Quantifying Side Effects in Multistate Discrete Networks

David Murrugarra\textsuperscript{a} \quad Elena Dimitrova\textsuperscript{b}

June 25, 2019

\textsuperscript{a}Department of Mathematics, University of Kentucky, Lexington, KY 40506-0027 USA.
\textsuperscript{b}School of Mathematical and Statistical Sciences, Clemson University, Clemson, SC 29634-0975 USA.

Abstract

Developing efficient computational methods to change the state of a cell from an undesirable condition, e.g. diseased, into a desirable, e.g. healthy, condition is an important goal of systems biology. The identification of potential interventions can be achieved through mathematical modeling of the state of a cell by finding appropriate input manipulations in the model that represent external interventions. This paper focuses on quantifying the unwanted or unplanned changes that come along with the application of an intervention to produce a desired effect, which we define as the \textit{side effects} of the intervention. The type of mathematical models that we will consider are discrete dynamical systems which include the widely used Boolean networks and their generalizations. The potential control targets can be represented by a set of nodes and edges that can be manipulated to produce a desired effect on the system. This paper presents practical tools along with applications for the analysis and control of multistate networks. The first result is a polynomial normal form representation for discrete functions that provides a partition of the inputs of the function into canalizing and non-canalizing variables and, within the canalizing ones, we categorize the input variables into layers of canalization. The second theoretical result is a set of formulas for counting the maximum number of transitions that will change in the state space upon an edge deletion in the wiring diagram. These formulas rely on the stratification of the inputs of the target function where the number of changed transitions depends on the layer of canalization that includes the input to be deleted. Applications from using these formulas to estimate the number of changes in the state space and comparisons with the actual number of changes are also presented.

1 Introduction

Boolean networks (BNs) have been proposed as an appropriate framework for modeling the state of cells due to their simplicity and the variety of tools available for model analysis. However, some complex gene interactions cannot be represented in the Boolean setting and several generalizations of the Boolean approach have been developed. Multistate models, a generalization of the BN framework, where the genes can attain more than two states have been proposed as appropriate models for capturing complex gene expression patterns, such as consideration of three states (low, medium, and high). This paper presents theoretical results along with applications for the analysis and control of multistate networks.

A Gene Regulatory Network (GRN) is a representation of the intricate relationships among genes, proteins, and other substances that are responsible for the expression levels of mRNA and proteins. Many dynamic systems theory approaches have been used over the last 5-6 decades to develop computational tools for analyzing the dynamics of GRNs. Prominently, Boolean networks have been successfully used to model and study the properties of GRNs [1, 16]. In particular, Boolean canalizing rules were introduced by S. Kauffman and collaborators [13, 14] and reflect the concept of canalization in evolutionary biology that Waddington pioneered in 1942 [30] – that organisms evolve developmental robustness, producing an invariant phenotype even under genetic or environmental perturbations.

In this article, we study the network-wide effect of an experimental intervention that prevents a regulation from happening. Such intervention is modeled through edge deletion and can be achieved via therapeutic drugs that target a specific gene interaction [4, 2]. In [18] we introduced methods
for quantifying side effects in Boolean networks. However, many of the more recently published
discrete dynamical models include variables that take on more than two states due to the need
for capturing mechanisms that are not binary in nature [25, 3, 6, 27]. Consequently, Boolean
nested and partially nested canalizing functions were generalized to multistate [19, 20, 11] which
enables the possibility to capture more complex interactions among the genes in the network. Such
functions can be viewed as a discrete dynamical system with a stratified structure which consists
of hierarchical layers of variables according to their relative influence over the system dynamics.

There are several published control methods for Boolean networks such as Stable Motifs [31],
Feedback Vertex Sets [33], Minimal Hitting Sets [29, 15], and several others [24, 17, 23, 7, 32].
While these control methods focus on finding control targets, there are very few studies focusing
on the consequences of applying a certain control beyond its immediate target. This is analogous
to developing cancer treatments that are only concerned with killing cancer cells without regard
of the effects on the organism. This paper contributes methods for measuring the impact of the
controllers on the dynamics of multistate networks.

