Annual Reviews in Control 45 (2018) 5–17

Contents lists available at ScienceDirect

Annual Reviews in Control

journal homepage: www.elsevier.com/locate/arcontrol

Vision article

Future systems and control research in synthetic biology

Domitilla Del Vecchio[a,*], Yili Qian[a], Richard M. Murray[b], Eduardo D. Sontag[c]

[a] Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
[b] California Institute of Technology, Pasadena, CA 91125, USA
[c] Department of Electrical and Computer Engineering and Department of Bioengineering, Northeastern University, Boston, MA 02115, USA

ARTICLE INFO

Article history:
Received 16 February 2018
Revised 20 April 2018
Accepted 20 April 2018
Available online 24 May 2018

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. Many 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 theories/tools that are required to tackle such problems. The aim is to attract more systems and control theorists to collaborate with molecular biologists and biophysicists and help synthetic biology reach its promise. 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.

Contents

* Corresponding author.
E-mail address: ddv@mit.edu (D. Del Vecchio).

https://doi.org/10.1016/j.arcontrol.2018.04.007
1367-5788/

1. Introduction

Synthetic biology is an emergent interdisciplinary field of research, whose aim is to engineer biomolecular 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 (Peralta-Yahya, Zhang, del Cardayre, & Keasling, 2012) and sense heavy metals (van der Meer & Belkin, 2010), genetic circuits that can reprogram cell identity to treat diabetes (Saxena et al., 2016), and engineered immune cells that can track and kill cancer cells (Chakravarti & Wong, 2015). 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. In fact, most synthetic genetic circuits constructed nowadays 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 this field.

The origins of these problems can to some extent be traced back to molecular biology issues, such as the reliability and orthogonality of genetic parts, and intense research efforts are underway in this direction (see Arpino et al., 2013; Kosuri et al., 2013, 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 beyond the scope of molecular biology. Comparatively, in these problems, very little research is being performed. In addition, as we discuss in more detail in Section 3, existing theoretical tools and mathematical frameworks adopted directly from engineering systems are often unsuitable and/or inefficient to deal with the level of complexity in biomolecular systems. The aim of this paper is to provide a perspective on future systems and control research that can help solve a wide range of 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.

This paper is not a comprehensive review of synthetic biology. Instead, it is a vision paper aimed at motivating future theoretical research and new mathematical frameworks that could facilitate the design, analysis, and verification of synthetic genetic circuits and is intended for readers with a background in systems and control theory. Nevertheless, we should clarify that mathematical tools are valuable to synthetic biology only if they are aware of the domain-specific constraints, such as limitations of a biophysical model and the available design parameter space. In fact, many of the problems we describe here reflect such needs. 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 paper 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 paper.

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 (Fig. 1).

2.1. Brief history

The roots 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 (Jacob & Monod, 1961). 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 an input and an output (Fig. 1-A), with a 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 (Cameron, Bashor, & Collins, 2014). In the year 2000, the first two synthetic genetic circuits were constructed: an oscillator (Elowitz & Leibler, 2000) and a toggle switch (Gardner, Cantor, & Collins, 2000). 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 (Fig. 1-B), including various forms of logic gates, cell-cell communication modules, cascades, feedback loops, and feedforward motifs (see Cameron et al., 2014; Del Vecchio, Dy, & Qian, 2016; Hsiao, Swaminathan, & Murray, 2018, in press; Qian, McBride, & Del Vecchio, 2018 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 (Xie, Wroblewska, Prochazka, Weiss, & Benenson, 2011), and biofuel production pathways (Peralta-Yahya et al., 2012) (see more examples in Ruder, Lu, and Collins (2011) and Khalil and Collins (2010)).

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 biological parts and tools (see, for example, Arpino et al., 2013). 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 (Purnick & Weiss, 2009) (Fig. 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 (Chakravarti & Wong, 2015); when using cell-fate reprogramming to produce insulin-secreting beta cells, the level

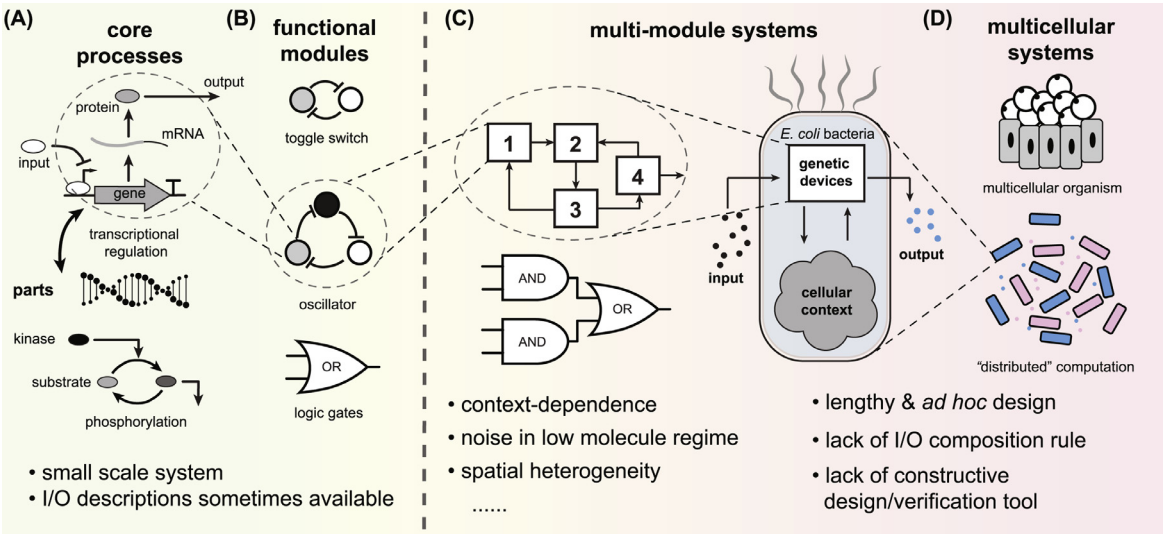

Fig. 1. Synthetic biology as the bottom-up/layered design of biological systems. Core biological processes are encoded in DNA (panel A). These processes can be engineered into functional modules (panel B). Largely due to context dependence in biological systems, composing functional modules together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

and timing of transcription factor production must be tightly controlled (Saxena et al., 2016); in regenerative medicine, synthetic gene circuits need to accomplish multicellular coordination to form spatial patterns (Berthiaume, Maguire, & Yarmush, 2011); and for smart drug delivery, cells producing the therapeutic protein need to coordinate the timing of their lysis to release drugs periodically (Din et al., 2016).

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 unpredictably (Cardinale & Arkin, 2012). Today, a designer is forced to re-characterize and re-tune the same circuit over and over again as new modules are added, as the host cell changes, as the growing media is varied, and as temperature changes, to name a few. We analyze in Section 3 some of the roots of these problems, taking a systems and control engineering angle.

2.2. The role of systems and control in synthetic biology

Systems-level concepts and analogies have permeated the field of synthetic biology since its inception. The abstraction hierarchy envisioned for the bottom-up design of a synthetic genetic circuit is constituted of different layers (Fig. 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 (Alon, 2006; Del Vecchio & Murray, 2014) (Fig. 1-A and Fig. 1-B). Next, we have "systems", which are obtained as the I/O interconnection of modules assembled in the cell, which is the circuit chassis (Baker et al., 2006; Purnick & Weiss, 2009) (Fig. 1-C). These systems could be implemented across multiple cells to lead to multicellular communities, tissues, and organs (Fig. 1-D).

Context dependence, the fact that a system's behavior 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 (Cardinale & Arkin, 2012; Del Vecchio, 2015; Del Vecchio et al., 2016; Qian et al., 2018).

For electronic systems, the ability to cope with uncertainty and noise, to maintain modularity, and to enforce compositionality are largely based on systems and control concepts. 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 compositional I/O relationship, hence allowing a designer to essentially "forget" about the specifics of a system's internal physical structure and treat it as a black box with inputs and outputs (Åström & Murray, 2008). An example of how feedback design has enabled to hide details of dynamics and uncertainty is that of Black's feedback amplifier (1920s) (Kline, 1993). 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 the uncertainty due to context dependent effects, 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 to a certain type of context dependence has been considered in synthetic biology (Mishra, Rivera, Lin, Del Vecchio, & Weiss, 2014), 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 system-level challenges that are impeding our current ability to perform robust and predictable design of synthetic biological systems.

