

On the use of mathematical programs with complementarity constraints in combined topological and parametric design of biochemical enzyme networks

Tinghao Guo and James T. Allison

Industrial and Enterprise Systems Engineering, Institute for Genomic Biology, University of Illinois at Urbana-Champaign, Urbana, USA

ABSTRACT

A new method is presented for biochemical enzyme network design based on direct transcription and mathematical programs with complementarity constraints. Topology and continuous parameters are optimized simultaneously. The case study design objective is to optimize adaptability while maintaining sufficient sensitivity to ensure input change detection. A three-node problem is solved using both simultaneous and single-shooting approaches. The simultaneous approach enables solution of four-node problems; this is a new capability not available through existing approaches such as exhaustive enumeration, and is a step toward designing larger systems. A conventional nested solution strategy was also investigated for a four-node problem where an outer loop solves the discrete topology optimization problem, and an inner loop solves the continuous problem for each candidate topology. The simultaneous approach yields robust network topological designs that are superior to those identified through the single-shooting and nested strategies.

ARTICLE HISTORY

Received 4 May 2015
Accepted 3 May 2016

KEYWORDS

synthetic biology; enzyme network; topology optimization; direct transcription; nonlinear programming

1. Introduction

Synthetic biology aims to invent and design novel devices, systems and organisms to achieve useful biological function (Weber and Fussenegger 2012). Although significant progress has been made in designing and constructing a variety of genetic regulatory networks, these have been limited to small-scale (normally single-function) networks. An important object in synthetic biology is to design and construct complex higher-order biological networks (Lu, Khalil and Collins 2009). In this article, a design strategy using direct transcription (DT) and mathematical programming with complementarity constraints (MPCCs) is presented to solve the combined topology and continuous parameter enzyme network design problem. Direct transcription is employed to solve for optimal continuous parameters. Complementarity constraints (CCs) are used to manage topological decisions. The authors' objective here is to extend existing engineering design optimization techniques in a way that supports the identification of genetic network designs that exhibit desired behaviour with much less computational effort. Furthermore, these efficiency improvements increase the dimension of biological networks that can be designed and implemented successfully. The proposed optimization-based approach offers demonstrated advantages over traditional design

CONTACT James T. Allison  jtalliso@illinois.edu

 Supplemental materials for this article can be accessed at <http://dx.doi.org/10.1080/0305215X.2016.1188091>

methodologies—such as exhaustive enumeration—when solving specific three-node and four-node enzyme network topologies.

1.1. Synthetic biology

Synthetic biology is a new research area that involves the design of biological devices and systems. It encompasses biology, chemistry, mathematics and engineering with a wide range of applications such as biochemical treatment (Weber and Fussenegger 2012), bio-energy (Anderson *et al.* 2010) and bio-sensors (Beal and Bachrach 2006). As a framework for conceptualizing synthetic biology, Andriantoandro *et al.* (2006) compares possible hierarchies between synthetic biology and computer engineering. Within the hierarchy, one can think of DNA, mRNA, genes, proteins and metabolites as the physical layers such as transistors, capacitors and resistors. The biochemical reactions in the device layer that manipulate interactions and processes are analogous to the logical gates in a computer. At the module layer, biologists assemble these biological devices into pathways resembling circuits, which can integrate into cells that help perform certain biological functions, such as specific tissue cells. The design process is similar to that in computer engineering where modules (electrical circuits or chips) are integrated together to form a complete system.

The first wave of synthetic biology occurred when basic parts and devices (promoters, ribosome binding sites and transcriptional repressors) were used to construct small modules with specified functions (Purnick and Weiss 2009). Early examples of genetic regulatory networks include a toggle switch and an oscillating genetic regulatory network developed by Gardner, Cantor and Collins (2000) and Elowitz and Leibler (2000). The toggle switch in *Escherichia coli* was constructed from two repressors and two promoters, where each promoter is inhibited by the repressor in transcription of the other promoter. The toggle exhibits bi-stable behaviour. An oscillating genetic network, also termed a repressilator, consists of three transcriptional repressor promoters, each of which transcribes one of the other promoters, resulting in an oscillating network. Through the combination of mathematical models and experiments, these efforts demonstrate progress towards a more general study of synthetic biology (Judd, Laub and McAdams 2000). In addition to switches and oscillators (Kramer 2004; Isaacs *et al.* 2004; Ham *et al.* 2008; Deans, Cantor and Collins 2007; Ajo-Franklin *et al.* 2007; Friedland *et al.* 2009; Stricker *et al.* 2008; Tigges *et al.* 2009), synthetic biologists have also engineered a wide range of artificial genetic regulatory networks including digital logic evaluators (Rinaudo *et al.* 2007), transcriptional cascades (Hooshangi, Thiberge and Weiss 2005), filters (Sohka *et al.* 2009), and cell–cell communicators (Basu *et al.* 2005; Kobayashi *et al.* 2004).

Computational tools have also contributed to the advancement of synthetic biology. BioBrick, proposed by Knight (2003), is a standard where units of DNA (BioBrick parts) can be used to design and assemble genetic networks. These parts, which include promoters and ribosome binding sites, are stored at the Registry of Standard Biological Parts (parts.igem.org). Despite its simplicity and popularity, the BioBrick standard has limitations due to its inability to construct protein fusions. Anderson *et al.* (2010) presented a similar but newer composition standard (BglBriks) that addresses this shortcoming. Eugene is a domain-specific language for creating and improving synthetic biological systems (Densmore 2014) that permits the use of part specifications and rule-based constraints (Bilitchenko *et al.* 2011) and enables the creation and assembly of biological devices with the aid of a platform-based design environment such as Clotho (Densmore *et al.* 2009). The bio-oriented language developed by Beal and Bachrach (2006, 2008) is particularly suitable for engineered genetic regulatory networks when describing various spatial patterns to construct tissues or organs. The high-level specification described in the Proto language is converted into an abstract genetic regulatory network, and the genetic network can be optimized using the compiler (Beal, Lu and Weiss 2011). Several other programming languages such as Genetic Engineering of Living Cells (GEC) (Pedersen and Phillips 2009) are summarized by Purnick and Weiss (2009).

Established strategies from other engineering domains have been extended to synthetic biology: for instance, feedback control plays an important role in genetic regulatory networks and has been applied to study perfect adaptation in bacterial chemotaxis (Yi *et al.* 2000), maintaining calcium homeostasis in mammals (El-Samad, Goff and Khammash 2002), biophasic responses in the chemo attractant in *Dictyostelium* cells (Yang and Iglesias 2006) and enhanced sensitivity and stability of the genetic networks (Stelling *et al.* 2004). Mathematical modelling and simulations are also crucial tools. Differential equations are used to characterize dynamic behaviours within biological systems, *e.g.* applications for *Bacillus subtilis* chemotaxis (Rao, Kirby and Arkin 2004) and the synthetic transcriptional cascade in *Escherichia coli* (Batt *et al.* 2007). Techniques such as parameter fitting and Monte Carlo sampling are also used, *e.g.* in microbial biofuel production modelling, with simulations implemented in the MATLAB[®] environment (Dunlop, Keasling and Mukhopadhyay 2010). Optimization has been used in the area of biochemical reaction networks (Banga 2008). Linear and nonlinear programs have been applied to metabolic flux balance analysis (Varma and Palsson 1994), metabolic networks (Patil *et al.* 2005), microbial cells (Price, Reed and Palsson 2004), biochemical pathways (Mendes and Kell 1998), and metabolism analysis (Vo and Palsson 2006; Vo, Lee and Palsson 2007). Both discrete and continuous decision variables have been investigated for biochemical networks in terms of mixed integer linear and nonlinear programs (MILP and MINLP), including time-delay inferring (Dasika *et al.* 2004), the inference of regulatory interactions (Thomas *et al.* 2007), metabolic networks (Lee *et al.* 2000; Hatzimanikatis, Floudas and Bailey 1996a; 1996b) and gene and enzymes manipulations based on kinetic models (Vital-Lopez *et al.* 2006). Evolutionary strategies proposed by François and Hakim (2004) were used to create genetic networks with functions of bistate switches and oscillators. AutoBioCAD, a design tool for genetic networks, contains a library of genetic elements and incorporates strategies of enumeration and optimization in the evolutionary scheme to design genetic networks with dynamic behaviours (Rodrigo and Jaramillo 2012).

Despite significant achievements, artificial genetic network design remains a challenging task. Many efforts have concentrated on achieving particular biological functions in genetic networks such as oscillations (Tsai *et al.* 2008; Wagner 2005) and adaptation (Alon *et al.* 1999; Barkai and Leibler 1997; Hornung and Barkai 2008). While exhaustive enumeration strategies have been used for robust design of a *Drosophila* segment polarity network (Ma *et al.* 2006), biochemical adaptation of enzyme networks (Ma *et al.* 2009), ultra-sensitivity and bi-stability of switch-like cellular responses (Shah and Sarkar 2011), AutoBioCAD (Rodrigo and Jaramillo 2012) and engineering system architecture (Herber, Guo and Allison 2016), enumeration is an inefficient strategy that is limited to solution of small-scale problems. For instance, Ma *et al.* (2009) carried out a series of analyses on three-node enzyme networks—a particular class of genetic regulatory network—and successfully identified two primary topologies that produce adaptive behaviour. They used exhaustive enumeration to explore all possible 16,038 three-node network topologies. Each topology was tested for adaptive properties by simulating the network 10,000 times to determine the parameter value combination that produced the best results. This nested exhaustive enumeration approach carries with it extreme computational expense, and the size of networks that can be explored in this manner is therefore very limited.

By 2009 the complexity of successfully implemented genetic regulatory networks, as measured by the number of regulatory regions, had plateaued at six (Purnick and Weiss 2009). While some optimization problems and approaches have been explored (Feng *et al.* 2004; Festa 2007; Balsa-Canto and Banga 2011; Handl, Kell and Knowles 2007; Banga 2008), to the authors' knowledge very limited research has been focused on dynamic optimization for both discrete and continuous variables. Adiwijaya, Barton and Tidor (2006) proposed a biological network design method that used dynamic optimization, but required manual adjustments to network topology. Dasika and Maranas (2008) developed an optimization-based method for computational design of genetic networks using mixed integer dynamic optimization (MIDO), requiring MINLP solvers.