The rest of the paper is structured as follows. In Section 2, we introduce discrete dynamical
systems and their representation as polynomial dynamical systems. In Section 3 we define the
control actions for multistate networks. In Section 4 we provide a polynomial normal form for
discrete functions and then we use this representation to derive a set of formulas for counting the
maximum number of transitions in the state space upon edge deletions. In Section 5 we apply our
formulas to two multistate models. Finally, in Section 6 we provide the conclusions of the paper.

2 Background

A discrete dynamical system can be defined as a dynamical system that is discrete in time as well
as in variable states. More formally, consider a collection $x_1, \ldots, x_n$ of variables, each of which can
take on values in finite sets $X_1, \ldots, X_n$. Let $X = X_1 \times \cdots \times X_n$ be their Cartesian product. A
discrete dynamical system in the variables $x_1, \ldots, x_n$ is a function

$$F = (f_1, \ldots, f_n) : X \to X$$

where each coordinate function $f_i$ is a discrete function on a subset of $\{x_1, \ldots, x_n\}$ which represents
how the future value of the $i$-th variable depends on the present values of the variables. If $X_i =
\{0, 1\}$, then each $f_i$ is a Boolean rule and $F$ is a Boolean network.

In this article, for the purpose of exploiting the algebraic properties of discrete functions, it
is assumed that the variables $x_1, \ldots, x_n$ take on values from a finite field $\mathbb{F}$. Then using the fact
that any discrete function $f_i : \mathbb{F}^n \to \mathbb{F}$ can be represented as a polynomial in $x_1, \ldots, x_n$, that is
$f_i \in \mathbb{F}[x_1, \ldots, x_n]$, the discrete network can be represented as

$$F = (f_1, \ldots, f_n) : \mathbb{F}^n \to \mathbb{F}^n$$

where each $f_i \in \mathbb{F}[x_1, \ldots, x_n]$. If any of the variables $x_1, \ldots, x_n$ take on values from a set that
cannot be directly identified with a finite field, then it is straightforward to embed the system
$F : X \to X$ into a system $\hat{F} : \mathbb{F}^n \to \mathbb{F}^n$, where $X \subset \mathbb{F}^n$, while preserving the attractor structure of
$F$; see [28].

Given a discrete network $F = (f_1, \ldots, f_n)$, a directed graph $W$ with $n$ nodes $x_1, \ldots, x_n$ is
associated to $F$. Thus, there is a directed edge in $W$ from $x_j$ to $x_i$ if $x_j$ appears in $f_i$, i.e. $x_j$ is in
the support of $f_i$, written $x_j \in \text{supp}(f_i)$. In the context of a molecular network model, this graph
represents the wiring diagram of the network.

The dynamics of discrete networks are given by the difference equation $x(t+1) = F(x(t))$; that is,
the dynamics is generated by iteration of $F$. More precisely, the dynamics of $F$ is represented
by the state space graph $S$, defined as the graph with vertices in $\mathbb{F}^n$ which has an edge from $x \in \mathbb{F}^n$
to $y \in \mathbb{F}^n$ if and only if $y = F(x)$. In this context, the problem of finding the states $x \in \mathbb{F}^n$ where
the system will get stabilized is of particular importance. These special points of the state space
are called attractors of a discrete network and these attractors may include steady states (fixed
points), where $F(x) = x$, or cycles, where $F^r(x) = x$ for some integer $r > 1$. Attractors in network
modeling might represent cell types [12] or cellular states such as apoptosis, proliferation, or cell
senescence [10, 26]. Identifying the attractors of a system is an important step towards the control
of that system.
3 Methods

Network interventions can be modeled through edge and node manipulations and can be achieved via therapeutic drugs that target a specific gene interaction [4, 2]. In [18, 21] we provided definitions for these actions in Boolean networks. These definitions are usually used for encoding the control parameters with the purpose of identifying control targets as shown in [21]. In this paper we will consider the deletion and constant expressions of edges and nodes in the multistate setting.

