often restricted to one or two forms of context-dependence described above, with very limited composability. As a consequence, no design-oriented model can effectively predict all possible interactions listed above. In addition, there are a number of interactions that we still cannot predict with sufficient accuracy. For instance, we cannot effectlively predict how protein-DNA interactions alter DNA structure, how a given protein-DNA structure impacts gene expression, how supercoiling results from genetic circuit layout, how DNA-encoded modules will perform due to possible emergence of mechanical interactions, and how location of a gene on the genome impacts gene expression [20]. Even if one could simulate the whole engineered cell through available software tools, the information obtained from these simulations would be hardly usable for design and verification. In fact, a plethora of experimental data and computational tools are available to characterize some of these effects, independently, but a design-oriented mathematical framework to describe these effects and their interactions is missing.

The ever-growing list of context-dependence effects that need to be considered during design leads us to the fundamental question of how to determine a suitable *mathematical framework to describe composition among parts, functional modules, and systems.* Currently, there is no consensus in the community as to what type of model is sufficiently descriptive to capture important biological phenomena, yet amenable to a constructive design and verification approach.

Our lack of understanding of *modularity and compositionality in natural biological systems* has

largely limited our ability to answer the above question. We, as engineers, are inclined to take a reductionist approach with a hope to fit the description of biological systems within the same convenient standards that we use for human-made systems. However, compositionality in biology may differ from the one we observe in engineered systems. Natural biological systems are in fact extremely *robust* to parameters fluctuations/uncertainty and external perturbations, such as (unfortunately) cancer pathways [37], which are very hard to eradicate. It is therefore important to learn from these natural systems to achieve similar robustness in engineered systems. For example, the fact that in biology, the behavior of a component may depend on the system around it may be exploited as a design "degree of freedom".

Performing *reverse engineering* of models from experimental data could provide a way to characterize engineered systems *in vivo* within their context [38, 39]. However, one significant problem is the nonunicity of solutions to the inverse problem, which may occur even with very large data sets and in biology we typically have sparse data [8]. In general, this will make it impossible to determine the "right" model. *Feedback control* has been instrumental in enabling layered design, allowing one to "forget" the specifics of a system's internal structure and to only focus on its I/O behavior [25]. But this requires the ability to sense, compute, and actuate precisely and accurately appropriate signals. In biology, on the one hand, we do not have sensors for many biological processes/signals and the available sensors are not sensitive enough (i.e., difficult to detect low numbers of molecules, may not be able to measure quickly enough). On the other hand, biomolecular sensors, controllers, and actuators are themselves often corrupted by disturbances, noise, and uncertainties.

Encapsulation and compartmentalization offer a way to enforce compositionality through protecting systems from interference. Biology uses this at various scales, including the cell itself, which could be used to enforce modular construction, so that the correct emergent behavior arises at the level of the cell population [40–43]. Yet, this concept is still underutilized in synthetic systems. This is because we don't know how to leverage the single agents dynamics and heterogeneity to obtain stable and robust emergent population phenotypes, especially in a situation in which we may have trillions of single agents (e.g., bacteria) as opposed to just a few as in current engineered systems.

3.2 Emergent Behaviors from Stochasticity

Biological systems are inherently stochastic due to the way in which the composing chemical reactions take place [22, 44, 45]. For a reaction to occur, molecules need to collide as a result of thermal noise, leading to a marked probabilistic behavior especially in low-molecule-count conditions [46, 47]. Stochastic effects also manifest themselves in cell populations, where gene expression is subject to substantial variability across genetically identical cells [48, 49]. Noise propagation can deteriorate circuit performance or even lead to complete circuit failure [50–53].

To theoretically study gene expression in the presence of intrinsic noise, biomolecular reactions are often treated as discrete state continuous time Markovian processes and modeled by the chemical master equations (CME) rather than ODEs [22, 54, 55]. Analytical solutions to CMEs are limited to a small set of systems often consisting of only one or two species, and finding robust and scalable approximations for larger systems is still an active area of research [56–58]. Existing approximation and simulation tools (i.e., stochastic simulation algorithms (SSA)) are often plagued by significant computational challenges and by our inability to map experimentally measured quantities to model parameters. Hence, there appear to be a general *lack of constructive tools* that can be used for verification and design. To make matters worse, another recurrent challenge is unknown biology and, in this case, the fact that it is often unknown *a priori* in which domain the system operates (i.e., high versus low molecular counts) [59], making it difficult to pick the appropriate modeling framework (i.e., ODE versus CME) to initiate the design process.

Although the modeling of monolithic stochastic biological systems can be done through the CME and simulated through stochastic simulation algorithms, most of the research (both experimental and theoretical) to date focuses primarily on noise characteristics of single gene expression, and only a very limited number of investigations on noise characteristics have been carried out on the system level. This is partly due to the fact that (de)composition of a stochastic network into the

interconnection of I/O stochastic modules is still largely unexplored. Specifically, the definition of I/O stochastic properties for modular design is generally lacking. Previous concepts of noise-to-state stability (reminiscent of the concept of input-to-state stability for nonlinear systems [60]) could be leveraged, in which the covariance of the noise appears as the "input" to the system [61, 62]. A theory for cascaded such systems has been initiated years ago, but more general interconnections have not been studied before.

In small scale circuits, feedback control has been demonstrated to be instrumental to attenuate noise [63–66]. However, the controllers, being implemented by chemical reactions, are often corrupted by noise. Specifically, the *reliability of the controller performance* in the low molecule count regime remains a significant challenge [67]. This is largely different from engineered systems, in which the "feedback" path in any controller is typically close to noise-free and highly precise/reliable.

While noise is typically regarded as a negative in engineered systems, biological systems often exploit noise and cell-cell heterogeneity in order to achieve a robust emergent phenotype [68, 69]. Examples include cell fate decisions, such as the lysis-lysogenic switch in phage lambda [70], or the process of cellular differentiation [52], which is extremely robust and reliable despite remarkable gene expression differences across cells. By contrast, noise is still underutilized in synthetic biology.

3.3 Interactions between Spatially Distributed Dynamics

Within a single cell, core processes tend to spatially localize in specific regions. For example, in bacterial cells, the DNA plasmids where genetic circuits are assembled tend to localize at the poles of the cell, the chromosome localizes at the center of the cell, and the ribosomes localize in ribosome-rich regions [71]. Similarly, in eukaryotic cells, the process of transcription occurs within the nucleus, while the process of translation takes place outside the nucleus in the cytosol [72]. In the first case, different expression of the same gene may occur depending on whether it is on the plasmid or on the chromosome, an example of how context-dependence may arise due to spatial dynamics. In the second case, the resulting "compartmentalization" may lead to a simpler sequential transcription/translation process, leading to, for example, less severe consequences of resource competition than what happens in bacteria.