3.1. Lack of modularity and compositionality

Many complex engineering systems can be regarded as a number of low-level components (i.e., modules) interconnected through a set of prescribed rules via suitable I/O interfaces (Tripakis, 2016). Design, analysis, and verification of such systems rely heavily on the property of system *compositionality*, which guarantees that system-level behaviors can be predicted from (1) the relevant I/O behaviors of the constituent components and (2) the interconnection rules. In contrast, a system is called *monolithic* if it contains no clear architecturally separate components whose functions can be composed to obtain the emergent system-level behavior. Compositional systems are therefore more efficient to design, analyze, and verify than monolithic ones. A key to compositionality is the ability to abstract possibly complicated component dynamics by I/O behaviors. Hence, a necessary condition for system compositionality is the *modularity* of all its constituent components, which implies that components' I/O behaviors are independent of the context, including the environment and the neighboring components.

While modularity and compositionality are taken for granted and have enabled layered/hierarchical design and verification in many engineering domains, in biological systems, these are still largely open questions. The reality is that genetic parts, functional modules, and systems are often influenced by their context (Cardinale & Arkin, 2012; Del Vecchio, 2015; Del Vecchio et al., 2016). 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 (Jayanthi, Nilgiriwala, & Del Vecchio, 2013; Jiang et al., 2011; Ventura et al., 2010), for which a mathematical analysis framework was developed (Del Vecchio, Ninfa, & Sontag, 2008); competition for shared resources, which creates subtle coupling among otherwise unconnected modules (Gyorgy et al., 2015; Qian, Huang, Jiménez, & Del Vecchio, 2017a); loading effects on the chassis (cell), which in turn impacts the functionality of modules and/or leads to mutations (Ceroni, Algar, Stan, & Ellis, 2015; Tan, Marguet, & You, 2009); and dependence of a circuit's function on the way DNA parts are assembled together (Yeung et al., 2017).

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 compositionality. 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 effectively 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 (Cardinale & Arkin, 2012). 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-dependent 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. Natural biological systems are in fact extremely robust to parameter fluctuations/uncertainty and external perturbations, such as (unfortunately) cancer pathways (Hanahan & Weinberg, 2011), which are very hard to eradicate. 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 man-made systems. While the use of modularity and compositionality is a convenient way to analyze and design robust systems, it may not be the only way and what we may view as undesirable "crosstalk" may actually be exploited by natural systems to attain robustness. For example, when faced with nutrient limitations, genetically identical cells in a biofilm can exhibit crosstalk and differentiate into different cell types depending on their locations within the colony (i.e., peripheral or interior) to achieve cooperative division of metabolic labor and consequently robustness of growth and survival (Bocci, Suzuki, Lu, & Onuchic, 2018; Liu et al., 2015; Yamagishi, Saito, & Kaneko, 2016). It is therefore important to learn from these natural systems and possibly learn how to exploit the context around a system as a design "degree of freedom".

Performing *reverse engineering* of models from experimental data provides a way to characterize engineered systems in vivo within their context. For example, by systematically perturbing a synthetic circuit in vivo to reconstruct the network topology and then compare it with the known regulatory interactions, context-dependent interactions can be isolated. Such information can then be used to develop models that account for context dependence. This strategy has been applied in Kang, Moore, Li, Sontag, and Bleris (2015) and Quarton, Kang, Sontag, and Bleris (2016) to identify hidden interactions due to resource competition, resulting in a model that matches well with a mechanistic resource competition model (Qian et al., 2017a). A universal challenge in biological inverse problems and system identification is the poor quality of available data. In fact, biological data are often either too sparse due to the lack of high quality sensors or too "dense" due to the lack of specific perturbations (e.g., correlation-based high-throughput gene expression data). In the latter case, one must be extremely cautious to differentiate causality and correlation (Feizi, Marbach, Médard, & Kellis, 2013; Kang et al., 2015).

Feedback control has been instrumental in enabling layered design, allowing one to "forget" the specifics of a system's internal structure and to treat it as a black box with a prescribed dynamic I/O relationship (Åström & Murray, 2008). However, this requires the ability to sense, compute, and actuate precisely and accurately appropriate signals. In synthetic biology, on the one hand, we do not have biomolecular sensors for many biological processes/signals and the available sensors are not sensitive enough (i.e., difficult to detect low numbers of molecules and unable to sample quickly enough). On the other hand, biomolecular sensors, controllers, and actuators are themselves often corrupted by disturbances, noise, and uncertainties.

Biology utilizes *encapsulation and compartmentalization* at various scales to enforce modularity and to increase robustness. For example, at the molecular scale, scaffold proteins regulate interaction of signaling molecules through localization with high specificity (Good, Zalatan, & Lim, 2011); at the cellular scale, eukaryotic cells pump toxic metal ions into specialized compartments (i.e., organelles) for detoxification to avoid metal-induced toxicity (Dameron & Harrison, 1998); and at the populational scale, division of metabolic labor among cells themselves enables robust growth and survival of the colony (Bocci et al., 2018; Liu et al., 2015; Yamagishi et al., 2016). At the molecular and cellular scale, synthetic biologists have created synthetic compartmentalizations to mimic the natural ones (see Chen & Silver, 2012 for a review). At

the populational scale, synthetic microbial ecosystems have been constructed (Balagaddé et al., 2008; Scott et al., 2017). Yet, regarding cells themselves as compartments to achieve modularity and division of labor is still an underutilized concept in synthetic biology. In addition to molecular technological bottlenecks such as the lack of orthogonal cell-cell communication modules (Payne & You, 2013), we still lack systematic understanding of how heterogeneous agent dynamics could give rise to robust emergent population phenotypes. The presence of large number of heterogeneous agents in biological systems challenges existing theories, which are often oriented to engineered systems that contain only a few agents and/or can be regarded as homogeneous.

3.2. Emergent behaviors from stochasticity

Biological systems are inherently stochastic due to the way in which the constituent chemical reactions take place (Del Vecchio & Murray, 2014; Khammash, 2009; Paulsson, 2004). 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 (Elowitz, 2002; Rosenfeld, Young, Alon, Swain, & Elowitz, 2005). Stochastic effects also manifest themselves in cell populations, where gene expression is subject to substantial variability across genetically identical cells (Norman, Lord, Paulsson, & Losick, 2013; Raj & van Oudenaarden, 2008). Noise propagation can deteriorate circuit performance or even lead to complete circuit failure (Atkinson, Savageau, Myers, & Ninfa, 2003; Hooshangi, Thiberge, & Weiss, 2005; Siciliano et al., 2013; Wu et al., 2013).

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 (CMEs) rather than ODEs (Del Vecchio & Murray, 2014; Van Kampen, 2007; Kurtz, 1972). 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 (Gupta, Briat, & Khammash, 2014; Naghnaeian & Del Vecchio, 2017; Singh & Hespanha, 2011). Existing approximation and simulation tools (i.e., stochastic simulation algorithms (SSAs)) are often plagued by significant computational challenges and by our inability to map experimentally measured quantities to model parameters. Hence, there appears to be a general *lack of constructive tools* that can be used for design and verification of stochastic biomolecular systems. 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) (Pahle, 2008), making it difficult to pick the appropriate modeling framework (e.g., 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 SSA, 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 (Sontag, 1989)) could be leveraged, in which the covariance of the noise appears as the "input" to the system (Deng, Krstic, & Williams, 2001; Krstic & Deng, 1998). 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 (Becskei & Serrano, 2000; Nevozhay, Adams, Murphy, Josic, & Balazsi, 2009; Shimoga, White, Li, Sontag, & Bleris, 2014; Singh, 2011). However, the controllers, being implemented by chemical reactions, are often corrupted by noise. The reliability of the controller performance in the low molecule count regime remains a major challenge. 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. However, certain biomolecular controller designs, such as the antithetic feedback in Briat, Gupta, and Khammash (2016a), can perform robustly in the presence of noise. It is therefore desirable to expand the toolbox of such controllers and/or to come up with reliable design principles in noisy environments.