The work presented here builds on previous efforts to apply new engineering design methods to genetic regulatory network design with the objectives of (1) increasing the scale and complexity

of networks that can be explored and designed effectively (a vital step in bringing synthetic biology capabilities toward functionally important systems), and (2) providing important new tools and understanding that support these advancements. This class of design problem is particularly difficult because it involves both topological and continuous design variables, and dynamic behaviour is nonlinear and stochastic. This results in a complex and difficult-to-navigate design space. Studies presented here are based on several simplifying assumptions, but fill important gaps in current design methods and knowledge. The first assumption made here is system class; studies are restricted here to a specific class of genetic regulatory networks—biochemical enzyme networks—and focus on identifying designs that are highly adaptable to network input changes. It is further assumed here that dynamic network behaviour is adequately represented by deterministic nonlinear differential equations. This is admittedly a significant assumption and associated simulation results will certainly deviate in important ways from realizable biochemical behaviour, but the assertion here is that this investigation strategy reveals valuable trends in design results, and, most importantly, *insight into new directions for design methodologies and tools for synthetic biology* with the aim of significantly increased scale and functionality of synthetic systems.

An additional contribution of the work presented here is a novel approach to combined topological and continuous parameter design optimization for dynamic systems. Instead of relying on enumeration or MINLP solvers, the problem is converted to a continuous formulation that simultaneously solves the topological and continuous elements of the problem. Dynamic optimization is managed using direct transcription (DT), and the DT formulation is augmented with complementarity constraints (CCs) to support topological decisions. DT has been combined in the past with CCs to model hybrid dynamic optimization problems to form mathematical programs with complementarity constraints, but not for the topological design of dynamic systems. MPCCs have been formulated for static topological design problems in other engineering domains, but not with direct consideration of system dynamics.

Investigation results are compared against other studies based on deterministic dynamic models, and are found to be consistent with earlier findings. In addition, the new method requires lower computational effort than existing alternatives. These comparisons can only be made for small three-node networks as earlier methods are not capable of managing larger network design problems. The new DT-MPCC solution method is then demonstrated successfully for the design of a larger network, providing evidence that advancements in new optimization-based methods may lead to successful design of complex biochemical networks. Sensitivity studies are performed and additional solution strategies are investigated. It is concluded here that DT-MPCC is useful for the design of moderately larger networks, but that alternatives should be explored creatively to strive for methods with the potential for navigating much larger network design problems. The work reported here is part of a larger ongoing effort to step back from existing design methods in synthetic biology¹ and evaluate whether fundamentally different strategies may help designers design and implement synthetic biological systems with completely new levels of complexity that are essential for practical applications. The specific formulation and method presented here has produced evidence that a new direction in design strategies is worth further investigation. The specific method and formulation presented here (DT-MPCC) is not the final result of the larger effort. It may be moderately generalizable, but is only a step toward a comprehensive design framework that applies to a wide variety of problems and is supported by experimental validation of not only predicted design behaviour, but also design method validation. Immediate next steps include both new strategies for navigating complex topological design spaces, and methods that account for stochastic behaviour—possibly leveraging theory and methods for design optimization under uncertainty (Aoues and Chateaufneuf 2010). Additional problem classes beyond adaptation, as well as transcriptional regulation networks, should also be studied.

The remainder of this section provides an overview of DT and CCs, the two essential components of the new design method. The problem formulation and model is then presented, followed by the results of numerical studies and conclusions.

1.2. Direct transcription

Direct transcription (DT) is a method for dynamic system optimization that has been applied widely to trajectory optimization, optimal control and parameter estimation problems (Hargraves and Paris 1987; Betts 2010) in a variety of domains, such as crystallization and batch distillation (Biegler 2007), aircraft trajectory design (Raghunathan *et al.* 2004; Betts and Cramer 1995), orbital trajectory design (Ozimek, Grebow and Howell 2008) and active suspension system co-design (Allison, Guo and Han 2014). Here the word ‘transcription’ in DT is distinct from the specific meaning in biology; it refers to the transcription of an infinite-dimensional optimal control problem into a large-scale nonlinear programming (NLP) problem via discretization (Betts 2010). It may also be classified as a special case of the all-at-once (AAO) multidisciplinary design optimization (MDO) formulation (Cramer *et al.* 1994; Allison and Herber 2014). DT is applied here for the first time to enzyme network design for optimal dynamic performance. Consider the following general continuous dynamic problem:

$$\min_{\xi(t), \mathbf{u}(t), t_F, \mathbf{p}} J = \phi(\xi(t), \mathbf{u}(t), t_F, \mathbf{p}) \quad (1a)$$

$$\text{s.t. } \dot{\xi} = \mathbf{f}_d(\xi(t), \mathbf{u}(t), t, \mathbf{p}) \quad (1b)$$

$$\mathbf{0} = \mathbf{f}_a(\xi(t), \mathbf{u}(t), t, \mathbf{p}) \quad (1c)$$

$$\mathbf{0} = \psi(\xi(t_F), \mathbf{u}(t_F), t_F, \mathbf{p}). \quad (1d)$$

Equation (1a) defines the cost function for open-loop optimal control problems, where $\xi(t)$ is the state vector and $\mathbf{u}(t)$ is the control input, both of which are functions of time $t \in [t_0, t_F]$. Time-independent parameters are represented by the vector \mathbf{p} . Constraints (1b) and (1c) are the differential and algebraic equations that model the governing system dynamics. These two sets of constraints combined form a system of differential algebraic equations (DAEs) (Brenan, Campbell and Petzold 1996). The system may have initial conditions $\xi(t_0) = \xi_0$, as well as a boundary condition at the final time t_F (Equation 1d).

A conventional solution to this problem involves application of first-order optimality conditions—*e.g.* Pontryagin’s maximum principle (PMP) (Pontryagin *et al.* 1962; Bryson and Ho 1975)—which produce a boundary value problem (BVP). In simple cases a closed-form solution of the BVP may be derived; in more general cases this BVP must be solved numerically. In addition, this solution strategy requires the derivation of system Hamiltonian derivatives, which is difficult in many cases.

A solution to the above problem produces an optimal trajectory $\mathbf{u}_*(t)$ (open-loop optimal control solution), as well as \mathbf{p}_* (optimal time-invariant parameter values) and the corresponding state trajectories and final time. Feedback control can be implemented by replacing $\mathbf{u}(t)$ with control design variables \mathbf{x}_c that parameterize a feedback controller; \mathbf{x}_c can be optimized together with the time-independent parameters \mathbf{p} :

$$\min_{\xi(t), \mathbf{x}_c, \mathbf{p}} J = \phi(\xi(t), \mathbf{u}(t), \mathbf{p}, t_F) \quad (2)$$

(DAE and boundary constraints have been omitted for brevity).

A closed-form solution exists for linear dynamic systems with quadratic objective functions. For nonlinear systems, a PMP-based approach may work, but if not other solution methods are required. One often-used solution method is a nested simulation approach, where the DAE (or ODE if no algebraic constraints exist) is solved using numerical simulation for every design alternative tested by the optimization algorithm. This nested simulation method is often referred to as single-shooting. Consider an optimal feedback control problem where t_F and \mathbf{p} are fixed, and no algebraic or boundary

constraints are present:

$$\min_{\mathbf{x}_c} J = \phi(\Xi, \mathbf{x}_c). \quad (3)$$

Constraints are omitted in this formulation because they are satisfied implicitly via simulation. Here Ξ denotes discretized state variable trajectories, which are solved for using numerical simulation for every new candidate control design \mathbf{x}_c . This problem may be solved using standard NLP algorithms (Bertsekas 1999). While straightforward and intuitive, this approach suffers from several drawbacks. Simulation can lead to cost and constraint functions that are non-smooth or arithmetically inconsistent (Betts 2010). Numerical integration of DAEs may become very time-consuming for large-scale problems (Biegler 2007). In addition, single-shooting cannot handle open loop instabilities (Ascher and Petzold 1998; Flores-Tlacuahuac, Biegler and Saldívar-Guerra 2005), and solutions may be numerically unstable for highly nonlinear or stiff systems.

Multiple-shooting is a related method that addresses instabilities by dividing the simulation into n_T smaller time segments. The system is simulated within each of these segments independently, and defect constraints are included in the NLP formulation to ensure consistency and state continuity between segments at optimization convergence. If the size of time segments is reduced to the limiting value of the individual time step sizes (*i.e.* one time step per time segment), the result is the direct transcription method. No forward simulation is performed. Instead, defect constraint satisfaction ensures approximate satisfaction of the underlying DAEs. More specifically, the time domain is divided into n_t segments or intervals $t_0 < t_1 < \dots < t_{n_t} = t_F$, and both discretized state (Ξ) and control (\mathbf{U}) trajectories are optimization variables in the DT formulation. Time invariant parameters \mathbf{p} may also be optimization variables.

The exact form of the defect constraints depends on the collocation method used to discretize the system equations. For example, the first-order Runge–Kutta method (forward Euler method) and higher-order Runge–Kutta methods, such as the implicit trapezoidal method, can be used to generate defect constraints:

$$\zeta_i(\Xi, \mathbf{x}_c, \mathbf{p}) = \xi_i - \xi_{i-1} - h_i \mathbf{f}_d(\xi_{i-1}, \mathbf{x}_c(t_{i-1}), t_{i-1}, \mathbf{p}), \quad (4)$$

$$\zeta_i(\Xi, \mathbf{x}_c, \mathbf{p}) = \xi_i - \xi_{i-1} - \frac{h_i}{2} (\mathbf{f}_d(\xi_{i-1}, \mathbf{x}_c, t_{i-1}, \mathbf{p}) + \mathbf{f}_d(\xi_i, \mathbf{x}_c, t_i, \mathbf{p})), \quad (5)$$

where h_i is the step size at time t_i . Higher-order collocation methods enable accurate solution with fewer time steps. This helps reduce problem dimension, but DT still involves large-scale NLPs. DT problems, however, often have well-behaved sparsity structures that support efficient solution through strategies such as sparse finite differences and exploiting parallel computations (Allison, Kokkolaras and Papalambros 2007; Betts 2010; Allison, Guo and Han 2014). In this article, a new DT solution method will be demonstrated for the enzyme network design problem.