Spatial heterogeneity also underlies many biological phenomena that are spatially distributed across different cells [73]. Examples includes cross-feeding among different bacterial species, which can lead to more stable and resilient communities in the presence of resource limitations [74], and morphogen gradients in early organism development that help differentiate stem cells into different cell types [75]. Therefore, whether we are designing circuits within a cell or are concerned about the emergent behavior of cell populations, spatial dynamics are important. However, depending on the specific design goal, small length-scale (e.g., intracellular) spatial information may not be as relevant/critical as longer length-scales (e.g., colonies, biofilms, tissues, and organs).

In general, current tools that consider spatial dynamics particularly for low copy number molecules are still crude. Discretization is a common way to analyze spatial dynamics. Often times, one discretizes neighboring interactions using voxel models or graph models [76]. Unfortunately, this can often lead to model artifacts that are not confirmed experimentally. Capturing spatial gradients is important and discretized models may miss these effects. The current computational burden of simulating models that include spatial dynamics is enormous; moreover, simulation tools are mostly unsuitable for design and verification. More generally, the theory to analyze reaction-diffusion PDEs is largely lacking. With this respect, a modeling framework that allows the analysis of *interconnected systems of PDEs* would be highly valuable [77, 78]. In particular, it is a challenge to connect ODE and/or CME models with PDE models in a meaningful way. In fact, agent-based simulations are not constructive and research on "spatially distributed" chemical master equations is still on-going [76, 79, 80]. These models would need to give rise to computable solutions and properties that can be efficiently used for verification and design, as opposed to simulation. Similarly, spatial context-dependence has hardly been explored. For example, cell signaling can lead to changes in cell morphology (e.g. movement/growth/division) which, in turn, affects the spatial domain in which the reactions take place.

Table 1: Potential systems and control research opportunities in synthetic biology

SYSTEM CHARACTERIZATION	SYSTEM DESIGN/CONTROL
new system ID techniques accustomed to biological models, inputs, and data acquisition identification for stochastic/PDE models **novel system descriptions** I/O description of stochastic/PDE systems composition of stochastic/PDE systems mixed stochastic/PDE I/O systems set-based I/O systems **compositional modeling framework** account for uncertainty and context	**feedback control for robustness** find realization of existing algorithms learn alternatives from nature cope with uncertainty and noise in controllers context-aware control design **system level considerations** search for robust circuit architectures exploit redundancy and compartmentalization account for mutation and biosafety **new design paradigms** evolutionary/evolutionary+modular design bioinspired design
more accurate, responsive, and standardized measurement exploit time scale separation population-level characterization, design, and control	

Engineering sufficient and efficient channels for transferring/controlling information flow from one cell to another is also still a largely open field of research. In fact, a particular grand challenge application is the design of the circuitry within each individual cell so that the community reaches a community-wide goal, an idea similar to cooperative control. Since communication among different cells relies on each individual cell producing diffusible small molecules, the challenge is to maintain a stable population of each cell type in order to coordinate reliable signal communication [81]. In particular, how to especially obtain robustness of the community emergent behavior, despite each agent (bacterium) being highly susceptible to perturbations (e.g. resource fluctuations) is an open question. Furthermore, cell-to-cell contact leads to mechanical interconnections, which necessitate hybrid (mechanical/biochemical) models. *Hybrid mechanical-biochemical models* that are computationally tractable and can educate design are still largely lacking.

4 Systems and Control Research Opportunities

In this section we highlight some potential research opportunities for the systems and control field, which stem from the challenges so far described. We hope that addressing these research questions will both advance synthetic biology and demonstrate the impact of systems and control tools in a yet largely unexplored application domain. Some highlights of these research opportunities are summarized in Table 1.

4.1 System Identification

The lack of appropriate system identification techniques is a key obstacle to solve many issues in synthetic biology, from context-dependence, to stochastic effects, to spatial effects. In order to carry out model-based design, there is a pressing need to establish accustomed system identification procedures for deterministic, stochastic and spatially distributed biomolecular models. A major hurdle is the fact that current technology only allows us to measure a few (often noisy) outputs (e.g., fluorescence) of a nonlinear dynamical process, where linearization is not an option, with limited time resolution. As a consequence, identifiability becomes a major issue: it is common to have multiple sets of parameters that all fit the experimental data equally well (see [8] for a technical review).

We therefore need to develop new system identification techniques that can (i) educate optimal placement of a small number of sensors, compatible with physical and technological constraints, (ii) provide rational selection criteria for candidate biological models, and (iii) exploit stochasticity in the model and the measurements, working synergistically with existing stochastic computational tools. For example, one opportunity is that stochasticity may provide persistency of excitation and may lead to further restrict the set of possible models. In addition, it is important to put context-dependence into picture to develop experimental practices that minimize contextual effects from measurements (e.g., retroactivity). Finally, we currently have a limited capability to identify the dynamics of spatially distributed biological systems.

4.2 Compositional Modeling Frameworks

A *compositional modeling framework* is critical to untangle complexity associated with design and verification problems. Design/verification of a system in a monolithic fashion is a combinatorial problem; in particular, for system design, lack of compositionality leads to a need for re-tuning the behavior of each component once a new module is added to the system. To establish a compositional modeling framework, it is important to correctly capture the boundary of a module, its I/O interfaces, and its context. With respect to the context, one could think of restricting the dynamics of a module to a different set of possibilities depending on the biological context so to account for the unknowns that are associated with it.

Differential inclusion models may be a possible framework to capture all the uncertainty that we have about such systems [82]. However, it is often the case that we do not even know how many states we have. Developing constructive observability and controllability tools for such models would be helpful to determine what can be designed/verified with the given level of uncertainty that we have to cope with. Another open question is whether differential inclusion models are useful for producing constructive tools for design and verification. In addition, time-scale separation is key in all these processes and could be leveraged for analysis and design and also to simplify the models or to help define the layers of abstraction [83].

One could consider hierarchical networks of dynamical systems, in which the links between nodes obey "contracts", which can be framed in logical (AND/OR) or modal ("before") properties of inputs and outputs, and use ideas from contract-based design [84]. The contracts can be composed according to the interconnection scheme of the components or reverse-engineered through interrogation and observation. For example, within a Markov decision process framework, it is possible to learn the cost functions of individual agents that collectively reach a steady state. Specifically, the cost functions used by metabolic systems can be inferred from data about metabolite fluxes [85]. It is less clear, however, how such approaches could be used for other kinds of cellular processes in which information, rather than matter, is being processed.