3.1 Edge Control in Multistate Networks

In the Boolean setting, the deletion of an edge was implemented by setting an input to zero so that the interaction of that input (represented by an edge) was being silenced. For the multistate case, the silencing of the interaction will be applied whenever the control variable is within a range of values of the possible discrete values.

Definition 3.1 (Edge Control). Consider the edge \(x_i \rightarrow x_j\) in the wiring diagram \(\mathcal{W}\). Let \(S_{i,j} \subset \mathbb{F}\) be a range of values within \(\mathbb{F}\) and \(Q_{S_{i,j}}(u)\) be the indicator function of \(S_{i,j}\). That is,

\[
Q_{S_{i,j}}(u) = \begin{cases} 
1 & \text{if } u \in S_{i,j}, \\
0 & \text{if } u \notin S_{i,j}.
\end{cases}
\]

For \(u \in S_{i,j}\), the control of the edge \(x_i \rightarrow x_j\) consists of manipulating the input variable \(x_i\) for \(f_j\) in the following way:

\[
\mathcal{F}_j(x, u) = f_j(x_{j1}, \ldots, (p-1)Q_{S_{i,j}}(u) + 1)x_{i1}, \ldots, x_{jm}).
\]

For each value of \(u \in \mathbb{F}\) we have the following control settings:

- For \(u \in S_{i,j}\), \(\mathcal{F}_j(x, u) = f_j(x_{j1}, \ldots, x_{i1} = 0, \ldots, x_{jm})\). That is, the control is active and the action represents the removal of the edge \(x_i \rightarrow x_j\).
- For \(u \notin S_{i,j}\), \(\mathcal{F}_j(x, u) = f_j(x_{j1}, \ldots, x_{i1}, \ldots, x_{jm})\). That is, the control is not active.

3.2 Node Control in Multistate Networks

Definition 3.2 (Node Control). Consider the node \(x_i\) in the wiring diagram \(\mathcal{W}\). Let \(S_i \subset \mathbb{F}\) be a range of values within \(\mathbb{F}\) and \(Q_{S_i}(u)\) be the indicator function of \(S_i\). That is,

\[
Q_{S_i}(u) = \begin{cases} 
1 & \text{if } u \in S_i, \\
0 & \text{if } u \notin S_i.
\end{cases}
\]

The function

\[
\mathcal{F}_i(x, u) := ((p-1)Q_{S_i}(u) + 1)f_i(x)
\]

encodes the control of the node \(x_i\) because for each possible value of \(u \in \mathbb{F}\) one has the following control settings:

- For \(u \in S_i\), \(\mathcal{F}_i(x, u) = 0\). This action represents the knock out of the node \(x_i\).
- For \(u \notin S_i\), \(\mathcal{F}_i(x, u) = f_i(x)\). That is, the control is not active.

3.3 Constant Expressions in Multistate Networks

We will also consider the constant expression of edges and nodes, which we define as follows.

Definition 3.3 (Constant edge expression). Consider the edge \(x_i \rightarrow x_j\) in the wiring diagram \(\mathcal{W}\) and \(a \in \mathbb{F}\). Let \(S_{i,j} \subset \mathbb{F}\) be a range of values within \(\mathbb{F}\) and \(Q_{S_{i,j}}(u)\) be the indicator function of \(S_{i,j}\). That is,

\[
Q_{S_{i,j}}(u) = \begin{cases} 
1 & \text{if } u \in S_{i,j}, \\
0 & \text{if } u \notin S_{i,j}.
\end{cases}
\]

For \(u \in S_{i,j}\), the control of the edge \(x_i \rightarrow x_j\) consists of manipulating the input variable \(x_i\) for \(f_j\) in the following way:

\[
\mathcal{F}_j(x, u) = f_j(x_{j1}, \ldots, (p-1)Q_{S_{i,j}}(u) + 1)x_{i1} + aQ_{S_{i,j}}(u), \ldots, x_{jm}).
\]

For each value of \(u \in \mathbb{F}\) we have the following settings:
For $u \in S_{i,j}$, $F_j(x,u) = f_j(x_{j_1},\ldots,x_{j_m})$. That is, the control is active and the action represents the constant expression (to a) of the edge $x_i \rightarrow x_j$.