1.3. Complementarity constraints

Complementarity constraints (CCs) are useful in many applications such as parameter estimation (Kameswaran, Staus and Biegler 2005), metabolic flux networks (Raghunathan 2004), truss topology design (Kočvara and Outrata 2006) and hybrid dynamic systems (Baumrucker and Biegler 2009). CCs are used here to manage enzyme network topology decisions.

Complementarity refers to a relationship between variables where either one (or both) are at the boundary. The mathematical program with complementarity constraints takes the following form

(Baumrucker, Renfro and Biegler 2007; Baumrucker and Biegler 2009):

$$\min_{\mathbf{x}, \mathbf{y}, \mathbf{z}} f(\mathbf{x}, \mathbf{y}, \mathbf{z}) \quad (6a)$$

$$g(\mathbf{x}, \mathbf{y}, \mathbf{z}) \geq 0 \quad (6b)$$

$$h(\mathbf{x}, \mathbf{y}, \mathbf{z}) = 0 \quad (6c)$$

$$0 \leq \mathbf{x} \perp \mathbf{y} \geq 0. \quad (6d)$$

The complementarity constraints are given in the last line and imply that either $x_i = 0$ or $y_i = 0$, and $\mathbf{x} \geq 0, \mathbf{y} \geq 0$. While complementarity constraints (CCs) can be used to represent discrete decisions, the resulting nonlinear programming problems do not satisfy regularity requirements such as linear independence constraint qualification (LICQ) or even the weaker Mangasarian–Fromovitz constraint qualification (MFCQ). As a result, multipliers may be non-unique or unbounded (Baumrucker, Renfro and Biegler 2007; Izmailov, Solodov and Uskov 2012). Mathematical programs with complementarity constraints (MPCCs) are therefore difficult to solve using standard NLP solvers.

Several equivalent reformulations have been investigated and analysed using relaxed complementarity constraints (Anitescu, Tseng and Wright 2007; Baumrucker, Renfro and Biegler 2007; Baumrucker and Biegler 2009; Hu and Ralph 2004; Scholtes 2001; Ralph and Wright 2004; Leyffer, López-Calva and Nocedal 2006). One approach is to approximate CCs using a non-negative scalar t ($Reg(t): x_i y_i \leq t, \mathbf{x} \geq 0, \mathbf{y} \geq 0$). Alternatively, the complementarity condition can be approximated by a single constraint ($RegComp(t): \mathbf{x}^T \mathbf{y} \leq t, \mathbf{x} \geq 0, \mathbf{y} \geq 0$) or replaced by equalities ($RegEq(t): x_i y_i = t, \mathbf{x} \geq 0, \mathbf{y} \geq 0$). In addition, an exact l_1 penalty function may be used; the complementarity constraint is enforced using a penalty term in the objective function ($PF(\rho): \min f(\mathbf{x}, \mathbf{y}, \mathbf{z}) + \rho \mathbf{x}^T \mathbf{y}, \mathbf{x} \geq 0, \mathbf{y} \geq 0$ for $\rho > 0$). These MPCC reformulations exhibit different local convergence properties under mild conditions and may be solved using standard optimization algorithms.

2. Design problem

2.1. Problem formulation

Consider a three-node enzyme network topology (Ma *et al.* 2009). Figure 1 illustrates one possible network topology where A receives I and node C produces the output. A third node B acts as a regulator. Figure 2 illustrates a step input change used in this study, and a representative output response. A network topology exhibits adaptation if its output returns to original levels after an input disturbance. Here it is desired to seek network designs with maximally adaptive behaviour, *i.e.* the error E between the steady-state output before and after the step input change is minimized. Equivalently, the precision P —the inverse of E —is to be maximized.

In addition, it is important to maintain an acceptable level of sensitivity (*i.e.* the largest relative output change compared to the relative input change) (Ma *et al.* 2009). Here the sensitivity is required to be greater than or equal to one. The resulting optimization formulation, with normalized error and sensitivity metrics, is

$$\min_{\mathbf{x}} E = \left| \frac{(O_2 - O_1)/O_1}{(I_2 - I_1)/I_1} \right|, \quad \text{subject to} \quad \left| \frac{(O_{\text{peak}} - O_1)/O_1}{(I_2 - I_1)/I_1} \right| \geq 1, \quad (7)$$

where \mathbf{x} represents a combination of topological and continuous design variables.

2.2. Dynamic system model

As discussed by Ma *et al.* (2009), the dynamic output response $O(t)$ of the enzyme network is obtained via numerical simulation to solve the Michaelis–Menten equations that correspond to the candidate

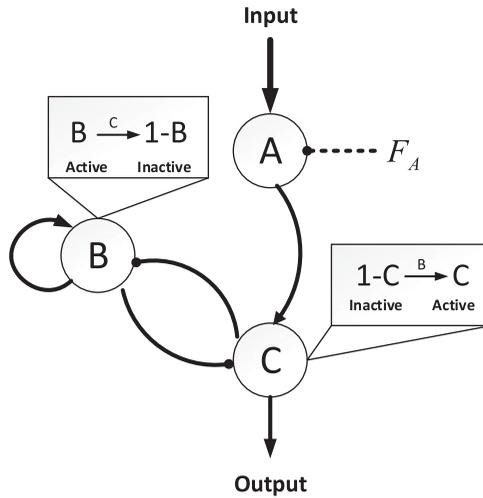


Figure 1. Three-node network topology.

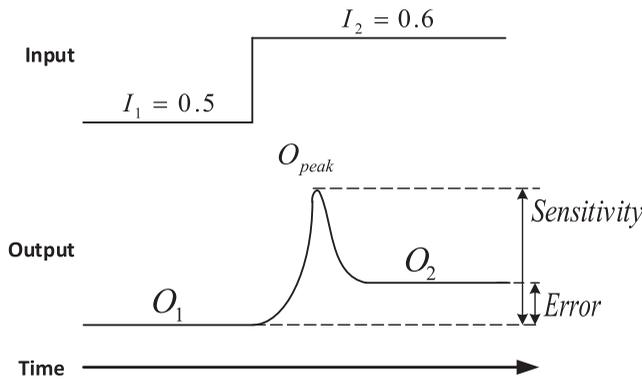


Figure 2. Input and output time responses (similar to Figure 1(a) in Ma et al. [2009]).

network design (Michaelis and Menten 1913; Alon 2007). Within each of this model’s nodes, basally available enzymes or other active network enzymes can oversee the conversion of a normalized enzyme concentration between active and inactive forms. For example, in Figure 1, T and $1-T$ represent the active and inactive levels at node T , for $T \in \{A, B, C\}$, respectively.² The arrow connecting node B to node C represents a positive regulation of node C by node B . This means that active level B (at node B) will convert the enzyme levels at node C from its inactive level $1-C$ to active level C . In other words, the active level B (at node B) will increase the production of the active level C (at node C). A negative regulation is also possible. In Figure 1, the connection between nodes C and B with a circle at node B represents a negative regulation (*i.e.* $C \rightarrow \bullet B$). This means that the active level C (at node C) catalyzes the transition of active form B to inactive form $1-B$ at node B .

Figure 3 illustrates these relations between nodes B and C in greater detail. In Figure 3(a), C and $1-C$ represent the levels of node C ’s active and inactive forms, respectively. The arc with the symbol B above it indicates that the active form level B (produced by node B) causes the transition from its inactive form level $1-C$ to active form level C at node C . The differential equation that models this transition is given at the bottom of Figure 3(a). Here k_{BC} is the catalytic rate constant, and K_{BC} is the Michaelis–Menten constant. Figure 3(b) also outlines the negative regulation of C on B .

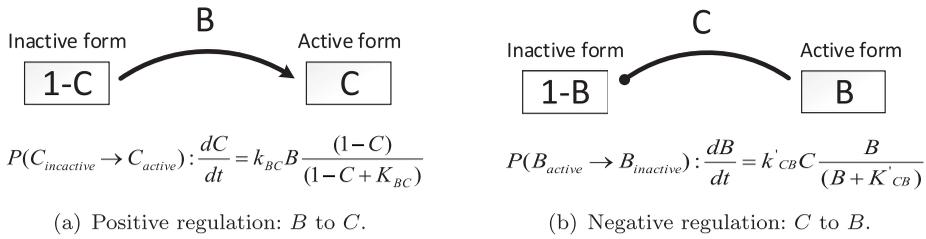


Figure 3. Inner conversion between nodes B and C .

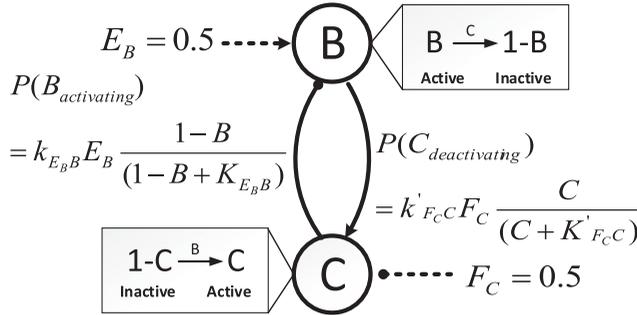


Figure 4. Regulations between nodes B and C .

Often the basal (minimum) enzyme production level is non-zero. If a node has only positive (negative) incoming regulations, it is assumed that a basal enzyme level would deactivate (activate) this node. The notation E_i and F_i is used to represent activating and deactivating basal enzymes receptively, where $i \in \{A, B, C\}$. The basal enzyme in this model is set to be a constant value (0.5), and the corresponding rate equations are illustrated in Figure 4.