We may need to create new mathematical systems descriptions along with the compositional rules for I/O systems, accounting for how nature has composed systems through the course of evolution, perhaps reasoning about I/O sets. The context needs to be accounted for as a mechanism to reduce uncertainty when more information from experimentalists is available. This framework should be constructive, that is, it should enable explicit design and verification procedures. The notion of *weak regulatory linkage* [86], which enables variation and selection in the course of evolution, may provide a new way of describing composition in engineered biological systems. These linkages reconfigure the interfaces between conserved core processes as opposed to enforcing agreed interfaces between varied components. However, no existing mathematical framework exists that captures this notion rigorously, especially because establishing such framework would require cross-fertilization between different communities, i.e., biologists who provide plenty of domain knowledge about the context and engineers who can create such design-oriented frameworks.

Additional traditional approaches that could be considered include computer science approaches to analyze multi-agent and distributed systems [87], the concept of reconfiguration (allow a system to re-wire itself under stress) [88], and stochastic safety verification tools [89], but these tools may not be constructive without leveraging domain-specific structure. Boolean networks capture the

ON/OFF of gene expression and provide a coarse model that is appropriate when a lot of the knowledge is missing [90, 91].

4.3 Stochastic and Spatially Distributed Systems

While there remains a strong need to improve analytical understanding and computational efficiency of standard stochastic models (in particular, CME and SSA), these standard tools can be expanded by a number of existing theoretical/computational tools developed in other engineering contexts, which may decrease model complexity and increase computational efficiency in certain situations. These include stochastic hybrid systems [92], interval-based methods (i.e., interval arithmetic) to capture uncertainty, especially in the parameters [93]; viability theory, leading to set-based approaches to verification and design [82]; and queuing theory [94]. However, an underlying challenge in these existing approaches is that explicit computation, which is often critical for design, is not possible. Hence, there appear to be a general *lack of constructive tools* that can be used for verification and design.

Mathematical frameworks for *I/O composition of stochastic system descriptions* is an interesting and challenging research avenue, which may require the development of new representations of systems and new rules of composition. New theory will need to be developed for producing constructive/compositional approaches to verification and design. Existing tools such as noise-to-state stability [61, 62] may be adopted towards this goal. However, this would first require a deeper understanding of the stochastic properties of biological systems, especially in the low molecular count regime. This understanding will also allow us to determine fundamental limitations that are likely to exist in the design of closed loop architectures, thus uncovering potential tradeoffs between stochasticity and sharp control.

A similar *I/O compositional framework of PDE models* along with constructive design and verification tools would be helpful to account for many of the spatial effects that today we cannot yet handle in design and verification [78, 95]. To this end, there is a need to establish input-to-state stability notions in PDE models in order to analyze robustness to disturbances (mechanical, thermal, chemical, biological), uncertain parameters, and unmodeled dynamics. Furthermore, a constructive and compositional mathematical framework that provides *mixed CME/PDE descriptions* is highly desirable to capture both effects efficiently and simultaneously.

4.4 Feedback Controller Design for Robustness and Modularity

Context-dependence and general lack of robustness of engineered biological circuits remain a major motivation for feedback control design in synthetic biology (see [6–8] for a more technical review). Robustness to perturbations (i.e., noise, parameter uncertainty, and environmental changes) is a remarkable property of virtually all living systems, yet current synthetic biology circuits are highly fragile to perturbations, poorly reliable, and often result in unpredictable behavior. Feedback control has been instrumental in achieving robustness and reliability since the onset of electrical circuit design.

However, feedback design in biomolecular systems faces a number of difficulties unique to biology, most noticeably: (1) there is often no clear and established realization of existing control algorithms through available biomolecular reactions and (2) like the process that needs to be controlled, the feedback path is often equally plagued by disturbance, uncertainty and noise, and there is a lack of fundamental understanding on how these deteriorate control performance. Therefore, it is unclear whether mapping our classical closed loop feedback diagram [25] directly to biological circuits is the most appropriate approach to take to handle robustness problems.

While it is definitely tempting to mimic classical engineering designs, the field of synthetic biology may be held back by pushing too far the analogy with existing engineered systems. Although recent studies have developed a small set of biomolecular tools to mimic essential mathematics operations used in classical feedback control, such as signal subtraction and integration [67, 96–98], it is possible that natural systems perform feedback in a more efficient way that is different from what we are used

to think. In addition to the lack of basic modules required for signal measurement and computation, control design in synthetic biology is further complicated by much larger parametric uncertainty (more than 10x) and noise than typically experienced in traditional engineering applications.

Strategies that relax the requirement for a precise feedback path in a closed loop design would be highly valuable since every path is subject to large parametric uncertainty and noise. More generally, a mathematical framework to achieve a robust emergent behavior by appropriately connecting highly uncertain components would be highly valuable. Components are subject to 10-100x variation and elaborate circuit design strategies to obtain a sharp emergent behavior are needed. Addressing these robustness questions may also require to move beyond traditional core processes in synthetic biology, such as gene expression and gene regulation, and move to other types of processes, such as protein-protein interactions and CRISPR/Cas-based systems. These biological tools are currently underutilized, yet they may allow faster responses and an easier tuning procedure. However, how to perform circuit design using these core precesses to achieve a desired functionality remains a largely unexplored research avenue.

Finally, biomolecular controllers, just like the process they are aimed to control, are context-dependent. Therefore, controller design must be carried out with the additional constraint that it should not, for example, impose heavy burden on the host cell and/or incur unexpected interactions when connected with the plant. On the one hand, this requires control engineers to be very familiar with the (context-dependent) characteristics of the available biological tools. For example, two genetic circuits realizing the same ODE model may result in completely different resource demands. On the other hand, context-aware control design principles still needs to be explored.

4.5 System-Level Considerations for Robustness

A different approach to increase circuits' robustness to uncertainty is to search for circuit architectures that are better suited to handle poor characterization and thus result in a robust emergent system behavior despite highly uncertain components [99]. On the one hand, systems and control theory can help identify subsystem properties and interconnection rules required to maintain robust system-level behaviors [100]. On the other hand, a bio-inspired approach, where we learn from biology how notions of modularity are used, how composition and wiring is performed, and how robustness is achieved may be promising.

Exploiting *redundancy and crosstalk* in the design of circuits may be an additional approach to reach robustness and resiliency. These are aspects of natural systems, in which multiple paths exists to transmit the same signal and signaling pathways often share components [101], which lead to significant cross-talk (lack of modularity). This may be a mechanism for robustness that engineered systems could also exploit. In this sense, less modularity may lead to increased robustness.

Encapsulation and compartmentalization offer a way to enforce compositionality through protecting systems from interference. In this respect, cells themselves could be used to enforce modular and robust construction, so that the correct emergent behavior arises at the level of the cell population without the need to worry about context-dependence at the single cell level. While this idea of "distributed computation" has been realized by synthetically producing small molecules that can diffuse through cell membranes [40–42], theoretical studies are largely missing. This results in limited system-level guidance available for implementation and therefore in solutions that are not generalizable beyond the specific experimental systems and conditions considered. Cell-cell variability, spatial heterogeneity, communication (diffusion) delay, the large number of agents (billions to trillions), and most importantly, the need to create stable cell populations are major hurdles where control engineers could substantially contribute.