For $u \notin S_{i,j}$, $F_j(x,u) = f_j(x_{j_1},\ldots,x_{i-1},x_{i+1},\ldots,x_{j_m})$. That is, the control is not active.

**Notation 3.4.** We will denote an edge deletion or a constant expression by indicating the constant that the input is being set to by $x_i \xrightarrow{u} x_j$.

Similarly, we encode constant expression of nodes.

**Definition 3.5 (Constant node expression).** Consider the node $x_i$ in the wiring diagram $W$ and $a \in \mathbb{F}$. Let $S_i \subset \mathbb{F}$ be a range of values within $\mathbb{F}$ and $Q_{S_i}(u)$ be the indicator function of $S_i$. That is,

$$Q_{S_i}(u) = \begin{cases} 1 & \text{if } u \in S_i, \\ 0 & \text{if } u \notin S_i. \end{cases}$$

The function

$$F_i(x,u) := ((p-1)Q_{S_i}(u) + 1)f_i(x) + aQ_{S_i}(u)$$

encodes the constant expression of the node $x_i$ to $a$. Note that,

- For $u \in S_i$, $F_i(x,u) = a$. This action represents the constant expression to $a$ of the node $x_i$.
- For $u \notin S_i$, $F_i(x,u) = f_i(x)$. That is, the control is not active.

**4 Results**

In this section we present a definition of $k$-canalizing functions for the multistate case and then we characterize these functions in terms of layers of canalizations. Subsequently, we use this canalizing layers representation to derive an upper bound for the number of changes in the state space of a discrete system upon an edge deletion in the wiring diagram.

**4.1 Multistate $k$-Canalizing Functions**

**Definition 4.1.** The function $f : \mathbb{F}^n \rightarrow \mathbb{F}$ is a $k$-canalizing function in the variable order $x_{\sigma(1)},\ldots,x_{\sigma(k)}$ with canalizing input sets $S_1,\ldots,S_k \subset \mathbb{F}$ and canalizing output values $b_1,\ldots,b_k \in \mathbb{F}$ if it can be represented in the form

$$f(x_1,\ldots,x_n) = \begin{cases} b_1 & \text{if } x_{\sigma(1)} \in S_1, \\ b_2 & \text{if } x_{\sigma(1)} \notin S_1, x_{\sigma(2)} \in S_2, \\ \vdots \\ b_k & \text{if } x_{\sigma(1)} \notin S_1,\ldots, x_{\sigma(k)} \in S_k, \\ g & \text{if } x_{\sigma(1)} \notin S_1,\ldots,x_{\sigma(k)} \notin S_k. \end{cases} \tag{3}$$

where $g = g(x_{\sigma(k+1)},\ldots,x_{\sigma(n)})$ is a multistate function on $n-k$ variables. When $g$ is not a canalizing function, the integer $k$ is the canalizing depth of $f$. If $g$ is not a constant function, then $g$ is called the core function of $f$ and is denoted by $P_C$.

**Remark 4.2.** Note that in Definition 4.1 of $k$-canalizing functions we require that the function $g$ be unique when all the canalizing variables are not in their corresponding canalizing input sets. For instance, a function could be canalizing but not 1-canalizing, see Example 4.3.

**Example 4.3.** Let $\mathbb{F} = \{0,1,2\}$ and $n = 2$. Consider the function

$$f(x_1,x_2) = 1 + 2x_1^2 + 2x_2 + 2x_1^2x_2 + 2x_1^2.$$  

For this function $x_2$ is canalizing (with $S_1 = \{2\}$) because $f(x_1,2) = 1$. However, $f$ is not a 1-canalizing function because $f(x_1,0) = 1 + 2x_1^2 \neq 2 + 2x_1^2 = f(x_1,1)$. Thus, even though $x_2$ is canalizing for $f$, the function $f$ has zero layers. Here $P_C = f$.  