An additional regulation type is the self-loop regulation (e.g. the positive self-loop for node B in Figure 1). In this study it is assumed for simplicity that self-loops do not occur. The multiple regulations on a node are additive, resulting in the complete rate equation for node T that can be written as

$$\frac{dT}{dt} = \sum_i X_i k_{X_i T} \frac{1-T}{(1-T+K_{X_i T})} - \sum_i Y_i k'_{Y_i T} \frac{T}{(T+K'_{Y_i T})}, \quad (8)$$

where $T \in \{A, B, C\}$, and X_i is the activating enzyme (positive regulator) of node T , and belongs to the set $\{A, B, C, E_B, E_C\}$. Y_i is the deactivating enzyme (negative regulator) of node T , and is a member of the set $\{A, B, C, F_A, F_B, F_C\}$. The particular values of the catalytic rate constants (k_{cat} 's) and the Michaelis–Menten constants (K_m 's) of the enzymes have the ranges $0 \leq k_{ij} \leq 10$ and $0 \leq K_{ij} \leq 100$, where i and j refer to particular nodes and basal activating/deactivating enzymes.

2.3. MPCC formulation

MPCCs have been applied to many engineering applications including contact mechanics problems, structural design problems, nonlinear obstacle problems, elastohydrodynamic lubrication problems, and traffic equilibrium problems (Ferris and Pang 1997). For instance, in topology optimization, unilateral and frictional contact problems can be formulated as CCs (Panagiotopoulos and Al-Fahed 1994; Stavroulakis 1995; Rodrigues 1993). Kočvara and Outrata (2006) presented a new formulation for truss topology using MPCC and derived optimality conditions for the problem.

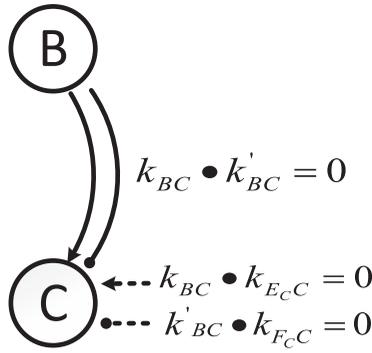


Figure 5. Complementarity constraints.

CCs are used here to model the existence regulations between different nodes in the enzyme network design problem (*i.e.* network topology). CCs are a natural problem formulation as the link between two nodes may only have positive, negative or no regulation. To clarify, node C cannot exert both positive *and* negative regulation on node B in Figure 1. This requirement can be formulated as a complementarity constraint. For example, such a constraint for the three-node network would be $k_{ij}k'_{ij} = 0, \forall \{i, j\} \in \{A, B, C\} \times \{A, B, C\}$.

Figure 5 illustrates how topology decisions may be modelled using CCs. At most one of k_{BC} and k'_{BC} can exist between nodes B and C. This is enforced by the constraint $k_{BC}k'_{BC} = 0$. Likewise, the constraints $k_{BC}k_{E_C C} = 0$ and $k'_{F_C C}k'_{BC} = 0$ ensure that a deactivating or activating enzyme is available when there are only positive or negative linkages coming into node C. Because the network depends on values of parameters \mathbf{k} and \mathbf{K} , the states and the parameters here are considered as optimization variables in the model. If a parameter is zero, its corresponding regulation does not exist; if it is non-zero, the corresponding regulation does exist. This strategy is essentially a ground structure method where possible nodes and interactions are defined *a priori*. Ground structure methods are used widely in structural design problems in topology optimization (Bendsøe and Sigmund 2003; Khetan, Lohan and Allison 2015). The MPCC formulation for the dynamic three-node enzyme network design problem is

$$\begin{aligned}
 \min_{A, B, C, \mathbf{k}, \mathbf{K}} \quad & E = \left| \frac{(O_2 - O_1)/O_1}{(I_2 - I_1)/I_1} \right| \\
 \text{s.t.} \quad & \left| \frac{(O_{\text{peak}} - O_1)/O_1}{(I_2 - I_1)/I_1} \right| \geq 1 \\
 & \frac{dA}{dt} = f_A(A) \quad \frac{dB}{dt} = f_B(B) \quad \frac{dC}{dt} = f_C(C) \\
 & 0 \leq \mathbf{k}_1 \perp \mathbf{k}_2 \geq 0, \quad 0 \leq \mathbf{k} \leq 10, \quad 0 \leq \mathbf{K} \leq 10,
 \end{aligned} \tag{9}$$

where \mathbf{k}_1 and \mathbf{k}_2 are defined for convenience for use with the CCs, and \mathbf{k} and \mathbf{K} are the design parameters to be optimized. The derivative functions $f_A(A)$, $f_B(B)$ and $f_C(C)$ characterize the dynamic interactions using Equation (8) between the nodes. More details regarding the formulation can be found in the supplemental materials for this article. The differential equations were formed into defect constraints $\zeta(\cdot)$ using the trapezoidal collocation method. The complementarity constraints are managed here using a penalty term with weight ρ (Baumrucker, Renfro and Biegler 2007; Baumrucker

and Biegler 2009), and the resulting DT-MPCC formulation is

$$\begin{aligned}
 \min_{\Xi, \mathbf{k}, \mathbf{K}} \quad & \hat{E} = \left(\frac{(O_2 - O_1)/O_1}{(I_2 - I_1)/I_1} \right)^2 + \frac{1}{2} \left(\rho \mathbf{k}_1^T \mathbf{k}_2 \right)^2 \\
 \text{s.t.} \quad & \left| \frac{(O_{\text{peak}} - O_1)/O_1}{(I_2 - I_1)/I_1} \right| \geq 1 \\
 & \zeta_i(\Xi, \mathbf{k}, \mathbf{K}) = \mathbf{0} \\
 & 0 \leq \mathbf{k}_1, \mathbf{k}_2, \mathbf{k} \leq 10, 0 \leq \mathbf{K} \leq 100, \text{ for } i = 1, 2, \dots, n_t - 1, n_t.
 \end{aligned} \tag{10}$$

3. Numerical results

The simultaneous direct transcription (DT) approach was applied to both three- and four-node enzyme network design problems. The successful solution of the four-node problem is an important development; the four-node problem is intractable via enumerative solution strategies due to the increased dimension of the topological search space. The four-node design space contains more than 4.3×10^7 unique design topologies, compared to the three-node design space investigated by Ma *et al.* (2009) that includes 16,038 possible topologies.

In this section, only the optimal topological designs and corresponding system responses are presented. Detailed formulations and optimal continuous parameters can be found in the supplemental materials for this article.

More results regarding steady-state behaviour and sensitivity analyses are also included in the supplemental materials. Complete MATLAB[®] code for this study can be found at the URL listed in the bibliography under Allison (2016). MATLAB[®] package ‘Tools for Network Analysis’, created by the MIT Strategic Engineering Research Group (SERG), was used for network analysis elements of the numerical studies presented here (de Weck 2015).

3.1. Three-node enzyme network design problem

Problem (10) was solved in two different ways for a three-node network, both using the optimization algorithm KNITRO (Byrd, Nocedal and Waltz 2006). First, the simultaneous DT approach was used where the optimization problem and simulation problem were solved at the same time. The second approach utilized single-shooting, where the state variables Ξ were solved for each candidate design $[\mathbf{k}, \mathbf{K}]$ using a multistep, implicit method for forward simulation.³ This second method is a nested simulation strategy, since simulation is nested within the optimization problem.

A multi-start method was used in each case to improve the probability of locating a global optimum. The simultaneous DT approach exhibited significantly better computational efficiency and robustness than the single-shooting method. 20.83% of starting points using the simultaneous approach produced topologies with $P_* \geq 1$, where P_* is the precision,⁴ while only 0.83% did so when using the single-shooting method. The best result of the simultaneous DT approach ($P_* = 1.6072$, $S = 1.016$) had better precision than the best feasible solution obtained using single-shooting ($P_* = 1.4584$, $S = 1.7963$). Observe that the sensitivity constraint in the single-shooting solution is inactive (*i.e.* > 1), indicating an opportunity for improvement that the single-shooting method was unable to identify. Previous studies have identified advantages to traversing state and design spaces simultaneously (Allison, Kokkolaras and Papalambros 2007; Allison, Guo and Han 2014). A more direct path to the solution can be traced, and often better solutions can be identified.

Figures 6 and 7 illustrate the network topologies and system responses produced by both solution approaches. Observe that the system is simulated long enough (40 s) for the output to achieve steady-state conditions before the input is perturbed.⁵ The peak occurs approximately 5 s after the input disturbance. Both responses shown here exhibit adaptation, *i.e.* outputs return to near their original

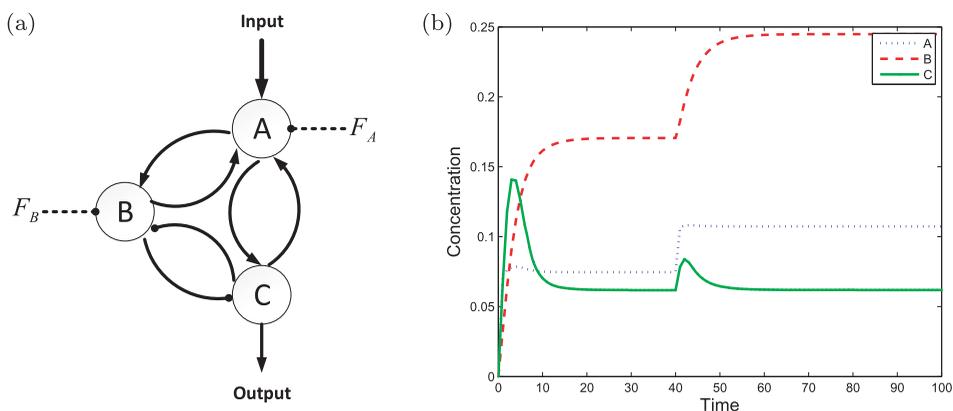


Figure 6. Resulting three-node topology and system response from the single-shooting method.