The problem of *mutations* is unique to engineering biology. Mutations are the result of selective pressure against the mutated components and could be theoretically captured by formulating a cost function that the cell is trying to optimize. Investigation of design approaches to make an engineered cell population robust to mutations could be very valuable. Specifically, some genetic circuit architectures are more robust to mutations than others because they evaluate to a better cost. How to determine such architectures is an interesting and challenging research question that

will most of all require a deeply intertwined theoretical (e.g., from optimal control and learning systems) and experimental approach.

Finally, *biosafety* is especially critical to move lab designs into real-world applications [102]. There is a need to learn from safety-critical system design and verification practices in traditional engineering systems and adopt them to the biological context.

4.6 New Circuit Design Paradigms

Model-based design may not be the only approach for biological systems. In fact, we can perform combinatorial search through a large set of circuit architectures with high-throughput experiments where the appropriate selective pressure is applied to cells. *Directed evolution* [103, 104] is an example of how this can be performed and is very effective for design space exploration. Engineering methods that synergistically combine *modular/layered design approaches with evolutionary design techniques* could be particularly promising and may lead to design/verification methodologies that are constructive and also in-line with how nature designs its systems. This may require establishing a new research direction, for example in optimal control, in which the cost function is implemented through a selective pressure applied to the engineered cells so that they reconfigure themselves, leading to the optimal desired design.

With this respect, an interesting research approach could be to merge *model-based design* techniques, which require a reasonable characterization of the physical process, with *machine learning-based design* techniques, which are mostly data-driven. Specifically, system identification and sensing techniques may be developed to better characterize biological systems used in synthetic biology, and machine learning techniques along with domain-specific knowledge may be used to reduce the uncertainty on the context.

4.7 Population-Level Design

Population-based computation, in which cells compute and coordinate among each other to obtain an emergent ecosystem is highly desirable for many applications of synthetic biology, from programmable probiotics to regenerative medicine [73]. However, while co-existence of cell communities is ubiquitous in natural systems, engineering such an eco-system where multiple cell populations stably co-exist in an equilibrium that is robust to extraneous species is a formidable challenge [105–109]. Attractors exist for these systems that are far more complicated than simple equilibria. The whole population may be "stable" according to some relaxed notion of stability, but each of the composing cells may not be and may, instead, dynamically changing its state under the laws of physics or simply due to noise. This could be a mechanism for resilience such that the whole system equilibrium is robust to external interference but its composing agents are highly susceptible to perturbation and continuously evolving. This begs the question of what mathematics may be appropriate to describe and analyze such systems. These communities have trillions of cells (i.e., in our guts we have about 100 trillion bacteria), therefore a multi-agent approach to the problem will likely be inapplicable. Perhaps a PDE-based approach may be more promising, but current PDE tools are most likely insufficient to describe and analyze the important properties of these systems. New formalisms and analysis/design tools are most likely needed to reason about these questions.

The theoretical foundations to perform "cooperative control" of trillion agents to obtain *resilient behavior at the population level* is an intriguing research direction. Each agent may implement a different component of a circuit so that the circuit becomes distributed across a number of different cellular communities [40–42], allowing to defeat several sources of context-dependence at the single cell level (e.g. resource limitation). This requires new compositional/descriptive frameworks that allow for efficient verification and design, despite the large size of the systems considered and a significant amount of communication delay due to slow molecular diffusion.

4.8 Exploiting Time Scale Separation

Time scale separation is a ubiquitous property of biological processes. Different core processes occur on very different time scales, ranging from subseconds for protein-DNA or protein-protein interactions, to minutes for enzymatic reactions, to hours for gene expression, to days for cell fate decisions, to weeks for tissue formation [21]. The dynamics at the different layers of the design abstraction hierarchy of synthetic biology also occur at well separated time scales, with dynamics of subseconds to seconds for DNA conformation changes to dynamics of days for cell population dynamics. Stochastic processes are also widely distributed on the temporal axis, especially depending on the regime (low versus high molecular counts), with molecular interactions occurring in the subsecond time scale and average noise affecting gene expression in the high-molecular counts in minutes to hours time scale. Finally, spatial dynamics are characterized by time scale separation as well, with diffusion of mRNA molecules within the cell occurring much faster than that of larger molecules (e.g., ribosomes). Spatial dynamics also occur on different length-scales, from sub-micron length for diffusion of molecules within bacteria, to tens of microns for diffusion of molecules in mammalian cells, to hundreds of microns for diffusion of molecules among cells in a population.

On the one hand, system identification, analysis and simulation are challenged by such a wide scale at which the phenomena of interest occur. How to remove dynamics that are tangential to the system of interest to simplify analysis and design remains an open question, especially for stochastic and spatially distributed systems [110, 111]. On the other hand, time scale separation may enable simplifying models and also reaching some level of "insulation" between processes that occur at different time scales. This property has been exploited to obtain modular synthetic systems (e.g., mitigate retroactivity [27, 112]) and to facilitate control design (e.g., approximate integral control [98]). More research in this direction should be explored to support synthetic biology design.

4.9 Instrumentation and Standardization

A fundamental issue impeding our ability to carry out high fidelity system identification and control design is the lack of *accurate and reliable sensors* that could enable better characterization of processes, parts, and interactions (i.e., stochastic, spatial, structural/mechanical). Unfortunately, we do not have a sufficient number of accurate sensors to monitor conformational DNA interactions, single molecules within a cell, cell growth, and stochastic fluctuations that may occur at faster time scales. Creation of more sensitive and responsive biosensors with minimal context-dependence (e.g., retroactivity and resource demand), as well as lab measurement techniques for single cell dynamics will be beneficial in this respect. Furthermore, there is still no agreement in the community on the standard units that should be used to measure biological signals, and therefore quantitative measurements are often lab-dependent and host/environment-dependent [113], preventing better cross-fertilization among different labs. Similarly, a high fidelity mapping from experimental measurement to parameters in stochastic/spatial models is still largely absent.

4.10 Simulation and Cell-Free Test-Beds

For most of the biological phenomena discussed, mechanistic simulation tools exist. For example, these include whole cell simulation models for bacterial cells, which incorporate cellular context [114]; the stochastic simulation algorithm to capture noise effects at all molecular counts [115]; agent-based simulation tools that can account for spatial and cell-cell dynamics [116]; and molecular simulation algorithms that account for forces and mechanical interactions involving DNAs [117]. However, *existing mechanistic simulation tools are often not suitable for design and verification* since they simulate the system as a monolithic entity and do not allow composition of simpler systems to create/verify the final system of interest. Therefore, a design/verification problem remains a combinatorial problem. Furthermore, simulation depends on the specific choice of parameters, thus providing little insight for design and verification, where closed-form analytical expressions, even if approximate, are much more useful.