4.2 Layers of canalization in multistate networks

In Theorem 4.4 we provide a polynomial normal description of discrete functions. Basically, this theorem gives a partition of the inputs of the function into canalizing and non-canalizing variables and, within the canalizing ones, we categorize the input variables into layers of canalization. This theorem is a generalization of a theorem in [9] from Boolean to the multistate case.

Let \( S \subset \mathbb{F} \) be a range of values within \( \mathbb{F} \) and \( \tilde{Q}_S(u) \) be the indicator function of the complement of \( S \). That is,
\[
\tilde{Q}_S(x) = \begin{cases} 
1 & \text{if } x \notin S, \\
0 & \text{if } x \in S.
\end{cases}
\]

**Theorem 4.4.** Every multistate function can be uniquely written as
\[
f(x_1, \ldots, x_n) = M_1(M_2(\ldots (M_{r-1}(M_r P_C + B_r) + B_{r-1}) \ldots) + B_2) + B_1, \tag{4}
\]
where \( M_i = \prod_{j=1}^{k_i} \tilde{Q}_{S_{i,j}}, \) \( k = k_1 + \cdots + k_r \) is the canalizing depth, \( P_C \) is a polynomial that has no canalizing variables, \( B_1, B_2, \ldots, B_r \in \mathbb{F} \), and \( B_r \neq 0 \). Each variable \( x_i \) appears in exactly one of the \( M_1, M_2, \ldots, M_r, P_C \).

**Proof.** If \( f(x_1, \ldots, x_n) \) is non-canalizing, then \( P_C = f \). If \( f(x_1, \ldots, x_n) \) is canalizing, then we proceed by induction. For \( n = 1 \), if \( f \) is canalizing in \( x_i \) but not 1-canalizing in \( x_i \), then we set \( P_C = f \). If \( f \) is 1-canalizing in \( x_i \), then it can be written as \( f = Q_{S_i}(x_i) + B_1 \). Then \( f \) has the form of Equation 4 by setting \( M_1 = \tilde{Q}_{S_i}(x_i) \) and \( P_C = 1 \). For \( n = 2 \), if \( f(x_i, x_j) \) is not 1-canalizing on any of its variables, then we set \( P_C = f \). If \( f \) is 1-canalizing on \( x_i \), then \( f \) can be written as \( f(x_i, x_j) = M_1(x_i) g(x_j) + B_1 \) for some \( g(x_j) \). Then \( f \) has the form of Equation 4 by setting \( P_C = g \). Now assume that Equation 4 is true for any canalizing function that is essential in at most \( n-1 \) variables (that is, for all functions that depend in at most \( n-1 \) variables). Let \( f \) be a function that is essential in \( n \) variables. If \( f \) is not 1-canalizing on any of its variables, then we set \( P_C = f \). If \( f \) is 1-canalizing in \( x_1, \ldots, x_{k_i} \), then \( f = M_1 g + B_1 \), where \( M_1 \) is the product of indicator functions of the complements of \( S_1, \ldots, S_{k_i} \) and \( g \) has \( n - k_i \) variables. If \( g \) has no canalizing variables, then \( f \) has the form of Equation 4 with \( P_C = g \). If \( g \) is canalizing, then by the inductive hypothesis \( g \) can be written as
\[
g = M_2(\ldots (M_{r-1}(M_r P_C + B_r) + B_{r-1}) \ldots) + B_2.
\]
Thus, \( f \) has the form of Equation 4. \( \square \)

**Remark 4.5.**
For a multistate nested canalizing function, the formula in Equation 4 reduces to
\[
f(x_1, \ldots, x_n) = M_1(M_2(\ldots (M_{r-1}(B_{r+1} M_r + B_r) + B_{r-1}) \ldots) + B_2) + B_1, \tag{5}
\]
as was shown in [11].

In the following example we describe a 2-canalizing function with noncanalizing variables.