While noise is typically regarded as undesirable in engineered systems, biological systems often exploit noise and cell-cell heterogeneity in order to achieve a robust emergent phenotype (Eldar & Elowitz, 2010; Pelkmans, 2012). Examples include cell fate decisions, such as the lysis-lysogenic switch in phage lambda (Ptashne, 2004), or the process of cellular differentiation (Wu et al., 2013), 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 prokaryotic cells, the chromosome and the RNA polymerase localize at the center of the cell, and the ribosomes localize in ribosome-rich regions, often near the two endcaps (Bakshi, Siryaporn, Goulian, & Weisshaar, 2012; Castellana, Li, & Wingreen, 2016). Therefore, different expression of the same gene may occur depending on (i) whether it is on the plasmid or on the chromosome and (ii) the localization of the plasmid in the cytoplasm (Pogliano, Ho, Zhong, & Helinski, 2001). As a consequence, gene expression is dependent on the spatial context. In comparison, although transcription and translation in eukaryotic cells are inherently more complex (e.g., they involve chromatin remodeling and RNA splicing), they happen in specialized compartments (i.e., inside and outside the nucleus) (Alberts, 2014). This spatial arrangement leads to easier access to transcriptional and translational resources and possibly more homogeneous gene expression with respect to spatial distribution of genes.

Spatial heterogeneity also underlies many biological phenomena that are spatially distributed across different cells (Payne & You, 2013). Examples include cross-feeding among different bacterial species, which can lead to more stable and resilient communities in the presence of resource limitations (Brenner, You, & Arnold, 2008; Liu et al., 2015; Wintermute & Silver, 2010), and morphogen gradients in early organism development that help differentiate stem cells into different cell types (Gurdon & Bourillot, 2001). 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 (Klann & Koeppl, 2012). Unfortunately, this can lead to model artifacts that cannot be confirmed experimentally. Capturing spatial gradients is important and discretized models may miss these effects. The current computational burden of simulating

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

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

models that include spatial dynamics is enormous, making them mostly unsuitable for design and verification. More generally, the theory to analyze reaction-diffusion PDEs is largely lacking, and a modeling framework that allows the analysis of *interconnected systems of PDEs* would be highly valuable (Aminzare & Sontag, 2014; 2016). 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" CMEs is still on-going (Ander et al., 2004; Hattne, Fange, & Elf, 2005; Klann & Koeppl, 2012). These models would need to give rise to computable solutions and properties that can be efficiently used for design and verification, 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.

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 (Li & You, 2011). In particular, how to 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 solving 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 customized 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 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 (Hsiao et al., 2018; Villaverde, Barreiro, & Papachristodoulou, 2016).

We therefore need to develop new system identification techniques that can (i) educate optimal placement of a small number of sensors, (ii) provide rational selection criteria for candidate biological models, and (iii) exploit stochasticity in the measurements to work synergistically with stochastic models. For example, one opportunity is that stochasticity may provide persistency of excitation and may further restrict the set of possible models. In addition, it is important to incorporate the consideration of context dependence, so as 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 and verification of a monolithic system is a combinatorial problem. For example, lack of compositionality leads to a need to retune the behavior of each module once a new component 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 (Aubin & Cellina, 1984). 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, timescale separation could be leveraged to simplify the models or to help define the layers of abstractions (Simon, 1962).

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 (Nuzzo, Sangiovanni-Vincentelli, Bresolin, Geretti, & Villa, 2015). The contracts can be composed according to the interconnection scheme of the modules 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 (Basan et al., 2015). 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* (Kirschner & Gerhart, 2006), 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 captures this notion rigorously, especially because establishing such framework would require cross-fertilization between different communities: biologists who provide domain-specific 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 (Wooldridge, 2009), the concept of reconfiguration (allow a system to re-wire itself under stress) (Steiger, Walder, & Platzner, 2004), and stochastic safety verification tools (Prajna, Jadbabaie, & Pappas, 2007), but these tools may not be constructive without leveraging domain-specific structures. Boolean networks capture the ON/OFF of gene expression and provide a coarse model that is appropriate when detailed mechanistic knowledge is missing (Chaves, Sontag, & Albert, 2006; Shao, Liu, Zhang, Wu, & Ouyang, 2015).

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 (Hespanha, 2006); interval-based methods (i.e., interval arithmetic) that capture uncertainty in the parameters (Moore, 1979); viability theory, leading to set-based approaches to design and verification (Aubin & Cellina, 1984); and queuing theory (Cookson et al., 2011). However, an underlying challenge in these existing approaches is that explicit computation, which is often critical for design, is not possible. Hence, there appears to be a general lack of constructive tools that can be used for design and verification.

Developing mathematical frameworks for *I/O composition of stochastic systems* are interesting and challenging research avenues, which may require new representations of systems and new rules of composition. New theory will need to be developed for producing constructive/compositional approaches to design and verification. Existing tools such as noise-to-state stability (Deng et al., 2001; Krstic & Deng, 1998) 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 we cannot handle today in design and verification (Aminzare & Sontag, 2016; Miller et al., 2012). 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 stochastic and spatial effects efficiently and simultaneously.

4.4. Feedback controller design for robustness and modularity

Context dependence and the general lack of robustness of synthetic genetic circuits remain major motivations for feedback control design (see Del Vecchio et al., 2016; Hsiao et al., 2018; Qian et al., 2018; Steel, Lillacci, Khammash, & Papachristodoulou, 2017 for more technical reviews). 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 and poorly reliable, often resulting in unpredictable behavior. While feedback control has been instrumental in achieving robustness and reliability since the onset of electrical circuit design, applying these existing control theoretic tools directly to biomolecular systems faces a number of difficulties.

Firstly, there are often no clear and established realizations of existing control algorithms through available biomolecular reactions. Fueled by earlier researches in adaptation and homoeostasis in natural systems, one of the most popular control mechanisms in biology is integral control. In particular, the integral control structure theoretically demonstrated in bacterial chemotaxis to achieve adaptation (Barkai & Leibler, 1997; Yi, Huang, Simon, & Doyle, 2000) motivated subsequent searches for biomolecular reaction motifs that can achieve or approximate integral control (Ang, Bagh, Ingalls, & McMillen, 2010; Briat et al., 2016a; Briat, Zechner, & Khammash, 2016b; Drengstig et al., 2012; Drengstig, Ueda, & Ruoff, 2008; Klavins, 2010). However, although several recent experiments have shown promising results (Chang, Armitage, Papachristodoulou, & Wadhams, 2013; Hsiao, De Los Santos, Whitaker, Dueber, & Murray, 2015; Lillacci, Aoki, Schweingruber, & Khammash, 2017), implementing a synthetic biomolecular controller in vivo that clearly demonstrates the adaptation property is still largely an open research question. This is because many proposed motifs realize integral control in the limit of infinitely strong binding between two molecules (Ang et al., 2010; Drengstig et al., 2008; Klavins, 2010) and engineering/finding molecular interaction pairs with sufficiently strong affinity remains a difficult task (Arpino et al., 2013). 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. In the case of integral control, dilution of biomolecular species due to cell growth makes it essentially impossible to construct a non-leaky memory variable to carry out perfect integration

of the error signal for feedback (Olsman et al., 2017; Qian & Del Vecchio, 2018). It is possible that natural systems perform feedback in a more efficient way that is different from the common practices in engineering.

Another grand challenge in biomolecular feedback system design is that, like the process that needs to be controlled, the feedback path is often equally plagued by disturbances, parametric uncertainty and noise, and there is a lack of fundamental understanding on how these factors deteriorate control performance. Therefore, it is unclear whether mapping our classical closed loop feedback diagram (Åström & Murray, 2008) directly to biological circuits is the most appropriate approach to handle robustness problems. 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. 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 processes to achieve a desired functionality remains a largely unexplored research avenue.

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 a heavy burden on the host cell and/or incur unexpected interactions when connected with the process to be regulated. 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.

Finally, feedback control typically trades off robustness and fragility: feedback systems designed to hedge against one type of disturbance are typically more fragile to other types of disturbances (Csete & Doyle, 2002; Kitano, 2007; Whitacre, 2012). A theoretical manifestation of this trade-off in classical linear control theory is Bode's sensitivity integral (Åström & Murray, 2008), where robustness to low-frequency disturbances, for example, must be achieved at the price of fragility to high-frequency ones. While it is generally believed that through evolution, natural systems have developed a set of complex mechanisms to balance this robustness-fragility trade-off in their context, it is crucial to factor this trade-off into consideration when transferring synthetic genetic circuits from simple laboratory environments to real-world applications (Chan, Lee, Cameron, Bashor, & Collins, 2015; Wright, Stan, & Ellis, 2013). In this respect, 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.