Cell-free systems could be a promising middle ground to test design ideas and even a potential medium to implement circuits in case these do not need a living cell [118]. Cell-free systems provide a means for "running" a circuit of interest without being susceptible to any unknown interaction with the host cell [119, 120]. However, there are still problems of standardization of cell extracts for meaningful quantitative analysis and comparison with *in vivo* circuits [119–122], in addition to the issue of resource depletion, which substantially shortens the life span of the circuit of interest.

5 Summary and Outlook

In this paper, we articulated a number of potential research opportunities for systems and control motivated from pressing problems in synthetic biology. These opportunities entail new theory that can have unprecedented impact by enabling ground-breaking applications of synthetic biology to health, energy, and environment. While the field of synthetic biology started with *model-based design* [11, 12], the field has progressively moved away from theory-based design. Modeling and analysis should be a precursor as opposed to an after-thought to the experiments, yet this is rarely the case nowadays as theory is lagging behind the quick progress in molecular biology and genetic engineering.

The contribution from the systems and control community to synthetic biology is conditioned on reaching out to experimentalists and opening a conversation that will provide control theorists the appropriate domain-specific knowledge to tackle relevant engineering problems. For example, the ability of designing dynamics has largely relied on systems and control theory, where a plant of interest that is intrinsically unstable (i.e., highly agile aircrafts) or underperforming can be made stable and robust to perturbations by suitable "closed loop" design. This ability is highly needed in synthetic biology, but the control theorist will have to learn how to go from design of a closed loop system on paper to concrete biological parts and to their reliable interconnection. A strong synergy is therefore required between control theorists and synthetic biologists for these research opportunities to unfold.

Acknowledgements

This report summarizes the discussions of a workshop titled "The Compositionality Problem in Synthetic Biology: New Directions for Control Theory" held on June 26-27, 2017 at MIT. The workshop was co-organized by D. Del Vecchio, R. M. Murray and E. D. Sontag and funded by the US Air Force Office of Scientific Research under grant number FA9550-17-1-0289.

Participants of the workshop were: Neda Bagheri (Northwestern University), Chase Beisel (North Carolina State University), Leonidas Bleris (University of Texas Dallas), Cullen Buie (MIT), James J Collins (MIT), Mary Dunlop (Boston University), Elisa Franco (University of California Riverside), Jeremy Gunawardena (Harvard Medical School), Joao Hespanha (University of California Santa Barbara), Pablo Iglesias (Johns Hopkins University), Ahmad Khalil (Boston University), Miroslav Krstic (University of California San Diego), Reinhard Laubenbacher (University of Connecticut Health Center), Frederick Leve (AFOSR, Dynamics and Control Program), Yannis Paschalidis (Boston University), Johan Paulsson (Harvard Medical School), Abhyudai Singh (University of Delaware), David Spivak (MIT), Mario Sznaier (Northeastern University), and Jeff Tabor (Rice University).

References

[1] P. P. Peralta-Yahya, F. Zhang, S. B. del Cardayre, and J. D. Keasling. Microbial engineering for the production of advanced biofuels. *Nature*, 488:320–328, 2012.

[2] J. R. van der Meer and S. Belkin. Where microbiology meets microengineering: design and applications of reporter bacteria. *Nat. Rev. Microb.*, 8:511–522, 2010.

[3] P. Saxena, B.-C. Heng, P. Bai, M. Folcher, H. Zulewski, and M. Fussenegger. A programmable synthetic lineage-control network that differentiates human IPSCs into glucose-sensitive insulin-secreting beta-like cells. *Nat. Commun.*, 7:11247, 2016.

[4] D. Chakravarti and W. W. Wong. Synthetic biology in cell-based cancer immunotherapy. *Trends Biotechnol.*, 33(8):449–461, 2015.

[5] S. Kosuri, D. B. Goodman, G. Cambray, V. K. Mutalik, Y. Gao, A. P. Arkin, D. Endy, and G. M. Church. Composability of regulatory sequences controlling transcription and translation inEscherichia coli. *Proc. Natl. Acad. Sci. U.S.A.*, 110(34):14024–14029, 2013.

[6] D. Del Vecchio, A. J. Dy, and Y. Qian. Control theory meets synthetic biology. *J. R. Soc. Interface*, 13(120), 2016.

[7] Y. Qian, C. McBride, and D Del Vecchio. Programming cells to work for us. *Annual Review of Control, Robotics, and Autonomous Systems*, 1(1), 2017.

[8] V. Hsiao, A. Swaminathan, and R. M. Murray. Control theory for synthetic biology. *IEEE Control Syst. Mag.*, 2017. in press.

[9] F. Jacob and J. Monod. Genetic regulatory mechanisms in the synthesis of proteins. *J. Mol. Biol.*, 3(3):318–356, 1961.

[10] D. E. Cameron, C. J. Bashor, and J. J. Collins. A brief history of synthetic biology. *Nat. Rev. Microbiol.*, 12(5):381–390, 2014.

[11] M. B. Elowitz and S. Leibler. A synthetic oscillatory network of transcriptional regulators. *Nature*, 403(6767):335–338, 2000.

[12] T. S. Gardner, C. R. Cantor, and J. J. Collins. Construction of a genetic toggle switch in Escherichia coli. *Nature*, 403(6767):339–342, 2000.

[13] Z. Xie, L. Wroblewska, L. Prochazka, R. Weiss, and Y. Benenson. Multi-input RNAi-based logic circuit. *Science*, 333:1307–1312, 2011.

[14] W. C. Ruder, T. Lu, and J. J. Collins. Synthetic biology moving into the clinic. *Science*, 333:1248–1252, 2011.

[15] A. S. Khalil and J. J. Collins. Synthetic biology: applications come of age. *Nat. Rev. Genet.*, 11:367–379, 2010.

[16] B. Jusiak, S. Cleto, P. Perez-Piñera, and T. K. Lu. Engineering synthetic gene circuits in living cells with CRISPR technology. *Trends Biotechnol.*, 34(7):535–547, 2016.

[17] P. E. M. Purnick and R. Weiss. The second wave of synthetic biology: from modules to systems. *Nat. Rev. Mol. Cell Biol.*, 10:410–422, 2009.

[18] F. Berthiaume, T. J. Maguire, and M. L. Yarmush. Tissue engineering and regenerative medicine: History, progress, and challenges. *Annu. Rev. Chem. Biomol. Eng.*, 2(1):403–430, 2011.