**Example 4.6.** Let \( \mathbb{F} = \{0, 1, 2\} \) and \( n = 4 \). Consider the function
\[
f(x_1, x_2, x_3, x_4) = 1 + x_1^2 + x_2^2 + 2x_1^2x_2 + x_1^2x_2^2 + 2x_1^2x_2x_3 + 2x_1^2x_2^2x_3 + x_1^2x_2x_4 + 2x_1^2x_2^2x_4.
\]
The function \( f \) can be written as in Equation 4 as
\[
f = M_1(M_2(P_C + 1) + 1) + 1,
\]
where \( M_1 = \tilde{Q}_{S_1}(x_1) = x_1^2 \), \( S_1 = \{0\} \), \( M_2 = \tilde{Q}_{S_2}(x_2) = x_2 + 2x_2^2 \), \( S_2 = \{0, 1\} \), and \( P_C = x_3 + x_4 \). Thus \( f \) has two layers and two noncanalizing variables. Note that \( f \) can also be written as in Equation 3 as
\[
f(x_1, x_2, x_3, x_4) = \begin{cases} 
1 & \text{if } x_1 \in S_1 = \{0\}, \\
2 & \text{if } x_1 \notin S_1, x_2 \in S_2 = \{0, 1\}, \\
P_C & \text{if } x_1 \notin S_1, x_2 \notin S_2.
\end{cases}
\]
4.3 Upper bounds

Using the polynomial normal form of multistate functions in Theorem 4.4, we derive a set of formulas for counting the maximum number of transitions that will change in the state space upon an edge deletion in the wiring diagram. The formulas presented here are generalizations from the Boolean case to the multistate setting of the formulas we presented in [18].

**Theorem 4.7.** Let $F = (f_1, \ldots, f_n) : \mathbb{F}^n \to \mathbb{F}^n$ be a multistate network and $f_i$ be a $k$-canalizing function where

$$f_i(x_1, \ldots, x_n) = M^i_1(M^i_2(\ldots(M^i_{\ell-1}(M^i_{\ell}P_C + B_r) + B_{\ell-1})\ldots) + B_2) + B_1,$$

where $M^i_\ell = \prod_{j=1}^{k_i} \bar{Q}_{j,t}$, $k = k_1 + \cdots + k_\ell$ is the canalizing depth, $P_C$ is a polynomial with no canalizing variables, $B_1, B_2, \ldots, B_\ell \in \mathbb{F}$, and $B_\ell \neq 0$. Each variable $x_i$ appears in exactly one of $M^i_1, M^i_2, \ldots, M^i_{\ell}, P_C$.

Let $x_s$ be in the $d^{th}$ layer, $d \leq r$. The maximum number of transitions in the state space that will change upon deletion of $x_s \to x_t$ is given by

$$
\frac{p^n}{p^{d+e} - \ell_d} \cdot \left( \prod_{i=1}^{d-1} \prod_{j=1}^{k_i} (p - |S_{j,t}|) \right) + \frac{1}{p^{d-1}} \left( \prod_{j=1}^{k_d} (p - |S_{j,t}|) \right) \frac{1}{p} (p - R) =

\frac{p^{n-d}}{p^{d+e} - \ell_d} \cdot \left( \prod_{i=1}^{d-1} \prod_{j=1}^{k_i} (p - |S_{j,t}|) \right) \left( \prod_{j=1}^{k_d} (p - |S_{j,t}|) \right) (p - R),
$$

where

$$R = \begin{cases} 
|S_{s,t}| & \text{if } 0 \in S_{s,t} \\
 p - |S_{s,t}| & \text{if } 0 \notin S_{s,t}.
\end{cases}$$

**Proof.** The term $\frac{p^n}{p^{d+e} - \ell_d}$ in Equation 7 accounts for the number of vectors containing the input vectors in the transition table of $f_i$. Then, we count the inputs where the variables in the first $d$ layers that are not $x_s$ do not take on values from their canalizing sets. The term inside the first set of parentheses of Equation 7 accounts for the potential number of changes in the previous layers to the layer containing $x_s$. The term inside the second set of parentheses of Equation 7 accounts for the potential number of changes in the layer containing $x_s$ but that are caused by the other variables in that layer. The last term in Equation 7 accounts for the potential number of changes caused by $x_s$. For the last term, notice that deleting $x_s \to x_t$ results in setting $x_s = 0$ in $f_i$. If 0 is in the canalizing set of $x_s$, $S_{s,t}$, then the rest of the values in $S_{s,t}$ will yield the same output as 0. Since $|S_{s,t}|/p$ of the input values in the transition table of $f_i$ contain a canalizing value for $x_s$, it is the remaining $\frac{p - |S_{s,t}|}{p}$ of the table that can potentially change as a result of the edge deletion. On the other hand, if 0 \notin S_{s,t}, then it is the inputs not in $S_{s,t}$ that have the potential to change the output as a result of deleting $x_s \to x_t$ which constitutes $1/p$ of the transition table, with $\frac{p-1}{p}$ of the table that can potentially change as a result of the edge deletion.