4.5. 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 (Ma, Trusina, El-Samad, Lim, & Tang, 2009). On the one hand, systems and control theory can help identify component properties and interconnection rules required to maintain robust system-level behaviors (Russo, di Bernardo, & Sontag, 2013). 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.

Redundancy is common in natural systems, in which multiple paths exist to transmit the same signal (Edelman & Gally, 2001; Whitacre, 2012). Exploiting redundancy in the design of circuits may help reach robustness and resiliency, but one has to carefully balance robustness and the size of the system. Another feature of natural signaling pathways is that they often share components, which leads to significant *crosstalk* (i.e., lack of modularity). One notable example is the bow-tie architecture in bacterial metabolic networks (Csete & Doyle, 2004), where several key species (i.e., "knots") are involved in most of the metabolic processes. The seemingly contradictory coexistence of modularity and crosstalk illustrate the complexity of natural systems. Completely modular systems are typically costly to build and to operate (e.g., a machine without a master switch), the lack of modularity in the crosstalk "knots" may serve as ways (i) to balance robustness and efficiency and (ii) to provide means for global regulation. For example, RNA polymerases can be regarded as a "knot" in gene networks as they are required for all transcriptional regulation actions. This property is taken advantage of in the bacterial stringent response, where redirecting RNA polymerases to different cellular tasks can globally change the cell growth mode when faced with nutrient starvation (Magnusson, Farewell, & Nyström, 2005). In this sense, less modularity leads to an efficient global regulation mechanism for increased robustness. Of course, a system is fragile when its key crosstalk "knots" are subject to attacks, a strategy utilized by certain viruses to invade host cells (Bushell & Sarnow, 2002). For synthetic genetic circuits, how to balance robustness (e.g., by enforcing modularity) and efficiency (e.g., by allowing cross-talk) may largely depend on application-specific design requirements.

Encapsulation and compartmentalization offer ways to enforce modularity 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 (Hsiao, Hori, Rothemund, & Murray, 2016; Regot et al., 2011; Tamsir, Tabor, & Voigt, 2011), theoretical studies are largely missing. This results in limited system-level guidance available for implementation and lack of engineering solutions that are 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.

Finally, the problem of *genetic 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., using tools from optimal control and learning systems) and experimental approach.

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* (Haseltine & Arnold, 2007; Yokobayashi, Weiss, & Arnold, 2002) 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 and 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 design.

Another 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 due to the context.

4.7. Population-level design

Population-based computation, in which cells compute and coordinate with each other to obtain an emergent ecosystem is highly desirable for many applications of synthetic biology, from programmable probiotics to regenerative medicine (Payne & You, 2013). However, while co-existence of cell communities is ubiquitous in natural systems, engineering such an ecosystem where multiple cell populations stably co-exist in a steady state that is robust to extraneous species is a formidable challenge (Armstrong & McGehee, 1976; Balagaddé et al., 2008; Foster & Bell, 2012; Scott et al., 2017; You, Cox, Weiss, & Arnold, 2004). Attractors exist for these systems that are far more complicated than simple steady states. The whole population may be "stable" according to some relaxed notion of stability, but each of the constituent cells may not be and may, instead, dynamically change its state under the laws of physics or simply due to noise. This could be a mechanism for resilience such that the whole system steady state is robust to external interference but its constituent 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 (e.g., 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, due to their heterogeneity. New problem formulations 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 (Hsiao et al., 2016; Regot et al., 2011; Tamsir et al., 2011), allowing to defeat several sources of context dependence at the single cell level, such as retroactivity and competition for transcriptional and translational resources. This requires new compositional/descriptive frameworks that allow for efficient design and verification, despite the large size of the systems considered and a significant amount of communication delay due to slow molecular diffusion.

4.8. Exploiting timescale separation

Timescale separation is a ubiquitous property of biological processes. Different core processes occur on very different timescales, 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, and to weeks for tissue formation (Alon, 2006). The dynamics at the different layers of the design abstraction hierarchy of synthetic biology also occur at well separated timescales, 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, with molecular interactions occurring on the subsecond timescale and average noise affecting gene expression on the minutes to hours timescale. Finally, spatial dynamics are characterized by timescale separation as well. For example, diffusion of mRNA molecules within the cell occurs much faster than that of larger molecules such as ribosomes.

On the one hand, system identification, analysis, and simulation are challenged by the 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 (Gómez-Uribe, Verghese, & Tzafriri, 2008; Herath & Del Vecchio, 2016). On the other hand, timescale separation may enable simplifying models and also reaching some level of "insulation" between processes that occur on different timescales. This property has been exploited to obtain modular synthetic systems (e.g., mitigate retroactivity Mishra et al., 2014; Shah & Del Vecchio, 2017) and to facilitate control design (e.g., approximate integral control Qian & Del Vecchio, 2018). 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 biomolecular 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 on faster timescales. 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 (Fahlberg, Groglio, Toxavidis, Israel, & Tigges, 2012), 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 testbed

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 (Karr et al., 2012); the SSAs to capture noise effects at all molecular counts (Gillespie, 1977); agent-based simulation tools that can account for spatial and cell-cell dynamics (Gorochowski, 2016); and molecular simulation algorithms that account for forces and mechanical interactions involving DNAs (Cheatham & Young, 2000). 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, the design and verification problems

remain combinatorial. 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 (Pardee et al., 2014). Cell-free systems provide a means for "running" a circuit of interest without being susceptible to any unknown interaction with the host cell (Shin & Noireaux, 2012; Takahashi et al., 2015). However, in addition to the issue of resource depletion, which substantially shortens the life span of the circuit of interest, there are still problems of standardization of cell extracts for meaningful quantitative analysis and comparison with in vivo circuits (Nagaraj, Greene, Sengupta, & Sontag, 2017; Niederholtmeyer et al., 2015; Shin & Noireaux, 2012; Takahashi et al., 2015).

5. Summary and outlook

In this paper, we articulated a number of potential systems and control research opportunities 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 (Elowitz & Leibler, 2000; Gardner et al., 2000), the field has progressively moved away from it. 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 modeling, design, and realization of any synthetic biological system relies fundamentally on our understanding and engineering capability of biology. Hence, the contribution from the systems and control community to synthetic biology is conditioned on reaching out to biologists and opening a conversation that will provide control theorists the appropriate domain-specific knowledge. For example, the ability to design dynamics has largely relied on systems and control theory, where a plant of interest that is intrinsically unstable (i.e., highly agile aircraft) 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 interconnections. They will also have to formulate the mathematical problems to reflect the biophysical constraints and to be suitable for biomolecular implementation (see Arpino et al., 2013 for discussions on tunability of common biomolecular parameters), which are significantly different from those in classical engineering settings. Therefore, a strong synergy is required between control theorists and synthetic biologists for these research opportunities to unfold.

Acknowledgements

This paper 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), João 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). We thank Carlos Barajas for helpful discussions on spatial heterogeneity of gene expression.

References

Alberts, B. (2014). *Molecular biology of the cell*. Garland Science.

Alon, U. (2006). *An introduction to systems biology: Design principles of biological circuits*. Chapman & Hall/CRC Press.

Aminzare, Z., & Sontag, E. D. (2014). Synchronization of diffusively-connected nonlinear systems: Results based on contractions with respect to general norms. *IEEE Transactions on Network Science and Engineering, 1*(2), 91–106. doi:10.1109/tnse.2015.2395075.

Aminzare, Z., & Sontag, E. D. (2016). Some remarks on spatial uniformity of solutions of reaction–diffusion PDEs. *Nonlinear Analysis, Theory, Methods & Applications, 147*, 125–144. doi:10.1016/j.na.2016.09.002.

Ander, M., Beltrao, P., Di Ventura, B., Ferkinghoff-Borg, J., Foglierini, M., Kaplan, A., et al. (2004). Smartcell, a framework to simulate cellular processes that combines stochastic approximation with diffusion and localisation: Analysis of simple networks. *Systems Biology (Stevenage), 1*(1), 129–138.