[19] M. O. Din, T. Danino, A. Prindle, M. Skalak, J. Selimkhanov, K. Allen, E. Julio, E. Atolia, L. S. Tsimring, S. N Bhatia, and J. Hasty. Synchronized cycles of bacterial lysis for *in vivo* delivery. *Nature*, 536:81–85, 2016.

[20] S. Cardinale and A. P. Arkin. Contextualizing context for synthetic biology- identifying causes of failure of synthetic biological systems. *Biotechnol. J.*, 7:856–866, 2012.

[21] U. Alon. *An Introduction to Systems Biology: Design Principles of Biological Circuits*. Chapman & Hall/CRC Press, 2006.

[22] D. Del Vecchio and R. M. Murray. *Biomolecular Feedback Systems*. Princeton University Press, Princeton, 2014.

[23] D. Baker, G. Church, J. Collins, D. Endy, J. Jacobson, J. Keasling, P. Modrich, C. Smolke, and R. Weiss. Engineering Life: Building a FAB for Biology. *Sci. Amer.*, pages 44–51, June 2006.

[24] D. Del Vecchio. Modularity, context-dependence, and insulation in engineered biological circuits. *Trends Biotechnol.*, 33(2):111–119, 2015.

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

[26] R. Kline. Harold Black and the negative-feedback amplifier. *IEEE Control Syst. Mag.*, 13(4):82–85, 1993.

[27] D. Mishra, P. M. Rivera, A. Lin, D. Del Vecchio, and R. Weiss. A load driver device for engineering modularity in biological networks. *Nat. Biotechnol.*, 32:1268–1275, 2014.

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

[29] A. C. Ventura, P. Jiang, L. Van Wassenhove, D. Del Vecchio, S. D. Merajver, and A. J. Ninfa. Signaling properties of a covalent modification cycle are altered by a downstream target. *Proc. Natl. Acad. Sci. U.S.A.*, 107(22):10032–10037, 2010.

[30] P. Jiang, A. C. Ventura, E. D. Sontag, S. D. Merajver, A. J. Ninfa, and D. Del Vecchio. Load-induced modulation of signal transduction networks. *Sci. Signal.*, 4(194):ra67–ra67, 2011.

[31] D. Del Vecchio, A. J. Ninfa, and E. D. Sontag. Modular cell biology: retroactivity and insulation. *Mol. Syst. Biol.*, 4(1), 2008.

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

[33] Y. Qian, H.-H Huang, J. I. Jiménez, and D. Del Vecchio. Resource competition shapes the response of genetic circuits. *ACS Synth. Biol.*, 6(7):1263–1272, 2017.

[34] F. Ceroni, R. Algar, G.-B. Stan, and T. Ellis. Quantifying cellular capacity identifies gene expression designs with reduced burden. *Nat. Methods*, 12(5):415–422, 2015.

[35] C. Tan, P. Marguet, and L. You. Emergent bistability by a growth-modulating positive feedback circuit. *Nat. Chem. Biol.*, 5(11):842–848, 2009.

[36] 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.

[37] D. Hanahan and R. A. Weinberg. Hallmarks of cancer: The next generation. *Cell*, 144(5):646–674, 2011.

[38] Q. Zhao, A. I. Stettner, E. Reznik, I. Ch. Paschalidis, and D. Segrè. Mapping the landscape of metabolic goals of a cell. *Genome Biol.*, 17(1), 2016.

[39] T. Kang, R. Moore, Y. Li, E. Sontag, and L. Bleris. Discriminating direct and indirect connectivities in biological networks. *Proc. Natl. Acad. Sci. U. S. A.*, 112(41):12893–12898, 2015.

[40] A. Tamsir, J. J. Tabor, and C. A. Voigt. Robust multicellular computing using genetically encoded nor gates and chemical 'wires'. *Nature*, 469(7329):212–5, 2011.

[41] S. Regot, J. Macia, N. Conde, K. Furukawa, J. Kjellén, T. Peeters, S. Hohmann, E. de Nadal, F. Posas, and R. Solé. Distributed biological computation with multicellular engineered networks. *Nature*, 469(7329):207–211, 2011.

[42] V. Hsiao, Y. Hori, P. W. Rothemund, and R. M Murray. A population-based temporal logic gate for timing and recording chemical events. *Mol. Syst. Biol.*, 12(5):869, 2016.

[43] Q. Zhao, D. Segre, and I. Ch. Paschalidis. Optimal allocation of metabolic functions among organisms in a microbial ecosystem. In *Proc. 55th IEEE Conf. Decis. Control*, pages 7063–7068, Las Vegas, NV, 2016.

[44] M. Khammash. Modeling and analysis of stochastic biochemical networks. In *Control Theory and Systems Biology*, pages 29–44. The MIT Press, 2009.

[45] J. Paulsson. Summing up the noise in gene networks. *Nature*, 427(6973):415–418, 2004.

[46] M. B. Elowitz. Stochastic gene expression in a single cell. *Science*, 297(5584):1183–1186, 2002.

[47] N. Rosenfeld, J. W. Young, U. Alon, P. S. Swain, and M. B. Elowitz. Gene regulation at the single-cell level. *Science*, 307(5717):1962–1965, 2005.

[48] T. M. Norman, N. D. Lord, J. Paulsson, and R. Losick. Memory and modularity in cell-fate decision making. *Nature*, 503(7477):481–486, 2013.

[49] A. Raj and A. van Oudenaarden. Nature, nurture, or chance: Stochastic gene expression and its consequences. *Cell*, 135(2):216–226, 2008.

[50] S. Hooshangi, S. Thiberge, and R. Weiss. Ultrasensitivity and noise propagation in a synthetic transcriptional cascade. *Proc. Natl. Acad. Sci. U. S. A.*, 102(10):3584–3586, 2005.

[51] M. R. Atkinson, M. A. Savageau, J. T. Myers, and A. J. Ninfa. Development of genetic circuitry exhibiting toggle switch or oscillatory behavior in Escherichia coli. *Cell*, 113(5):597–607, 2003.

[52] M. Wu, R.-Q. Su, X. Li, T. Ellisc, Y.-C. Lai, and X. Wang. Engineering of regulated stochastic cell fate determination. *Proc. Natl. Acad. Sci. U. S. A.*, 110(26):10610–10615, 2013.

[53] V. Siciliano, I. Garzilli, C. Fracassi, S. Criscuolo, S. Ventre, and D. di Bernardo. miRNAs confer phenotypic robustness to gene networks by suppressing biological noise. *Nat. Commun.*, 4, 2013.

[54] N. G. Van Kampen. *Stochastic Processes in Physics and Chemistry, Third Edition (North-Holland Personal Library)*. North Holland, 2007.

[55] T. G. Kurtz. The relationship between stochastic and deterministic models for chemical reactions. *J. Chem. Phys.*, 57(7):2976–2978, 1972.