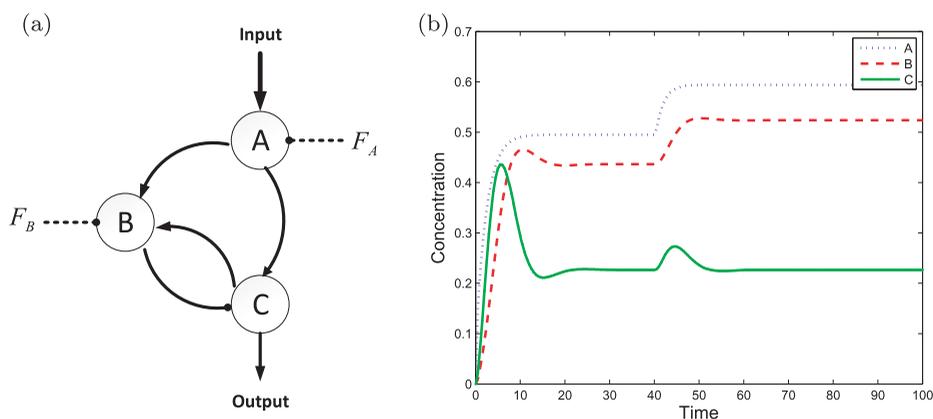


Figure 7. Resulting three-node topology and system response from the simultaneous DT approach.

values. The relative height of the output peak after $t = 40$ s corresponds to the sensitivity metric. A higher peak indicates that a change in input is easier to detect.

3.2. Four-node enzyme network design problem

In this section the investigation is extended to four-node enzyme networks as a step towards larger systems. Notice that 36 time-independent parameters must be optimized in the three-node case. The number of optimization variables increases to 64 for the four-node case. Here it is assumed that node A receives the input signal and node D produces the network output.

As part of this study the DT-MPCC solution method is compared to nested discrete-continuous optimization, a more conventional solution strategy. This is distinct from the nested simulation, or single-shooting, method described above. Nested optimization is described briefly here before presenting numerical results. Nested optimization involves the nested solution of outer- and inner-loop optimization problems. The outer-loop problem seeks to identify the optimal topology, and for each candidate topology tested by the outer-loop problem an inner-loop problem is solved to determine the best possible performance by optimizing continuous parameter values. Using the best performance as determined by the inner loop enables a fair comparison between candidate topologies.

Several solution strategies for the outer-loop integer program were considered. It is a general nonlinear program. The objective function is a black-box function (*i.e.* the result of the inner-loop

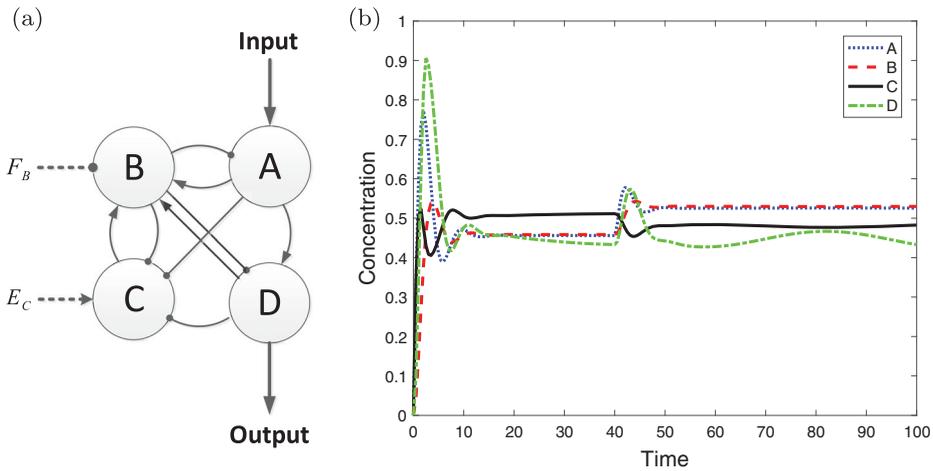


Figure 8. Resulting four-node topology and system response from the nested optimization.

problem) that is evaluated for every candidate topology. Branch and bound is an algorithm for general integer programs, but was discovered to be too computationally expensive for practical solution of this problem. A genetic algorithm (GA) was tested as well for solving the outer-loop problem⁶ (Deb 2000; Deep *et al.* 2009). While optimality cannot be proven when using a GA, it is a practical solution strategy that improves the probability of identifying a global optimal solution in a reasonable time period for difficult problems that are resistant to more efficient integer programming solution methods. Figure 8 illustrates the optimal topology and corresponding system response for the nested optimization method. The optimal topology results in the scaled precision $P_* = 3.3678$ and the active sensitivity constraint $S = 1.0158$.

The nested optimization implementation was difficult to solve and involved significant computational expense compared to the DT-MPCC method (144,000 versus 2,500 simulations). Several factors contribute to this difficulty. The discrete variables in the nested strategy cannot be relaxed to continuous variables and have no special structure, so must be solved using a general nonlinear discrete algorithm. In addition, the inner loop must be solved for each topology candidate, adding to the expense.

Figure 9 shows the optimal topology and corresponding system response using the simultaneous DT-MPCC approach. The optimal-scaled precision in this case is $P_* = 0.9568$, and the sensitivity constraint is active ($S = 1.0048$). Figure 9(b) shows that the steady-states for the system prior to and after the input disturbance have changed less significantly, indicating that a more robust topology is obtained from the simultaneous DT-MPCC approach.

3.3. Steady-state concentrations and sensitivity analyses

For dynamic systems, one can explore the steady-state operating points in a variety of ways. Steady-state behaviour of the optimal three-node and four-node topologies was found both through simulation with sufficient settling time, and also by setting the differential equations to zero and solving for equilibrium points. A multi-start solution method was used to test for multiple equilibrium points. The only physically meaningful equilibrium points found, $[0.5948, 0.5251, 0.2287]$ (three-node, Figure 7(a)) and $[0.5857, 0.5560, 0.5491, 0.2538]$ (four-node, Figure 9(a)), are consistent with the steady-state values obtained via simulation. These tests demonstrate that the simultaneous DT approach obtains correct steady-state values at optimization convergence. Detailed results of this equilibrium study are presented in the supplemental materials for this article.

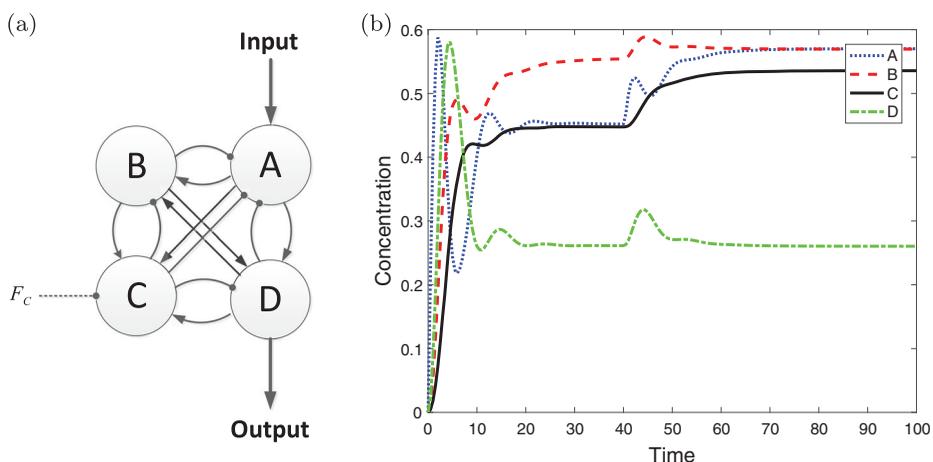


Figure 9. Resulting four-node topology and system response from the simultaneous DT approach.

Local sensitivity analysis for the parameters k_{ij} and K_{ij} was also performed for the same two enzyme networks used for equilibrium analysis. The goal of local sensitivity analysis is to understand how system output responses vary with small changes to parameters and state variables. The reader is referred to Cacuci (2003) and Cacuci, Ionescu-Bujor and Navon (2005) for relevant theory and local methods such as adjoint modelling. Note that this sensitivity analysis is different from the sensitivity function in the optimization formulation that quantifies the ability to detect a change in input from $I = 0.5$ to $I = 0.6$. The R Package FME was used to implement local sensitivity analysis (Soetaert and Petzoldt 2010). As with the equilibrium analysis, the three-node enzyme network of Figure 7(a) and the four-node enzyme network of Figure 9(a) were studied. The sensitivity analysis showed that the incoherent feed-forward loop motif plays a key role in maintaining system stability. In addition, feedback loops were also shown to be important, and the catalytic rate constants k_{ij} had great influence on the system behaviour. Detailed sensitivity analysis findings are presented in the supplemental materials for this article.

4. Conclusion

In this article a dynamic optimization technique based on direct transcription (DT) and mathematical programming with complementarity constraints (MPCCs) was introduced for the simultaneous solution of the enzyme network design problem. DT and CCs were used in tandem to optimize simultaneously both topology and continuous dynamic network design variables. DT was used to transform the ordinary differential equations into algebraic defect constraints, and the MPCC formulation provided a means for determining optimal network topologies. The case study concentrated on enzyme network topologies that could produce adaptive behaviour. Studies were restricted to deterministic models based on the Michaelis–Menten differential equations. The investigation centred on two case studies: three- and four-node enzyme network design problems. In the three-node enzyme network design problem, two solution strategies were tested: direct transcription (simultaneous analysis and design) and single-shooting (simulation performed for each design candidate). The simultaneous DT approach produced better results than single-shooting for the three-node enzyme regulatory network design problem, and was found to be highly efficient. The simultaneous DT optimization approach was then extended to the design of a four-node design problem, which is a complex enough design problem to be intractable in practice via enumerative solution strategies. A conventional nested optimization strategy was also investigated for the four-node problem for comparison. In the nested optimization strategy, an outer loop solves the discrete topology optimization problem,

and a continuous inner loop problem is solved to obtain the best dynamic network performance for each topology candidate tested by the outer-loop problem. Both the nested optimization and simultaneous DT-MPCC strategies were demonstrated to be effective solutions methods with potential to scale up to more complicated problems than can be solved via enumeration. The latter strategy was computationally less expensive, whereas the former produced a somewhat better performing network design.

The solutions obtained here have verified the importance of two core motifs: the feedback and the incoherent feed-forward loops that were identified by Ma *et al.* (2009). This result confirms that optimization-based methods are useful for identifying important topologies in the enzyme networks with reduced computational expense, as well as for supporting the effective design of larger networks. The equilibrium and sensitivity studies provided further insights into resulting optimal network designs. The proposed simultaneous DT approach identifies both the optimal topology and steady-state concentrations. The sensitivity analyses also revealed that the catalytic constants k_{ij} have significant impact on system performance.