**Remark 4.8.**

1. The bound in Equation 7 is sharp.

2. When $p = 2$, the formula in Equation 7 reduces to $2^{n-d} - \ell_d$.

3. If instead of edge deletion, we consider constant expression to $a \in \mathbb{F}$ (see Section 3.3) of $x_s \to x_t$, then the formula in Equation 7 remains the same except for the $R$ term which will become

$$R = \begin{cases} 
|S_{s,t}| & \text{if } a \in S_{s,t} \\
 p - |S_{s,t}| & \text{if } a \notin S_{s,t}.
\end{cases}$$
Proposition 4.9. Let $F = (f_1, \ldots, f_n): F^n \to F^n$ be a multistate network and $f_t$ be a $k$-canalizing function where

$$f_t(x_1, \ldots, x_m) = M_t^1(M_2^1(\ldots(M_{r-1}^1(M_r^1P_C + B_r) + B_{r-1})\ldots) + B_2) + B_1$$

(8)

where $M_t^1 = \prod_{j=1}^k \tilde{Q}_{j,t}$, $k = k_1 + \cdots + k_r$ is the canalizing depth, $P_C$ is a polynomial with no canalizing variables, $B_1, B_2, \ldots, B_r \in F$, and $B_r \neq 0$. Each variable $x_i$ appears in exactly one of $M_1^t, M_2^t, \ldots, M_r^t, P_C$.

Remark 4.10. 1. This upper bound is sharp.

2. When $p = 2$, the expression reduces to $2^{n-d-1}$.

3. If $f$ has no canalizing variables, then the formula in Equation 9 reduces to $p^{n-1}(p-1)$.

Proposition 4.11. Let $F = (f_1, \ldots, f_n): F^n \to F^n$ be a multistate network and $f_t$ be a $k$-canalizing function where

$$f_t(x_1, \ldots, x_m) = M_t^1(M_2^1(\ldots(M_{r-1}^1(M_r^1P_C + B_r) + B_{r-1})\ldots) + B_2) + B_1$$

(10)

where $M_t^1 = \prod_{j=1}^k \tilde{Q}_{j,t}$, $k = k_1 + \cdots + k_r$ is the canalizing depth, $P_C$ is a polynomial with no canalizing variables, $B_1, B_2, \ldots, B_r \in F$, and $B_r \neq 0$. Each variable $x_i$ appears in exactly one of $M_1^t, M_2^t, \ldots, M_r^t, P_C$.

If $P_C$ is canalizing but not 1-canalizing with canalizing variable $x_s$ and input set $S_{s,t}$, then there are two cases to consider

1. The deletion of $x_s \to x_t$ will result in up to

$$= p^{n-1} \prod_{i=1}^r (p - |S_{i,t}|) (p - R)$$

transitions, where

$$R = \begin{cases} |S_{s,t}| & \text{if } 0 \in S_{s,t} \\ 1 & \text{if } 0 \notin S_{s,t}. \end{cases}$$

2. Let $x_a \in \text{supp}(P_C)$ that is not canalizing. Then the maximum number of transitions in the state space that will change upon deletion of $x_a \to x_t$ is

$$= p^{n-1} \prod_{i=1}^r (p - |S_{i,t}|) (p - |S_{s,t}|)(p-1).$$

(12)

Remark 4.12. 1. This upper bound is sharp, Table 1.