[56] A. Singh and J. P. Hespanha. Approximate moment dynamics for chemically reacting systems. *IEEE Trans. Autom. Control.*, 56(2):414–418, 2011.

[57] A. Gupta, C. Briat, and M. Khammash. A scalable computational framework for establishing long-term behavior of stochastic reaction networks. *PLOS Comput. Biol.*, 10(6):e1003669, 2014.

[58] M. Naghnaeian and D. Del Vecchio. Robust moment closure method for the chemical master equation. In *2017 IEEE Conference on Control Technology and Applications (CCTA)*. IEEE, 2017.

[59] J. Pahle. Biochemical simulations: stochastic, approximate stochastic and hybrid approaches. *Brief. Bioinform.*, 10(1):53–64, 2008.

[60] E.D. Sontag. Smooth stabilization implies coprime factorization. *IEEE Trans. Automat. Contr.*, 34(4):435–443, 1989.

[61] H. Deng, M. Krstic, and R.J. Williams. Stabilization of stochastic nonlinear systems driven by noise of unknown covariance. *IEEE Trans. Autom. Control.*, 46(8):1237–1253, 2001.

[62] M. Krstic and H. Deng. *Stabilization of nonlinear uncertain systems (communications and control engineering)*. Springer, 1998.

[63] A. Becskei and L. Serrano. Engineering stability in gene networks by autoregulation. *Nature*, 405(6786):590–593, 2000.

[64] D. Nevozhay, R. M. Adams, K. F. Murphy, K. Josic, and G. Balazsi. Negative autoregulation linearizes the dose-response and suppresses the heterogeneity of gene expression. *Proc. Natl. Acad. Sci. U.S.A.*, 106(13):5123–5128, 2009.

[65] V. Shimoga, J. T. White, Y. Li, E. Sontag, and L. Bleris. Synthetic mammalian transgene negative autoregulation. *Mol. Syst. Biol.*, 9(1):670–670, 2014.

[66] A. Singh. Negative feedback through mRNA provides the best control of gene-expression noise. *IEEE Trans. Nanobiosci.*, 10(3):194–200, 2011.

[67] C. Briat, A. Gupta, and M. Khammash. Antithetic integral feedback ensures robust perfect adaptation in noisy biomolecular networks. *Cell Systems*, 2(1):15–26, 2016.

[68] A. Eldar and M. B. Elowitz. Functional roles for noise in genetic circuits. *Nature*, 467(7312):167–173, 2010.

[69] L. Pelkmans. Using cell-to-cell variability–a new era in molecular biology. *Science*, 336(6080):425–426, 2012.

[70] M. Ptashne. *A Genetic Switch, Third Edition: Phage Lambda Revisited*. Cold Spring Harbor Laboratory Press, 2004.

[71] M. Castellana, S. H.-J. Li, and N. S. Wingreen. Spatial organization of bacterial transcription and translation. *Proc. Natl. Acad. Sci. U.S.A.*, 113(33):9286–9291, 2016.

[72] B. Alberts. *Molecular Biology of the Cell*. Garland Science, 2014.

[73] S. Payne and L. You. Engineered cell-cell communication and its applications. In *Productive Biofilms*, pages 97–121. Springer, 2013.

[74] K. Brenner, L. You, and F. H. Arnold. Engineering microbial consortia: a new frontier in synthetic biology. *Trends Biotechnol.*, 26(9):483–489, 2008.

[75] J. B. Gurdon and P.-Y. Bourillot. Morphogen gradient interpretation. *Nature*, 413(6858):797–803, 2001.

[76] M. Klann and H. Koeppl. Spatial simulations in systems biology: From molecules to cells. *Int. J. Mol. Sci*, 13(12):7798–7827, 2012.

[77] Z. Aminzare and E. D. Sontag. Synchronization of diffusively-connected nonlinear systems: Results based on contractions with respect to general norms. *IEEE Trans. Netw. Sci. Eng.*, 1(2):91–106, 2014.

[78] Z. Aminzare and E. D. Sontag. Some remarks on spatial uniformity of solutions of reaction–diffusion PDEs. *Nonlinear Anal. Theory Methods Appl.*, 147:125–144, 2016.

[79] J. Hattne, D. Fange, and J. Elf. Stochastic reaction-diffusion simulation with MesoRD. *Bioinformatics*, 21(12):2923–2924, 2005.

[80] M. Ander, P. Beltrao, B. Di Ventura, J. Ferkinghoff-Borg, M. Foglierini, A. Kaplan, C. Lemerle, I. Tomas-Oliveira, and L. Serrano. SmartCell, a framework to simulate cellular processes that combines stochastic approximation with diffusion and localisation: analysis of simple networks. *Syst Biol (Stevenage)*, 1(1):129–138, 2004.

[81] B. Li and L. You. Synthetic biology: division of logic labour. *Nature*, 469(7329):171–172, 2011.

[82] J.-P. Aubin and A. Cellina. *Differential Inclusions*. Springer Berlin Heidelberg, 1984.

[83] H. A. Simon. The architecture of complexity. *Proc. Am. Philos. Soc.*, 106(6):467–482, 1962.

[84] P. Nuzzo, A. L. Sangiovanni-Vincentelli, D. Bresolin, L. Geretti, and T. Villa. A platform-based design methodology with contracts and related tools for the design of cyber-physical systems. *Proc. IEEE.*, 103(11):2104–2132, 2015.

[85] M. Basan, S. Hui, H. Okano, Z. Zhang, Y. Shen, J. R. Williamson, and T. Hwa. Overflow metabolism in Escherichia coli results from efficient proteome allocation. *Nature*, 528(7580):99–104, 2015.

[86] M. W. Kirschner and J. C. Gerhart. *The Plausibility of Life: Resolving Darwins Dilemma*. Yale University Press, 2006.

[87] M. Wooldridge. *An Introduction to MultiAgent Systems*. Wiley, 2009.

[88] C. Steiger, H. Walder, and M. Platzner. Operating systems for reconfigurable embedded platforms: online scheduling of real-time tasks. *IEEE Trans. Comput.*, 53(11):1393–1407, 2004.

[89] S. Prajna, A. Jadbabaie, and G. J. Pappas. A framework for worst-case and stochastic safety verification using barrier certificates. *IEEE Trans. Autom. Control.*, 52(8):1415–1428, 2007.

[90] M. Chaves, E. D. Sontag, and R. Albert. Methods of robustness analysis for boolean models of gene control networks. *IEE Proceedings - Systems Biology*, 153(4):154, 2006.

[91] B. Shao, X. Liu, D. Zhang, J. Wu, and Q. Ouyang. From boolean network model to continuous model helps in design of functional circuits. *PLOS ONE*, 10(6):e0128630, 2015.

[92] J. P. Hespanha. Modelling and analysis of stochastic hybrid systems. *IEE Proceedings - Control Theory and Applications*, 153(5):520–535, 2006.