While the DT-MPCC method is an important improvement in design methodologies for dynamic network design problems, it so far has resulted only in an incremental increase in the size of networks that can be designed. Realizing the potentially transformative impacts of synthetic biology will require the design of much larger multi-function enzyme regulatory networks with complex interactions. As a step toward the design of larger-scale synthetic biological systems, ongoing work is testing a new design strategy based on generative design algorithms (GDAs). GDAs are a design abstraction that allow the indirect representation of highly complex, variable-dimension designs using a small (fixed) number of parameters by leveraging emergent properties. Another important topic for future work is to investigate design methods that can account for the significant levels of uncertainty in synthetic biology. A large body of work exists in robust design and design under uncertainty, and these existing methods and theory should be studied in the context of advancing synthetic biology.

Notes

1. For example, application of control theory led initially to significant breakthroughs by leveraging its rich heritage, but natural biochemical networks are fundamentally more complex than existing engineering control systems. Continued progress in design strategies based on control theory will certainly lead to important new insights and refinement of synthetic biological systems with current levels of complexity, but is unlikely to support successful creation of systems with new levels of complexity.
2. The notation here follows that used by Ma *et al.* (2009).
3. MATLAB[®] ode15s/ode45 solver.
4. Here the precision reported is scaled as $\log(E^{-1})$.
5. An alternative strategy for finding steady-state conditions before perturbation is to solve $\dot{\xi} = \mathbf{f}_d(\xi(t), \mathbf{p}) = 0$ for ξ . See Section 3.3 for results of this additional equilibrium analysis.
6. MATLAB[®] Global Optimization Toolbox, genetic algorithms for mixed integer optimization.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

The authors would like to recognize that this work was supported in part by funding from the Department of Industrial and Enterprise Systems Engineering at the University of Illinois at Urbana-Champaign (UIUC), and that this work was performed as part of the authors' affiliation with the Institute for Genomic Biology at UIUC.

References

- Adiwijaya, Bambang S., Paul I. Barton, and Bruce. Tidor. 2006. "Biological Network Design Strategies: Discovery Through Dynamic Optimization." *Molecular Biosystems* 2 (12): 650–659. doi:10.1039/B610090B.

- Ajo-Franklin, Caroline M., David A. Drubin, Julian A. Eskin, Elaine P. S. Gee Dirk Landgraf, Ira. Phillips, and Pamela A. Silver. 2007. "Rational Design of Memory in Eukaryotic Cells." *Genes & Development* 21 (18): 2271–2276. doi:10.1101/gad.1586107.
- Allison, J. T. 2016. MPCPC Solution of Genetic Regulatory Circuit Design Problems. <http://www.mathworks.com/matlabcentral/fileexchange/41764-mpcc-solution-of-genetic-regulatory-circuit-design-problems>.
- Allison, J. T., T. Guo, and Z. Han. 2014. "Co-Design of an Active Suspension Using Simultaneous Dynamic Optimization." *Journal of Mechanical Design* 136 (8): Article ID 081003. doi:10.1115/1.4027335.
- Allison, James T., and Daniel R. Herber. 2014. "Multidisciplinary Design Optimization of Dynamic Engineering Systems." *AIAA Journal* 52 (4): 691–710. doi:10.2514/1.J052182.
- Allison, J. T., M. Kokkolaras, and P. Y. Papalambros. 2007. "On Selecting Single-Level Formulations for Complex System Design Optimization." *Journal of Mechanical Design* 129 (9): 898–906. doi:10.1115/1.2747632
- Alon, U. 2007. *An Introduction to Systems Biology: Design Principles of Biological Circuits*. Boca Raton, FL: Chapman and Hall/CRC, Taylor and Francis Group, I.I.C.
- Alon, U., M. G. Surette, N. Barkai, and S. Leibler. 1999. "Robustness in Bacterial Chemotaxis." *Nature* 397 (6715): 168–171. doi:10.1038/16483.
- Anderson, J. C., J. E. Dueber, M. Leguia, G. C. Wu, J. A. Goler, A. P. Arkin, and J. D. Keasling. 2010. "BglBricks: A Flexible Standard for Biological Part Assembly." *Journal of Biological Engineering* 4 (1): 1. doi:10.1186/1754-1611-4-1.
- Andrianantoandro, E., S. Basu, D. K. Karig, and R. Weiss. 2006. "Synthetic Biology: New Engineering Rules for an Emerging Discipline." *Molecular Systems Biology* 2 (1). doi:10.1038/msb4100073.
- Anitescu, Mihai., Paul. Tseng, and Stephen J. Wright. 2007. "Elastic-Mode Algorithms for Mathematical Programs with Equilibrium Constraints: Global Convergence and Stationarity Properties." *Mathematical Programming* 110 (2): 337–371. doi:10.1007/s10107-006-0005-4.
- Aoues, Younes., and Alaa Chateaneuf. 2010. "Benchmark Study of Numerical Methods for Reliability-Based Design Optimization." *Structural and Multidisciplinary Optimization* 41 (2): 277–294. doi:10.1007/s00158-009-0412-2
- Ascher, U. M., and L. R. Petzold. 1998. *Computer Methods for Ordinary Differential Equations and Differential-Algebraic Equations*. Philadelphia, PA: SIAM.
- Balsa-Canto, Eva, and Julio R. Banga. 2011. "AMIGO, A Toolbox for Avanced Model Identification in Systems Biology Using Global Optimization." *Bioinformatics* 27 (16): 2311–2313. doi:10.1093/bioinformatics/btr370
- Banga, Julio R. 2008. "Optimization in Computational Systems Biology." *BMC Systems Biology* 2: doi:10.1186/1752-0509-2-47.
- Barkai, N., and S. Leibler. 1997. "Robustness in Simple Biochemical Networks." *Nature* 387 (6636): 913–917. doi:10.1038/43199
- Basu, S., Y. Gerchman, C. Collins, F. H. Arnold, and R. Weiss. 2005. "A Synthetic Multicellular System for Programmed Pattern Formation." *Nature* 434: 1130–1134. doi:10.1038/nature03461.
- Batt, G., B. Yordanov, R. Weiss, and C. Belta. 2007. "Robustness Analysis and Tuning of Synthetic Gene Networks." *Bioinformatics* 23 (18): 2415–2422. doi:10.1093/bioinformatics/btm362
- Baumrucker, B. T., and L. T. Biegler. 2009. "MPEC Strategies for Optimization of a Class of Hybrid Dynamic Systems." *Journal of Process Control* 19 (8): 1248–1256. doi:10.1016/j.jprocont.2009.02.006
- Baumrucker, B. T., J. G. Renfro, and L. T. Biegler. 2007. "MPEC Problem Formulations in Chemical Engineering Applications." <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.583.7747&rep=rep1&type=pdf>.
- Beal, J., and J. Bachrach. 2006. "Infrastructure for Engineered Emergence on Sensor/Actuator Networks." *IEEE Intelligent Systems* 21 (2): 10–19. doi:10.1109/MIS.2006.29
- Beal, J., and J. Bachrach. 2008. "Cells Are Plausible Targets for High-Level Spatial Languages." In *Proceedings of the 2008 Second IEEE International Conference on Self-Adaptive and Self-Organizing Systems Workshops (SASOW 2008)*, 20–24 October 2008, Venice, 284–291. doi:10.1109/SASOW.2008.14.
- Beal, J., T. Lu, and R. Weiss. 2011. "Automatic Compilation from High-Level Biologically-Oriented Programming Language to Genetic Regulatory Networks." *PloS ONE* 6 (8): Article ID e22490. doi:10.1371/journal.pone.0022490.
- Bendsoe, Martin Philip, and Ole Sigmund. 2003. *Topology Optimization: Theory, Methods and Applications*. Berlin: Springer Science Business Media.
- Bertsekas, D. P. 1999. *Nonlinear Programming*. Belmont, MA: Athena Scientific.
- Betts, J. T. 2010. *Practical Methods for Optimal Control and Estimation Using Nonlinear Programming*. Philadelphia, PA: SIAM.
- Betts, John T., and Evin J. Cramer. 1995. "Application of Direct Transcription to Commercial Aircraft Trajectory Optimization." *Journal of Guidance, Control, and Dynamics* 18 (1): 151–159. doi:10.2514/3.56670
- Biegler, L. T. 2007. "An Overview of Simultaneous Strategies for Dynamic Optimization." *Chemical Engineering and Processing: Process Intensification* 46 (11): 1043–1053. doi:10.1016/j.cep.2006.06.021
- Bilitchenko, L., A. Liu, S. Cheung, E. Weeding, B. Xia, M. Leguia, J. C. Anderson, and D. Densmore. 2011. "Eugene—A Domain Specific Language for Specifying and Constraining Synthetic Biological Parts, Devices, and Systems." *PloS One* 6 (4): Article ID e18882. doi:10.1371/journal.pone.0018882.
- Brenan, K. E., S. L. Campbell, and L. R. Petzold. 1996. *Numerical Solution of Initial-Value Problems in Differential-Algebraic Equations*. Philadelphia, PA: SIAM.