Ang, J., Bagh, S., Ingalls, B. P., & McMillen, D. R. (2010). Considerations for using integral feedback control to construct a perfectly adapting synthetic gene network. *Journal of Theoretical Biology, 266*(4), 723–738. doi:10.1016/j.jtbi.2010.07.034.

Armstrong, R. A., & McGehee, R. (1976). Coexistence of species competing for shared resources. *Theoretical Population Biology, 9*(3), 317–328. doi:10.1016/0040-5809(76)90051-4.

Arpino, J. A. J., Hancock, E. J., Anderson, J., Barahona, M., Stan, G.-B. V., Papachristodoulou, A., & Polizzi, K. (2013). Tuning the dials of synthetic biology. *Microbiology, 159*(7), 1236–1253.

Atkinson, M. R., Savageau, M. A., Myers, J. T., & Ninfa, A. J. (2003). Development of genetic circuitry exhibiting toggle switch or oscillatory behavior in *escherichia coli*. *Cell, 113*(5), 597–607. doi:10.1016/s0092-8674(03)00346-5.

Aubin, J.-P., & Cellina, A. (1984). *Differential inclusions*. Springer Berlin Heidelberg. doi:10.1007/978-3-642-69512-4.

Baker, D., Church, G., Collins, J., Endy, D., Jacobson, J., Keasling, J., et al. (2006). Engineering life: Building a FAB for biology. *Scientific American*, 44–51.

Bakshi, S., Siryaporn, A., Goulian, M., & Weisshaar, J. C. (2012). Superresolution imaging of ribosomes and RNA polymerase in live *escherichia coli* cells. *Molecular Microbiology, 85*(1), 21–38. doi:10.1111/j.1365-2958.2012.08081.x.

Balagaddé, F. K., Song, H., Ozaki, J., Collins, C. H., Barnet, M., Arnold, F. H., et al. (2008). A synthetic *escherichia coli* predator-prey ecosystem. *Molecular Systems Biology, 4*. doi:10.1038/msb.2008.24.

Barkai, N., & Leibler, S. (1997). Robustness in simple biochemical networks. *Nature, 387*(6636), 913–917. doi:10.1038/43199.

Basan, M., Hui, S., Okano, H., Zhang, Z., Shen, Y., Williamson, J. R., & Hwa, T. (2015). Overflow metabolism in *escherichia coli* results from efficient proteome allocation. *Nature, 528*(7580), 99–104. doi:10.1038/nature15765.

Becskei, A., & Serrano, L. (2000). Engineering stability in gene networks by autoregulation. *Nature, 405*(6786), 590–593. doi:10.1038/35014651.

Berthiaume, F., Maguire, T. J., & Yarmush, M. L. (2011). Tissue engineering and regenerative medicine: History, progress, and challenges. *Annual Review of Chemical and Biomolecular Engineering, 2*(1), 403–430. doi:10.1146/annurev-chembioeng-061010-114257.

Bocci, F., Suzuki, Y., Lu, M., & Onuchic, J. N. (2018). Role of metabolic spatiotemporal dynamics in regulating biofilm colony expansion. *Proceedings of the National Academy of Sciences of the United States of America*, 201706920. doi:10.1073/pnas.1706920115.

Brenner, K., You, L., & Arnold, F. H. (2008). Engineering microbial consortia: A new frontier in synthetic biology. *Trends in Biotechnology, 26*(9), 483–489. doi:10.1016/j.tibtech.2008.05.004.

Åström, K. J., & Murray, R. M. (2008). *Feedback systems: An introduction for scientists and engineers*. Princeton University Press.

Briat, C., Gupta, A., & Khammash, M. (2016a). Antithetic integral feedback ensures robust perfect adaptation in noisy biomolecular networks. *Cell Systems, 2*(1), 15–26. doi:10.1016/j.cels.2016.01.004.

Briat, C., Zechner, C., & Khammash, M. (2016b). Design of a synthetic integral feedback circuit: Dynamic analysis and DNA implementation. *ACS Synthetic Biology, 5*(10), 1108–1116. doi:10.1021/acssynbio.6b00014.

Bushell, M., & Sarnow, P. (2002). Hijacking the translation apparatus by RNA viruses. *Journal of Cell Biology, 158*(3), 395–399. doi:10.1083/jcb.200205044.

Cameron, D. E., Bashor, C. J., & Collins, J. J. (2014). A brief history of synthetic biology. *Nature Reviews Microbiology, 12*(5), 381–390.

Cardinale, S., & Arkin, A. P. (2012). Contextualizing context for synthetic biology-identifying causes of failure of synthetic biological systems. *Biotechnology Journal, 7*, 856–866. doi:10.1002/biot.201200085.

Castellana, M., Li, S. H.-J., & Wingreen, N. S. (2016). Spatial organization of bacterial transcription and translation. *Proceedings of the National Academy of Sciences of the United States of America, 113*(33), 9286–9291. doi:10.1073/pnas.1604995113.

Ceroni, F., Algar, R., Stan, G.-B., & Ellis, T. (2015). Quantifying cellular capacity identifies gene expression designs with reduced burden. *Nature Methods, 12*(5), 415–422. doi:10.1038/nmeth.3339.

Chakravarti, D., & Wong, W. W. (2015). Synthetic biology in cell-based cancer immunotherapy. *Trends in Biotechnology, 33*(8), 449–461. https://doi.org/10.1016/j.tibtech.2015.05.001.

Chan, C. T. Y., Lee, J. W., Cameron, D. E., Bashor, C. J., & Collins, J. J. (2015). 'Deadman' And 'Passcode' microbial kill switches for bacterial containment. *Nature Chemical Biology, 12*, 82–86. doi:10.1038/nchembio.1979.

Chang, Y.-C., Armitage, J. P., Papachristodoulou, A., & Wadhams, G. H. (2013). A single phosphatase can convert a robust step response into a graded, tunable or adaptive response. *Microbiology, 159*(Pt_7), 1276–1285. doi:10.1099/mic.0.066324-0.

Chaves, M., Sontag, E. D., & Albert, R. (2006). Methods of robustness analysis for boolean models of gene control networks. *IEE Proceedings - Systems Biology, 153*(4), 154. doi:10.1049/ip-syb:20050079.

Cheatham, T. E., & Young, M. A. (2000). Molecular dynamics simulation of nucleic acids: successes, limitations, and promise. *Biopolymers, 56*(4), 232–256. doi:10.1002/1097-0282(2000)56:4(232::aid-bip10037)3.0.co;2-h.

Chen, A. H., & Silver, P. A. (2012). Designing biological compartmentalization. *Trends in Cell Biology, 22*(12), 662–670. doi:10.1016/j.tcb.2012.07.002.

Cookson, N. A., Mather, W. H., Danino, T., Mondragón-Palomino, O., Williams, R. J., Tsimring, L. S., & Hasty, J. (2011). Queueing up for enzymatic processing: Correlated signaling through coupled degradation. *Molecular Systems Biology, 7*, 561. doi:10.1038/msb.2011.94.

Csete, M. E., & Doyle, J. (2004). Bow ties, metabolism and disease. *Trends Biotechnology, 22*(9), 446–450. doi:10.1016/j.tibtech.2004.07.007.

Csete, M. E., & Doyle, J. (2002). Reverse engineering of biological complexity. *Science, 295*(5560), 1664–1669. doi:10.1126/science.1069981.

Dameron, C. T., & Harrison, M. D. (1998). Mechanisms for protection against copper toxicity. *The American Journal of Clinical Nutrition, 67*(5), 1091S–1097S. doi:10.1093/ajcn/67.5.1091s.

Del Vecchio, D. (2015). Modularity, context-dependence, and insulation in engineered biological circuits. *Trends Biotechnology, 33*(2), 111–119. doi:10.1016/j.tibtech.2014.11.009.

Del Vecchio, D., Dy, A. J., & Qian, Y. (2016). Control theory meets synthetic biology. *Journal of the Royal Society Interface, 13*(120). doi:10.1098/rsif.2016.0380.

Del Vecchio, D., & Murray, R. M. (2014). *Biomolecular feedback systems*. Princeton: Princeton University Press.

Del Vecchio, D., Ninfa, A. J., & Sontag, E. D. (2008). Modular cell biology: Retroactivity and insulation. *Molecular Systems Biology, 4*(1).