[93] R. E. Moore. *Methods and Applications of Interval Analysis*. Society for Industrial and Applied Mathematics, 1979.

[94] N. A. Cookson, W. H. Mather, T. Danino, O. Mondragón-Palomino, R. J. Williams, L. S. Tsimring, and J. Hasty. Queueing up for enzymatic processing: correlated signaling through coupled degradation. *Mol. Syst. Biol.*, 7:561, 2011.

[95] M. Miller, M. Hafner, E. D. Sontag, N. Davidsohn, S. Subramanian, P. E. M. Purnick, D. Lauffenburger, and R. Weiss. Modular design of artificial tissue homeostasis: Robust control through synthetic cellular heterogeneity. *PLOS Comput. Biol.*, 8(7):e1002579, 2012.

[96] V. Hsiao, E. L. C. De Los Santos, W. R. Whitaker, J. E. Dueber, and R. M. Murray. Design and implementation of a biomolecular concentration tracker. *ACS Synth. Biol.*, 4(2):150–161, 2015.

[97] C. Zechner, G. Seelig, M. Rullan, and M. Khammash. Molecular circuits for dynamic noise filtering. *Proc. Natl. Acad. Sci. U.S.A.*, 113(17):4729–4734, apr 2016.

[98] Y. Qian and D. Del Vecchio. Realizing 'integral control' in living cells: how to overcome leaky integration due to dilution? *J. R. Soc. Interface.*, 15(139), 2018.

[99] W. Ma, A. Trusina, H. El-Samad, W. A. Lim, and C. Tang. Defining network topologies that can achieve biochemical adaptation. *Cell*, 138(4):760–773, 2009.

[100] G. Russo, M. di Bernardo, and E. D. Sontag. A contraction approach to the hierarchical analysis and design of networked systems. *IEEE Trans. Autom. Control.*, 58(5):1328–1331, 2013.

[101] R. Albert, B. DasGupta, R. Hegde, G. S. Sivanathan, A. Gitter, G. Gürsoy, P. Paul, and E. D. Sontag. Computationally efficient measure of topological redundancy of biological and social networks. *Phys. Rev. E.*, 84(3), 2011.

[102] O. Wright, G.-B. Stan, and T. Ellis. Building-in biosafety for synthetic biology. *Microbiology*, 159(Pt_7):1221–1235, 2013.

[103] E. L. Haseltine and F. H. Arnold. Synthetic gene circuits: Design with directed evolution. *Annu. Rev. Biophys. Biomol. Struct.*, 36(1):1–19, 2007.

[104] Y. Yokobayashi, R. Weiss, and F. H. Arnold. Directed evolution of a genetic circuit. *Proc. Natl. Acad. Sci. U.S.A.*, 99(26):16587–16591, 2002.

[105] L. You, R. S. Cox, R. Weiss, and F. H. Arnold. Programmed population control by cell-cell communication and regulated killing. *Nature*, 428(6985):868–871, 2004.

[106] F. K. Balagaddé, H. Song, J. Ozaki, C. H. Collins, M. Barnet, F. H. Arnold, S. R. Quake, and L. You. A synthetic Escherichia coli predator-prey ecosystem. *Mol. Syst. Biol.*, 4, 2008.

[107] S. R. Scott, M. O. Din, P. Bittihn, L. Xiong, L. S. Tsimring, and J. Hasty. A stabilized microbial ecosystem of self-limiting bacteria using synthetic quorum-regulated lysis. *Nat. Microbiol.*, 2(17083), 2017.

[108] R. A. Armstrong and R. McGehee. Coexistence of species competing for shared resources. *Theor. Popul. Biol.*, 9(3):317–328, 1976.

[109] K. R. Foster and T. Bell. Competition, not cooperation, dominates interactions among culturable microbial species. *Curr. Biol.*, 22(19):1845–1850, 2012.

[110] C. A. Gómez-Uribe, G. C. Verghese, and A. R. Tzafriri. Enhanced identification and exploitation of time scales for model reduction in stochastic chemical kinetics. *J. Chem. Phys.*, 129(24):244112, 2008.

[111] N. Herath and D. Del Vecchio. Model order reduction for linear noise approximation using time-scale separation. In *2016 IEEE 55th Conference on Decision and Control (CDC)*. IEEE, 2016.

[112] R. Shah and D. Del Vecchio. Signaling architectures that transmit unidirectional information despite retroactivity. *Biophys. J.*, 113(3):728–742, 2017.

[113] M. Fahlberg, K. Groglio, V. Toxavidis, B. Israel, and John Tigges. Standardization of flow cytometry instrumentation for the analysis of microparticles. *J. Biomol. Tech.*, 23(Suppl):S38–S38, 2012. 147[PII].

[114] J. R. Karr, J. C. Sanghvi, D. N. Macklin, M. V. Gutschow, J. M. Jacobs, B. Jr Bolival, N. Assad-Garcia, J. I. Glass, and M. W. Covert. A whole-cell computational model predicts phenotype from genotype. *Cell*, 150(2):389–401, 2012.

[115] D. T. Gillespie. Exact stochastic simulation of coupled chemical reactions. *J. Phys. Chem.*, 81(25):2340–2361, 1977.

[116] T. E. Gorochowski. Agent-based modelling in synthetic biology. *Essays Biochem.*, 60(4):325–336, 2016.

[117] T. E. Cheatham and M. A. Young. Molecular dynamics simulation of nucleic acids: Successes, limitations, and promise. *Biopolymers*, 56(4):232–256, 2000.

[118] K. Pardee, A. A. Green, T. Ferrante, D. Ewen Cameron, A. DaleyKeyser, P. Yin, and J. J. Collins. Paper-based synthetic gene networks. *Cell*, 159(4):940–954, 2014.

[119] M. K. Takahashi, C. A. Hayes, J. Chappell, Z. Z. Sun, R. M. Murray, V. Noireaux, and J. B. Lucks. Characterizing and prototyping genetic networks with cell-free transcription–translation reactions. *Methods*, 86:60–72, 2015.

[120] J. Shin and V. Noireaux. An e. coli cell-free expression toolbox: Application to synthetic gene circuits and artificial cells. *ACS Synth. Biol.*, 1(1):29–41, 2012.

[121] H. Niederholtmeyer, Z. Z. Sun, Y. Hori, E. Yeung, A. Verpoorte, R. M. Murray, and S. J. Maerkl. Rapid cell-free forward engineering of novel genetic ring oscillators. *eLife*, 4, 2015.

[122] V. H. Nagaraj, J. M. Greene, A. M. Sengupta, and E. D. Sontag. Translation inhibition and resource balance in the TX-TL cell-free gene expression system. *Synthetic Biology*, 2(1), 2017.