- Bryson, A. E., and Y.-C. Ho. 1975. *Applied Optimal Control: Optimization, Estimation and Control*. Boca Raton, FL: CRC Press.
- Byrd, R. H., J. Nocedal, and R. A. Waltz. 2006. “KNITRO: An Integrated Package for Nonlinear Optimization.” In *Large-scale Nonlinear Optimization*, edited by Gianni Pillo and Massimo Roma, 35–59. New York: Springer.
- Cacuci, Dan G. 2003. *Sensitivity and Uncertainty Analysis*. Vol. I: *Theory*. Boca Raton, FL: CRC Press.
- Cacuci, Dan G., Mihaela Ionescu-Bujor, and Ionel. Michael Navon. 2005. *Sensitivity and Uncertainty Analysis*. Vol. II: *Applications to Large-Scale Systems*. Boca Raton, FL: CRC Press.
- Cramer, Evin, J. J. E. Dennis Jr Paul D. Frank, Robert. Michael Lewis, and Gregory R. Shubin. 1994. “Problem Formulation for Multidisciplinary Optimization.” *SIAM Journal on Optimization* 4 (4): 754–776. doi:10.1137/0804044
- Dasika, Madhukar S., Anshuman. Gupta, Costas D. Maranas, and Jeffrey D. Varner. 2004. “A Mixed Integer Linear Programming (MILP) Framework for Inferring Time Delay in Gene Regulatory Networks.” In *Pacific Symposium on Biocomputing*, 474–485. Hackensack, NJ: World Scientific.
- Dasika, Madhukar S., and Costas D. Maranas. 2008. “OptCircuit: An Optimization Based Method for Computational Design of Genetic Circuits.” *BMC Systems Biology* 2 (1): Article ID 24. doi:10.1186/1752-0509-2-24.
- de Weck, O. L. 2015. “Strategic Engineering Research Group (SERG).” <http://strategic.mit.edu>.
- Deans, T. L., C. R. Cantor, and J. J. Collins. 2007. “A Tunable Genetic Switch Based on RNAi and Repressor Proteins for Regulating Gene Expression in Mammalian Cells.” *Cell* 130 (2): 363–372. doi:10.1016/j.cell.2007.05.045
- Deb, Kalyanmoy. 2000. “An Efficient Constraint Handling Method for Genetic Algorithms.” *Computer Methods in Applied Mechanics and Engineering* 186 (2): 311–338. doi:10.1016/S0045-7825(99)00389-8
- Deep, Kusum., Krishna Pratap Singh, M. L. Kansal, and C. Mohan. 2009. “A Real Coded Genetic Algorithm for Solving Integer and Mixed Integer Optimization Problems.” *Applied Mathematics and Computation* 212 (2): 505–518. doi:10.1016/j.amc.2009.02.044
- Densmore, D. 2014. “CIDAR at Boston University—Eugene.” <http://cidarlab.org/eugene/>.
- Densmore, D., A. Van Devender, M. Johnson, and N. Sritanyaratana. 2009. “A Platform-Based Design Environment for Synthetic Biological Systems.” In *Proceedings of the Fifth Richard Tapia Celebration of Diversity in Computing Conference: Intellect, Initiatives, Insight, and Innovations (TAPIA '09)*, 1–4 April 2009, Portland, OR, 24–29. New York: ACM.
- Dunlop, M. J., J. D. Keasling, and A. Mukhopadhyay. 2010. “A Model for Improving Microbial Biofuel Production Using a Synthetic Feedback Loop.” *Systems and Synthetic Biology* 4 (2): 95–104. doi:10.1007/s11693-010-9052-5.
- El-Samad, H., J. P. Goff, and M. Khammash. 2002. “Calcium Homeostasis and Parturient Hypocalcemia: An Integral Feedback Perspective.” *Journal of Theoretical Biology* 214 (1): 17–29. doi:10.1006/jtbi.2001.2422
- Elowitz, M. B., and S. Leibler. 2000. “A Synthetic Oscillatory Network of Transcriptional Regulators.” *Nature* 403: 335–338. doi:10.1038/35002125
- Feng, X., S. Hooshangi, D. Chen, G. Li, R. Weiss, and H. Rabitz. 2004. “Optimizing Genetic Circuits by Global Sensitivity Analysis.” *Biophysical Journal* 87 (4): 2195–2202. doi:10.1529/biophysj.104.044131.
- Ferris, Michael C., and Jong-Shi Pang. 1997. “Engineering and Economic Applications of Complementarity Problems.” *SIAM Review* 39 (4): 669–713. doi:10.1137/S0036144595285963
- Festa, P. 2007. “On Some Optimization Problems in Molecular Biology.” *Mathematical Biosciences* 207 (2): 219–234. doi:10.1016/j.mbs.2006.11.012
- Flores-Tlacuahuac, A., L. T. Biegler, and E. Saldívar-Guerra. 2005. “Dynamic Optimization of HIPS Open-Loop Unstable Polymerization Reactors.” *Industrial and Engineering Chemistry Research* 44 (8): 2659–2674. doi:10.1021/ie049534p
- François, Paul., and Vincent Hakim. 2004. “Design of Genetic Networks with Specified Functions by Evolution in Silico.” *Proceedings of the National Academy of Sciences of the United States of America* 101 (2): 580–585. doi:10.1073/pnas.0304532101
- Friedland, A. E., T. K. Lu, X. Wang, D. Shi, G. Church, and J. J. Collins. 2009. “Synthetic Gene Networks that Count.” *Science* 324 (5931): 1199–1202. doi:10.1126/science.1172005.
- Gardner, T. S., C. R. Cantor, and J. J. Collins. 2000. “Construction of a Genetic Toggle Switch in *Escherichia coli*.” *Nature* 403 (6767): 339–342. doi:10.1038/35002131
- Ham, T. S., S. K. Lee, A. P. Keasling, and J. D. Arkin. 2008. “Design and Construction of a Double Inversion Recombination Switch for Heritable Sequential Genetic Memory.” *PLoS ONE* 3 (7): Article ID e2815. doi:10.1371/journal.pone.0002815.
- Handl, Julia, Douglas B. Kell, and Joshua Knowles. 2007. “Multiobjective Optimization in Bioinformatics and Computational Biology.” *IEEE/ACM Transactions on Computational Biology and Bioinformatics (TCBB)* 4 (2): 279–292. doi:10.1109/TCBB.2007.070203
- Hargraves, C. R., and S. W. Paris. 1987. “Direct Trajectory Optimization Using Nonlinear Programming and Collocation.” *Journal of Guidance, Control, and Dynamics* 10 (4): 338–342. doi:10.2514/3.20223
- Hatzimanikatis, Vassily., Christodoulos A. Floudas, and James E. Bailey. 1996a. “Analysis and Design of Metabolic Reaction Networks via Mixed-Integer Linear Optimization.” *AIChE Journal* 42 (5): 1277–1292. doi:10.1002/aic.690420509

- Hatzimanikatis, Vassily., Christodoulos A. Floudas, and James E. Bailey. 1996b. "Optimization of Regulatory Architectures in Metabolic Reaction Networks." *Biotechnology and Bioengineering* 52 (4): 485–500. doi:10.1002/(SICI)1097-0290(19961120)52:4 < 485::AID-BIT4 > 3.0.CO;2-L
- Herber, D. R., T. Guo, and J. T. Allison. 2016. "Enumeration of Architectures with Perfect Matchings." In *ASME 2016 International Design Engineering Technical Conference and Computer and Information in Engineering Conference (IDETC/CIE)*, 21–24 August 2016, Charlotte, NC, Paper No. IDETC2016-60212.
- Hooshangi, S., S. Thiberge, and R. Weiss. 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): 3581–3586. doi:10.1073/pnas.0408507102
- Hornung, G., and N. Barkai. 2008. "Noise Propagation and Signaling Sensitivity in Biological Networks: A Role for Positive Feedback." *PLoS Computational Biology* 4 (1): Article ID e8. doi:10.1371/journal.pcbi.0040008.
- Hu, X. M., and D. Ralph. 2004. "Convergence of a Penalty Method for Mathematical Programming with Complementarity Constraints." *Journal of Optimization Theory and Applications* 123 (2): 365–390. doi:10.1007/s10957-004-5154-0.
- Isaacs, F. J., D. J. Dwyer, C. Ding, D. D. Pervouchine, C. R. Cantor, and J. J. Collins. 2004. "Engineered Riboregulators Enable Post-Transcriptional Control of Gene Expression." *Nature Biotechnology* 22 (7): 841–847. doi:10.1038/nbt986.
- Izmailov, A., M. V. Solodov, and E. I. Uskov. 2012. "Global Convergence of Augmented Lagrangian Methods Applied to Optimization Problems with Degenerate Constraints, Including Problems with Complementarity Constraints." *SIAM Journal on Optimization* 22 (4): 1579–1606. doi:10.1137/120868359
- Judd, E. M., M. T. Laub, and H. H. McAdams. 2000. "Toggles and Oscillators: New Genetic Circuit Designs." *BioEssays* 22 (6): 507–509. doi:10.1002/(SICI)1521-1878(200006)22:6 < 507::AID-BIES3 > 3.0.CO;2-0
- Kameswaran, S., G. Staus, and L. T. Biegler. 2005. "Dynamic Optimization for the Core-Flooding Problem in Reservoir Engineering." *Computers & Chemical Engineering* 29 (8): 1787–1800. doi:10.1016/j.compchemeng.2005.02.038
- Khetan, Ashish., Danny J. Lohan, and James T. Allison. 2015. "Managing Variable-Dimension Structural Optimization Problems Using Generative Algorithms." *Structural and Multidisciplinary Optimization* 52 (4): 695–715. doi:10.1007/s00158-015-1262-8.
- Knight, Tom. 2003. *Idempotent Vector Design for Standard Assembly of Biobricks*. Technical report, Artificial Intelligence Laboratory: MIT. <http://dspace.mit.edu/bitstream/handle/1721.1/21168/biobricks.pdf>
- Kobayashi, H., M. Karen, M. Araki, K. Chung, T. S. Gardner, and C. R. Cantor. 2004. "Programmable Cells: Interfacing Natural and Engineered Gene Networks." *Proceedings of the National Academy of Sciences of the United States of America* 101 (22): 8414–8149. doi:10.1073/pnas.0402940101.
- Kočvara, M., and J. Outrata. 2006. "Effective Reformulations of the Truss Topology Design Problem." *Optimization and Engineering* 7 (2): 201–219. <http://www.springerlink.com/index/10.1007/s11081-006-6839-z>.
- Kramer, B. P. 2004. "An Engineered Epigenetic Transgene Switch in Mammalian Cells." *Nature Biotechnology* 22 (7): 867–870. <http://www.ncbi.nlm.nih.gov/pubmed/15184906>.
- Lee, Sangbum., Chan Phalakornkule, Michael M. Domach, and Ignacio E. Grossmann. 2000. "Recursive MILP Model for Finding All the Alternate Optima in LP Models for Metabolic Networks." *Computers & Chemical Engineering* 24 (2): 711–716. doi:10.1016/S0098-1354(00)00323-9
- Leyffer, Sven, Gabriel López-Calva, and Jorge Nocedal. 2006. "Interior Methods for Mathematical Programs with Complementarity Constraints." *SIAM Journal on Optimization* 17 (1): 52–77. doi:10.1137/040621065
- Lu, T. K., A. S. Khalil, and J. J. Collins. 2009. "Next-Generation Synthetic Gene Networks." *Nature Biotechnology* 27 (12): 1139–1150. doi:10.1038/nbt.1591.
- Ma, W., L. Lai, Q. Ouyang, and C. Tang. 2006. "Robustness and Modular Design of the Drosophila Segment Polarity Network." *Molecular Systems Biology* 2: 70. doi:10.1038/msb4100111
- Ma, W., A. Trusina, H. El-Samad, W. Lim, and C. Tang. 2009. "Defining Network Topologies that Can Achieve Biochemical Adaptation." *Cell* 138 (4): 760–773. doi:10.1016/j.cell.2009.06.013.
- Mendes, Pedro, and D. Kell. 1998. "Non-Linear Optimization of Biochemical Pathways: Applications to Metabolic Engineering and Parameter Estimation." *Bioinformatics* 14 (10): 869–883. doi:10.1093/bioinformatics/14.10.869
- Michaelis, L., and M. L. Menten. 1913. "Die Kinetik der Invertinwirkung." *Biochemische Zeitschrift* 49: 333–369.
- Ozimek, M. T., D. J. Grebow, and K. C. Howell. 2008. "Solar Sails and Lunar South Pole Coverage." In *Proceedings of the 2008 AIAA/AAS Astrodynamics Specialist Conference and Exhibit*, 18–21 August 2008, Honolulu, HI. https://engineering.purdue.edu/people/kathleen.howell.1/Publications/Conferences/2008_AIAA_OziGreHow.pdf.
- Panagiotopoulos, P. D., and A. M. Al-Fahed. 1994. "Robot Hand Grasping and Related Problems: Optimal Control and Identification." *The International Journal of Robotics Research* 13 (2): 127–136. doi:10.1177/027836499401300203
- Patil, Kiran R., Isabel Rocha, Jochen Forster, and Jens Nielsen. 2005. "Evolutionary Programming as a Platform for *in Silico* Metabolic Engineering." *BMC Bioinformatics* 6: Article ID 308. doi:10.1186/1471-2105-6-308.
- Pedersen, M., and A. Phillips. 2009. "Towards Programming Languages for Genetic Engineering of Living Cells." *Journal of the Royal Society Interface* 6 (Suppl 4): S437–S450. doi:10.1098/rsif.2008.0516.focus
- Pontryagin, V., R. Boltyanskii, E. Gamkrelidze, E. F. Mishchenko, K. N. Trirogoff, and L. W. Neustadt. 1962. *The Mathematical Theory of Optimal Processes*. New York: Interscience.