Deng, H., Krstic, M., & Williams, R. (2001). Stabilization of stochastic nonlinear systems driven by noise of unknown covariance. *IEEE Transactions on Automatic Control, 46*(8), 1237–1253. doi:10.1109/9.940927.

Din, M. O., Danino, T., Prindle, A., Skalak, M., Selimkhanov, J., Allen, K., et al. (2016). Synchronized cycles of bacterial lysis for in vivo delivery. *Nature, 536*, 81–85. doi:10.1038/nature18930.

Drengstig, T., Jolma, I. W., Ni, X. Y., Thorsen, K., Xu, X. M., & Ruoff, P. (2012). A basic set of homeostatic controller motifs. *Biophysical Journal, 103*, 2000–2010. doi:10.1016/j.bpj.2012.09.033.

Drengstig, T., Ueda, H. R., & Ruoff, P. (2008). Predicting perfect adaptation motifs in reaction kinetic networks. *The Journal of Physical Chemistry B, 112*(51), 16752–16758. doi:10.1021/jp806818c.

Edelman, G. M., & Gally, J. A. (2001). Degeneracy and complexity in biological systems. *Proceedings of the National Academy of Sciences of the United States of America, 98*(24), 13763–13768. doi:10.1073/pnas.231499798.

Eldar, A., & Elowitz, M. B. (2010). Functional roles for noise in genetic circuits. *Nature, 467*(7312), 167–173. doi:10.1038/nature09326.

Elowitz, M. B. (2002). Stochastic gene expression in a single cell. *Science, 297*(5584), 1183–1186. doi:10.1126/science.1070919.

Elowitz, M. B., & Leibler, S. (2000). A synthetic oscillatory network of transcriptional regulators. *Nature, 403*(6767), 335–338. doi:10.1038/35002125.

Fahlberg, M., Groglio, K., Toxavidis, V., Israel, B., & Tigges, J. (2012). Standardization of flow cytometry instrumentation for the analysis of microparticles. *Journal of Biomolecular Techniques, 23*(Suppl). S38–S38

Feizi, S., Marbach, D., Médard, M., & Kellis, M. (2013). Network deconvolution as a general method to distinguish direct dependencies in networks. *Nature Biotechnology, 31*(8), 726–733. doi:10.1038/nbt.2635.

Foster, K. R., & Bell, T. (2012). Competition, not cooperation, dominates interactions among culturable microbial species. *Current Biology, 22*(19), 1845–1850. doi:10.1016/j.cub.2012.08.005.

Gardner, T. S., Cantor, C. R., & Collins, J. J. (2000). Construction of a genetic toggle switch in *escherichia coli*. *Nature, 403*(6767), 339–342. doi:10.1038/35002131.

Gillespie, D. T. (1977). Exact stochastic simulation of coupled chemical reactions. *Journal of Physical Chemistry, 81*(25), 2340–2361. doi:10.1021/j100540a008.

Gómez-Uribe, C. A., Verghese, G. C., & Tzafriri, A. R. (2008). Enhanced identification and exploitation of time scales for model reduction in stochastic chemical kinetics. *Journal of Chemical Physics, 129*(24), 244112. doi:10.1063/1.3050350.

Good, M. C., Zalatan, J. G., & Lim, W. A. (2011). Scaffold proteins: Hubs for controlling the flow of cellular information. *Science, 332*(6030), 680–686. doi:10.1126/science.1198701.

Gorochowski, T. E. (2016). Agent-based modelling in synthetic biology. *Essays in Biochemistry, 60*(4), 325–336. doi:10.1042/ebc20160037.

Gupta, A., Briat, C., & Khammash, M. (2014). A scalable computational framework for establishing long-term behavior of stochastic reaction networks. *PLOS Computational Biology, 10*(6), e1003669. doi:10.1371/journal.pcbi.1003669.

Gurdon, J. B., & Bourillot, P.-Y. (2001). Morphogen gradient interpretation. *Nature, 413*(6858), 797–803. doi:10.1038/35101500.

Gyorgy, A., Jiménez, J. I., Yazbek, J., Huang, H.-H., Chung, H., Weiss, R., & Del Vecchio, D. (2015). Isocost lines describe the cellular economy of gene circuits. *Biophysical Journal, 109*(3), 639–646. doi:10.1016/j.bpj.2015.06.034.

Hanahan, D., & Weinberg, R. A. (2011). Hallmarks of cancer: The next generation. *Cell, 144*(5), 646–674. doi:10.1016/j.cell.2011.02.013.

Haseltine, E. L., & Arnold, F. H. (2007). Synthetic gene circuits: Design with directed evolution. *Annual Review of Biophysics and Biomolecular Structure, 36*(1), 1–19. doi:10.1146/annurev.biophys.36.040306.132600.

Hattne, J., Fange, D., & Elf, J. (2005). Stochastic reaction-diffusion simulation with mesord. *Bioinformatics, 21*(12), 2923–2924. doi:10.1093/bioinformatics/bti431.

Herath, N., & Del Vecchio, D. (2016). Model order reduction for linear noise approximation using time-scale separation. In *Proceedings of the IEEE 55th conference on decision and control (CDC)*. IEEE. doi:10.1109/cdc.2016.7799173.

Hespanha, J. P. (2006). Modelling and analysis of stochastic hybrid systems. *IEE Proceedings - Control Theory and Applications, 153*(5), 520–535. doi:10.1049/ip-cta:20050088.

Hooshangi, S., Thiberge, S., & Weiss, R. (2005). Ultrasensitivity and noise propagation in a synthetic transcriptional cascade. *Proceedings of the National Academy of Sciences of the United States of America, 102*(10), 3584–3586.

Hsiao, V., De Los Santos, E. L. C., Whitaker, W. R., Dueber, J. E., & Murray, R. M. (2015). Design and implementation of a biomolecular concentration tracker. *ACS Synthetic Biology, 4*(2), 150–161. doi:10.1021/sb500024b.

Hsiao, V., Hori, Y., Rothemund, P. W., & Murray, R. M. (2016). A population-based temporal logic gate for timing and recording chemical events. *Molecular Systems Biology, 12*(5), 869. doi:10.15252/msb.20156663.

Hsiao, V., Swaminathan, A., & Murray, R. M. (2018). Control theory for synthetic biology. *IEEE Control Systems*. In press.

Jacob, F., & Monod, J. (1961). Genetic regulatory mechanisms in the synthesis of proteins. *Journal of Molecular Biology, 3*(3), 318–356. doi:10.1016/s0022-2836(61)80072-7.

Jayanthi, S., Nilgiriwala, K., & Del Vecchio, D. (2013). Retroactivity controls the temporal dynamics of gene transcription. *ACS Synthetic Biology, 2*(8), 431–441. doi:10.1021/sb300098w.

Jiang, P., Ventura, A. C., Sontag, E. D., Merajver, S. D., Ninfa, A. J., & Del Vecchio, D. (2011). Load-induced modulation of signal transduction networks. *Science Signaling, 4*(194). doi:10.1126/scisignal.2002152. ra67–ra67

Kang, T., Moore, R., Li, Y., Sontag, E., & Bleris, L. (2015). Discriminating direct and indirect connectivities in biological networks. *Proceedings of the National Academy of Sciences of the United States of America, 112*(41), 12893–12898. doi:10.1073/pnas.1507168112.

Karr, J. R., Sanghvi, J. C., Macklin, D. N., Gutschow, M. V., Jacobs, J. M., Bolival, B. J., et al. (2012). A whole-cell computational model predicts phenotype from genotype. *Cell, 150*(2), 389–401. doi:10.1016/j.cell.2012.05.044.

Khalil, A. S., & Collins, J. J. (2010). Synthetic biology: Applications come of age. *Nature Reviews Genetics, 11*, 367–379.

Khammash, M. (2009). Modeling and analysis of stochastic biochemical networks. In *Control theory and systems biology* (pp. 29–44). The MIT Press. doi:10.7551/mitpress/9780262013345.003.0002.

Kirschner, M. W., & Gerhart, J. C. (2006). *The plausibility of life: Resolving darwins dilemma*. Yale University Press.

Kitano, H. (2007). Towards a theory of biological robustness. *Molecular Systems Biology, 3*. doi:10.1038/msb4100179.

Klann, M., & Koeppl, H. (2012). Spatial simulations in systems biology: From molecules to cells. *International Journal of Molecular Sciences, 13*(12), 7798–7827. doi:10.3390/ijms13067798.

Klavins, E. (2010). Proportional-integral control of stochastic gene regulatory networks. In *Proceeding of the 49th IEEE conference on decision and control* (pp. 2547–2553). doi:10.1109/CDC.2010.5717525. Atlanta, GA

Kline, R. (1993). Harold black and the negative-feedback amplifier. *IEEE Control Systems Magazine, 13*(4), 82–85. doi:10.1109/37.229565.

Kosuri, S., Goodman, D. B., Cambray, G., Mutalik, V. K., Gao, Y., Arkin, A. P., et al. (2013). Composability of regulatory sequences controlling transcription and translation in *escherichia coli*. *Proceedings of the National Academy of Sciences of the United States of America, 110*(34), 14024–14029. doi:10.1073/pnas.1301301110.

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

Kurtz, T. G. (1972). The relationship between stochastic and deterministic models for chemical reactions. *Journal of Chemical Physics, 57*(7), 2976–2978. doi:10.1063/1.1678692.

Li, B., & You, L. (2011). Synthetic biology: Division of logic labour. *Nature, 469*(7329), 171–172. doi:10.1038/469171a.

Lillacci, G., Aoki, S., Schweingruber, D., & Khammash, M. (2017). A synthetic integral feedback controller for robust tunable regulation in bacteria. *bioRxiv*. doi:10.1101/170951.

Liu, J., Prindle, A., Humphries, J., Gabalda-Sagarra, M., Asally, M., Lee, D.-Y. D., et al. (2015). Metabolic co-dependence gives rise to collective oscillations within biofilms. *Nature, 523*(7562), 550–554. doi:10.1038/nature14660.

Ma, W., Trusina, A., El-Samad, H., Lim, W. A., & Tang, C. (2009). Defining network topologies that can achieve biochemical adaptation. *Cell, 138*(4), 760–773. doi:10.1016/j.cell.2009.06.013.

Magnusson, L. U., Farewell, A., & Nyström, T. (2005). ppGpp: A global regulator in *escherichia coli*. *Trends in Microbiology, 13*(5), 236–242. doi:10.1016/j.tim.2005.03.008.

Miller, M., Hafner, M., Sontag, E. D., Davidsohn, N., Subramanian, S., Purnick, P. E. M., et al. (2012). Modular design of artificial tissue homeostasis: robust control through synthetic cellular heterogeneity. *PLOS Computational Biology, 8*(7), e1002579. doi:10.1371/journal.pcbi.1002579.

Mishra, D., Rivera, P. M., Lin, A., Del Vecchio, D., & Weiss, R. (2014). A load driver device for engineering modularity in biological networks. *Nature Biotechnology, 32*, 1268–1275. doi:10.1038/nbt.3044.

Moore, R. E. (1979). *Methods and applications of interval analysis*. Society for Industrial and Applied Mathematics. doi:10.1137/1.9781611970906.

Nagaraj, V. H., Greene, J. M., Sengupta, A. M., & Sontag, E. D. (2017). Translation inhibition and resource balance in the TX-TL cell-free gene expression system. *Synthetic Biology, 2*(1). doi:10.1093/synbio/ysx005.

Naghnaeian, M., & Del Vecchio, D. (2017). Robust moment closure method for the chemical master equation. *Proceeding of the IEEE conference on control technology and applications (CCTA)*. IEEE. doi:10.1109/ccta.2017.8062585.

Nevozhay, D., Adams, R. M., Murphy, K. F., Josic, K., & Balazsi, G. (2009). Negative autoregulation linearizes the dose-response and suppresses the heterogeneity of gene expression. *Proceedings of the National Academy of Sciences of the United States of America, 106*(13), 5123–5128. doi:10.1073/pnas.0809901106.

Niederholtmeyer, H., Sun, Z. Z., Hori, Y., Yeung, E., Verpoorte, A., Murray, R. M., & Maerkl, S. J. (2015). Rapid cell-free forward engineering of novel genetic ring oscillators. *eLife, 4*. doi:10.7554/elife.09771.

Norman, T. M., Lord, N. D., Paulsson, J., & Losick, R. (2013). Memory and modularity in cell-fate decision making. *Nature, 503*(7477), 481–486. doi:10.1038/nature12804.

Nuzzo, P., Sangiovanni-Vincentelli, A. L., Bresolin, D., Geretti, L., & Villa, T. (2015). A platform-based design methodology with contracts and related tools for the design of cyber-physical systems. *Proceedings of the IEEE, 103*(11), 2104–2132. doi:10.1109/jproc.2015.2453253.

Olsman, N., Baetica, A.-A., Xiao, F., Leong, Y. P., Doyle, J., & R. , M. (2017). Hard limits and performance tradeoffs in a class of sequestration feedback systems. *bioRxiv*. doi:10.1101/222042.

Pahle, J. (2008). Biochemical simulations: Stochastic, approximate stochastic and hybrid approaches. *Briefings in Bioinformatics, 10*(1), 53–64. doi:10.1093/bib/bbn050.

Pardee, K., Green, A. A., Ferrante, T., Cameron, D. E., DaleyKeyser, A., Yin, P., & Collins, J. J. (2014). Paper-based synthetic gene networks. *Cell, 159*(4), 940–954. doi:10.1016/j.cell.2014.10.004.

Paulsson, J. (2004). Summing up the noise in gene networks. *Nature, 427*(6973), 415–418. doi:10.1038/nature02257.

Payne, S., & You, L. (2013). Engineered cell-cell communication and its applications. In *Productive biofilms* (pp. 97–121). Springer. doi:10.1007/10_2013_249. https://doi.org/10.1007/10_2013_249

Pelkmans, L. (2012). Using cell-to-cell variability–a new era in molecular biology. *Science, 336*(6080), 425–426. doi:10.1126/science.1222161.

Peralta-Yahya, P. P., Zhang, F., del Cardayre, S. B., & Keasling, J. D. (2012). Microbial engineering for the production of advanced biofuels. *Nature, 488*, 320–328.

Pogliano, J., Ho, T. Q., Zhong, Z., & Helinski, D. R. (2001). Multicopy plasmids are clustered and localized in *escherichia coli*. *Proceedings of the National Academy of Sciences of the United States of America, 98*(8), 4486–4491. doi:10.1073/pnas.081075798.

Prajna, S., Jadbabaie, A., & Pappas, G. J. (2007). A framework for worst-case and stochastic safety verification using barrier certificates. *IEEE Transactions on Automatic Control, 52*(8), 1415–1428. doi:10.1109/tac.2007.902736.

Ptashne, M. (2004). *A genetic switch, third edition: Phage lambda revisited*. Cold Spring Harbor Laboratory Press.

Purnick, P. E. M., & Weiss, R. (2009). The second wave of synthetic biology: From modules to systems. *Nature Reviews Molecular Cell Biology, 10*, 410–422. doi:10.1038/nrm2698.

Qian, Y., & Del Vecchio, D. (2018). Realizing 'integral control' in living cells: How to overcome leaky integration due to dilution? *Journal of the Royal Society Interface, 15*(139). doi:10.1098/rsif.2017.0902.

Qian, Y., Huang, H.-H., Jiménez, J. I., & Del Vecchio, D. (2017a). Resource competition shapes the response of genetic circuits. *ACS Synthetic Biology, 6*(7), 1263–1272. doi:10.1021/acssynbio.6b00361.

Qian, Y., McBride, C., & Del Vecchio, D. (2018). Programming cells to work for us. *Annual Review of Control, Robotics, and Autonomous Systems, 1*(1). doi:10.1146/annurev-control-060117-105052.

Quarton, T., Kang, T., Sontag, E. D., & Bleris, L. (2016). Exploring the impact of resource limitations on gene network reconstruction. *Proceeding of the IEEE 55th conference on decision and control (CDC)*. IEEE. doi:10.1109/cdc.2016.7798773.

Raj, A., & van Oudenaarden, A. (2008). Nature, nurture, or chance: Stochastic gene expression and its consequences. *Cell, 135*(2), 216–226. doi:10.1016/j.cell.2008.09.050.