- Price, Nathan D., Jennifer L. Reed, and Bernhard Ø. Palsson. 2004. "Genome-scale Models of Microbial Cells: Evaluating the Consequences of Constraints." *Nature Reviews Microbiology* 2 (11): 886–897. doi:10.1038/nrmicro1023
- Purnick, P., and R. Weiss. 2009. "The Second Wave of Synthetic Biology: From Modules to Systems." *Nature Reviews Molecular Cell Biology* 10 (6): 410–422. doi:10.1038/nrm2698
- Raghunathan, A. 2004. "Mathematical Programs with Equilibrium Constraints (MPECs) in Process Engineering." PhD diss., Carnegie Mellon University.
- Raghunathan, Arvind U., Vipin Gopal, Dharmashankar Subramanian, Lorenz T. Biegler, and Tariq Samad. 2004. "Dynamic Optimization Strategies for Three-Dimensional Conflict Resolution of Multiple Aircraft." *Journal of Guidance, Control, and Dynamics* 27 (4): 586–594. doi:10.2514/1.11168
- Ralph, D., and S. J. Wright. 2004. "Some Properties of Regularization and Penalization Schemes for MPECs." *Optimization Methods and Software* 19 (5): 527–556. doi:10.1080/10556780410001709439.
- Rao, C. V., J. R. Kirby, and A. P. Arkin. 2004. "Design and Diversity in Bacterial Chemotaxis: A Comparative Study in *Escherichia coli* and *Bacillus subtilis*." *PLoS Biology* 2 (2): Article ID E49. doi:10.1371/journal.pbio.0020049.
- Rinaudo, Keller., Leonidas Bleris, Rohan Maddamsetti, Sairam Subramanian, Ron Weiss, and Yaakov Benenson. 2007. "A Universal RNAi-based Logic Evaluator that Operates in Mammalian Cells." *Nature Biotechnology* 25: 795–801. doi:10.1038/nbt1307.
- Rodrigo, Guillermo., and Alfonso. Jaramillo. 2012. "AutoBioCAD: Full Biodesign Automation of Genetic Circuits." *ACS Synthetic Biology* 2 (5): 230–236. doi:10.1021/sb300084h.
- Rodrigues, H. C.. 1993. "A Mixed Variational Formulation for Shape Optimization of Solids with Contact Conditions." *Structural Optimization* 6 (1): 19–28. doi:10.1007/BF01743171
- Scholtes, S. 2001. "Convergence Properties of a Regularization Scheme for Mathematical Programs with Complementarity Constraints." *SIAM Journal on Optimization* 11 (4): 918–936. doi:10.1137/S1052623499361233
- Shah, N.A., and C. A. Sarkar. 2011. "Robust Network Topologies for Generating Switch-Like Cellular Responses." *PLoS Computational Biology* 7 (6): Article ID e1002085. doi:10.1371/journal.pcbi.1002085.
- Soetaert, K. E. R., and Thomas Petzoldt. 2010. "Inverse Modelling, Sensitivity and Monte Carlo Analysis in R Using Package FME." *Journal of Statistical Software* 33 (3): 1–38. http://www.jstatsoft.org/v33/i03.
- Sohka, T., R. A. Heins, R. M. Phelan, J. M. Greisler, C. A. Townsend, and M. Ostermeier. 2009. "An Externally Tunable Bacterial Band-Pass Filter." *Proceedings of the National Academy of Sciences of the United States of America* 106 (25): 10135–10140. doi:10.1073/pnas.0901246106.
- Stavroulakis, G. 1995. "Optimal Prestress of Structures with Frictional Unilateral Contact Interfaces." *Archive of Applied Mechanics* 66 (1-2): 71–81. doi:10.1007/BF00786690
- Stelling, J., U. Sauer, Z. Szallasi, F. J. Doyle, and J. Doyle. 2004. "Robustness of Cellular Functions." *Cell* 118 (6): 675–685. doi:10.1016/j.cell.2004.09.008
- Stricker, J., S. Cookson, M. R. Bennett, W. H. Mather, L. S. Tsimring, and J. Hasty. 2008. "A Fast, Robust and Tunable Synthetic Gene Oscillator." *Nature* 456 (7221): 516–519. doi:10.1038/nature07389.
- Thomas, Reuben., Carlos J. Paredes, Sanjay Mehrotra, Vassily Hatzimanikatis, and Eleftherios T. Papoutsakis. 2007. "A Model-Based Optimization Framework for the Inference of Regulatory Interactions Using Time-Course DNA Microarray Expression Data." *BMC Bioinformatics* 8: Article ID 228. doi:10.1186/1471-2105-8-228.
- Tigges, M., T. T. Marquez-Lago, J. Stelling, and M. Fussenegger. 2009. "A Tunable Synthetic Mammalian Oscillator." *Nature* 457 (7227): 309–312. doi:10.1038/nature07616.
- Tsai, T. Y., Y. S. Choi, W. Ma, J. R. Pomeroy, C. Tang, and J. E. Ferrell. 2008. "Robust, Tunable Biological Oscillations from Interlinked Positive and Negative Feedback Loops." *Science* 321 (5885): 126–129. doi:10.1126/science.1156951.
- Varma, Amlt, and Bernhard Ø Palsson. 1994. "Metabolic Flux Balancing: Basic Concepts, Scientific and Practical Use." *Nature Biotechnology* 12: 994–998. doi:10.1038/nbt1094-994.
- Vital-Lopez, Francisco G., Antonios Armaou, Evgeni V. Nikolaev, and Costas D. Maranas. 2006. "A Computational Procedure for Optimal Engineering Interventions Using Kinetic Models of Metabolism." *Biotechnology Progress* 22 (6): 1507–1517. doi:10.1021/bp060156o
- Vo, Thuy D., W. N. Paul Lee, and Bernhard Ø Palsson. 2007. "Systems Analysis of Energy Metabolism Elucidates the Affected Respiratory Chain Complex in Leigh's Syndrome." *Molecular Genetics and Metabolism* 91 (1): 15–22. doi:10.1016/j.ymgme.2007.01.012
- Vo, Thuy D., and Bernhard Ø Palsson. 2006. "Isotopomer Analysis of Myocardial Substrate Metabolism: A Systems Biology Approach." *Biotechnology and Bioengineering* 95 (5): 972–983. doi:10.1002/bit.21063
- Wagner, A. 2005. "Circuit Topology and the Evolution of Robustness in Two-Gene Circadian Oscillators." *Proceedings of the National Academy of Sciences of the United States of America* 102 (33): 11775–11780. doi:10.1073/pnas.0501094102.
- Weber, W., and M. Fussenegger. 2012. "Emerging Biomedical Applications of Synthetic Biology." *Nature Reviews Genetics* 13 (1): 21–35. doi:10.1038/nrg3094
- Yang, L., and P. A. Iglesias. 2006. "Positive Feedback May Cause the Biphasic Response Observed in the Chemoattractant-Induced Response of *Dictyostelium* Cells." *Systems and Control Letters* 55 (4): 329–337. doi:10.1016/j.sysconle.2005.08.009.

Yi, T., Y. Huang, M. I. Simon, and J. Doyle. 2000. "Robust Perfect Adaptation in Bacterial Chemotaxis through Integral Feedback Control." *Proceedings of the National Academy of Sciences of the United States of America* 97 (9): 4649–4653. [doi:10.1073/pnas.97.9.4649](https://doi.org/10.1073/pnas.97.9.4649).