Regot, S., Macia, J., Conde, N., Furukawa, K., Kjellén, J., Peeters, T., et al. (2011). Distributed biological computation with multicellular engineered networks. *Nature, 469*(7329), 207–211. doi:10.1038/nature09679.

Rosenfeld, N., Young, J. W., Alon, U., Swain, P. S., & Elowitz, M. B. (2005). Gene regulation at the single-cell level. *Science, 307*(5717), 1962–1965. doi:10.1126/science.1106914.

Ruder, W. C., Lu, T., & Collins, J. J. (2011). Synthetic biology moving into the clinic. *Science, 333*, 1248–1252.

Russo, G., di Bernardo, M., & Sontag, E. D. (2013). A contraction approach to the hierarchical analysis and design of networked systems. *IEEE Transactions on Automatic Control, 58*(5), 1328–1331. doi:10.1109/tac.2012.2223355.

Saxena, P., Heng, B.-C., Bai, P., Folcher, M., Zulewski, H., & Fussenegger, M. (2016). A programmable synthetic lineage-control network that differentiates human IPSCs into glucose-sensitive insulin-secreting beta-like cells. *Nature Communications, 7*, 11247. doi:10.1038/ncomms11247.

Scott, S. R., Din, M. O., Bittihn, P., Xiong, L., Tsimring, L. S., & Hasty, J. (2017). A stabilized microbial ecosystem of self-limiting bacteria using synthetic quorum-regulated lysis. *Nature Microbiology, 2*(17083). doi:10.1038/nmicrobiol.2017.83.

Shah, R., & Del Vecchio, D. (2017). Signaling architectures that transmit unidirectional information despite retroactivity. *Biophysical Journal, 113*(3), 728–742. doi:10.1016/j.bpj.2017.06.019.

Shao, B., Liu, X., Zhang, D., Wu, J., & Ouyang, Q. (2015). From boolean network model to continuous model helps in design of functional circuits. *PLOS ONE, 10*(6), e0128630. doi:10.1371/journal.pone.0128630.

Shimoga, V., White, J. T., Li, Y., Sontag, E., & Bleris, L. (2014). Synthetic mammalian transgene negative autoregulation. *Molecular Systems Biology, 9*(1). doi:10.1038/msb.2013.27. 670–670

Shin, J., & Noireaux, V. (2012). An e. coli cell-free expression toolbox: Application to synthetic gene circuits and artificial cells. *American Chemical Society, 1*(1), 29–41. doi:10.1021/sb200016s.

Siciliano, V., Garzilli, I., Fracassi, C., Criscuolo, S., Ventre, S., & di Bernardo, D. (2013). Mirnas confer phenotypic robustness to gene networks by suppressing biological noise. *Nature Communications, 4*. doi:10.1038/ncomms3364.

Simon, H. A. (1962). The architecture of complexity. *Proceedings of the American Philosophical Society, 106*(6), 467–482.

Singh, A. (2011). Negative feedback through mrna provides the best control of gene-expression noise. *IEEE Transactions on NanoBioscience, 10*(3), 194–200. doi:10.1109/tnb.2011.2168826.

Singh, A., & Hespanha, J. P. (2011). Approximate moment dynamics for chemically reacting systems. *IEEE Transactions on Automatic Control, 56*(2), 414–418. doi:10.1109/tac.2010.2088631.

Sontag, E. (1989). Smooth stabilization implies coprime factorization. *IEEE Transactions on Automatic Control, 34*(4), 435–443. doi:10.1109/9.28018.

Steel, H., Lillacci, G., Khammash, M., & Papachristodoulou, A. (2017). Challenges at the interface of control engineering and synthetic biology. *Proceeding of the IEEE 56th annual conference on decision and control (CDC)*. IEEE. doi:10.1109/cdc.2017.8263791.

Steiger, C., Walder, H., & Platzner, M. (2004). Operating systems for reconfigurable embedded platforms: Online scheduling of real-time tasks. *IEEE Transactions on Computers, 53*(11), 1393–1407. doi:10.1109/tc.2004.99.

Takahashi, M. K., Hayes, C. A., Chappell, J., Sun, Z. Z., Murray, R. M., Noireaux, V., & Lucks, J. B. (2015). Characterizing and prototyping genetic networks with cell-free transcription–translation reactions. *Methods, 86*, 60–72. doi:10.1016/j.ymeth.2015.05.020.

Tamsir, A., Tabor, J. J., & Voigt, C. A. (2011). Robust multicellular computing using genetically encoded nor gates and chemical 'wires'. *Nature, 469*(7329). doi:10.1038/nature09565. 212–5

Tan, C., Marguet, P., & You, L. (2009). Emergent bistability by a growth-modulating positive feedback circuit. *Nature Chemical Biology, 5*(11), 842–848. doi:10.1038/nchembio.218.

Tripakis, S. (2016). Compositionality in the science of system design. *Proceedings of the IEEE, 104*(5), 960–972. doi:10.1109/jproc.2015.2510366.

van der Meer, J. R., & Belkin, S. (2010). Where microbiology meets microengineering: Design and applications of reporter bacteria. *Nature Reviews Microbiology, 8*, 511–522. doi:10.1038/nrmicro2392.

Van Kampen, N. G. (2007). *Stochastic processes in physics and chemistry, third edition (north-Holland personal library)*. North Holland.

Ventura, A. C., Jiang, P., Wassenhove, L. V., Vecchio, D. D., Merajver, S. D., & Ninfa, A. J. (2010). Signaling properties of a covalent modification cycle are altered by a downstream target. *Proceedings of the National Academy of Sciences of the United States of America, 107*(22), 10032–10037. doi:10.1073/pnas.0913815107.

Villaverde, A. F., Barreiro, A., & Papachristodoulou, A. (2016). Structural identifiability of dynamic systems biology models. *PLOS Computational Biology, 12*(10), e1005153. doi:10.1371/journal.pcbi.1005153.

Whitacre, J. M. (2012). Biological robustness: Paradigms, mechanisms, and systems principles. *Frontiers in Genetics, 3*. doi:10.3389/fgene.2012.00067.

Wintermute, E. H., & Silver, P. A. (2010). Emergent cooperation in microbial metabolism. *Molecular Systems Biology, 6*. doi:10.1038/msb.2010.66.

Wooldridge, M. (2009). *An introduction to multiagent systems*. Wiley.

Wright, O., Stan, G.-B., & Ellis, T. (2013). Building-in biosafety for synthetic biology. *Microbiology, 159*(Pt_7), 1221–1235. doi:10.1099/mic.0.066308-0.

Wu, M., Su, R.-Q., Li, X., Ellisc, T., Lai, Y.-C., & Wang, X. (2013). Engineering of regulated stochastic cell fate determination. *Proceedings of the National Academy of Sciences of the United States of America, 110*(26), 10610–10615.

Xie, Z., Wroblewska, L., Prochazka, L., Weiss, R., & Benenson, Y. (2011). Multi-input RNAi-based logic circuit. *Science, 333*, 1307–1312.

Yamagishi, J. F., Saito, N., & Kaneko, K. (2016). Symbiotic cell differentiation and cooperative growth in multicellular aggregates. *PLOS Computational Biology, 12*(10), e1005042. doi:10.1371/journal.pcbi.1005042.

Yeung, E., Dy, A. J., Martin, K. B., Ng, A. H., Del Vecchio, D., Beck, J. L., et al. (2017). Biophysical constraints arising from compositional context in synthetic gene networks. *Cell Systems, 5*(1), 11–24.e12. doi:10.1016/j.cels.2017.06.001.

Yi, T.-M., Huang, Y., Simon, M. I., & Doyle, J. (2000). Robust perfect adaptation in bacterial chemotaxis through integral feedback control. *Proceedings of the Na-*

tional Academy of Sciences of the United States of America, 97(9), 4649–4653. doi:10.1073/pnas.97.9.4649.

Yokobayashi, Y., Weiss, R., & Arnold, F. H. (2002). Directed evolution of a genetic circuit. *Proceedings of the National Academy of Sciences of the United States of America, 99*(26), 16587–16591. doi:10.1073/pnas.252535999.

You, L., Cox, R. S., Weiss, R., & Arnold, F. H. (2004). Programmed population control by cell-cell communication and regulated killing. *Nature, 428*(6985), 868–871. doi:10.1038/nature02491.