2. If $P_C$ has more than one canalizing variables (but it is still not 1-canalizing), then the formula in Equation 12 becomes

$$= p^{n-1} \prod_{i=1}^r (p - |S_{i,t}|) \prod_{i=1}^c (p - |S_{s,t}|)(p-1),$$

(13)

where each $x_{s_i}$ is a canalizing variable and $c$ is the number of canalizing variables of $P_C$. 
5 Applications

To provide further insight into the results presented above and to illustrate the use of the formulas we apply our methods to two published model.

Example 5.1 (Tissue model). Here we consider an example of a multistate network with six nodes where each variable can take on three states. This model was adapted from the model presented in [5] and its wiring diagram is given in Figure 1.

\[
\begin{align*}
\text{Injury} & \in \{0, 1\} \\
\text{ROS} & \in \{0, 1\} \\
M2 & \in \{0, 1\} \\
DAMPs & \in \{0, 1, 2\} \\
M1 & \in \{0, 1, 2\} \\
CCL2 & \in \{0, 1, 2\}
\end{align*}
\]  

(14)

The original model in [5] is a mixed-state model where half of the variables are Boolean and the other half are ternary, see Equation 14. Here, we used the method in [28] to have all the variables with same number of possible states. In this case, we used the finite field with three elements: \( \mathbb{F} = \{0, 1, 2\} \). Here we rename the variables as follows,

\[
\begin{align*}
x_1 &= \text{Injury} \\
x_2 &= \text{ROS} \\
x_3 &= M2 \\
x_4 &= DAMPs \\
x_5 &= M1 \\
x_6 &= CCL2.
\end{align*}
\]

The polynomial functions are listed below:

- \( f_1 = x_1^2 - x_1^2x_3^2 \),
- \( f_2 = x_4 - x_4^2 \),
- \( f_3 = 2x_5 - x_5^3 \),
- \( f_4 = x_1^2 + x_2^2 + 2x_2^2x_3^2 + 2x_2^2x_3^2 + x_1^2x_3^4 + x_2^2x_3 + x_2^2x_4 + x_1^2x_3^4 + 2x_1^2x_3^4x_4 + 2x_2^2x_3^2x_4 + 2x_1^2x_3^2x_4 + 2x_1^2x_3^2x_4 + x_1^2x_3^2x_4 + x_2^2x_3^2x_4 + 2x_2^2x_3^2x_4 + 2x_1^2x_3^2x_4, \)

8
Table 1: Upper bound versus the actual number of changes for each edge deletion in the wiring diagram of Figure 1. Rows in green indicate an accurate estimate.

| Deleted edge | Formula used | Canalizing structure | Upper bound | Exact number |
|--------------|--------------|----------------------|-------------|--------------|
| 1 \(\rightarrow\) 1 | Eq. 7 | \(r = 1, \ell_1 = 2\) | 162 | 162 |
| 3 \(\rightarrow\) 1 | Eq. 7 | \(r = 1, \ell_1 = 2\) | 324 | 324 |
| 4 \(\rightarrow\) 2 | Eq. 7 | \(r = 1, \ell_1 = 1\) | 243 | 243 |
| 5 \(\rightarrow\) 3 | Eq. 7 | \(r = 1, \ell_1 = 1\) | 243 | 243 |
| 1 \(\rightarrow\) 4 | Eq. 9 | \(r = 0, \ell_1 = 0\) | 486 | 342 |
| 2 \(\rightarrow\) 4 | Eq. 9 | \(r = 0, \ell_1 = 0\) | 486 | 342 |
| 3 \(\rightarrow\) 4 | Eq. 9 | \(r = 0, \ell_1 = 0\) | 486 | 432 |
| 4 \(\rightarrow\) 4 | Eq. 9 | \(r = 0, \ell_1 = 0\) | 486 | 36 |
| 6 \(\rightarrow\) 5 | Eq. 9 | \(r = 0, \ell_1 = 0\) | 486 | 486 |
| 3 \(\rightarrow\) 6 | Eq. 12 | \(r = 0, \ell_1 = 0\) | 162 | 162 |
| 4 \(\rightarrow\) 6 | Eq. 11 | \(r = 0, \ell_1 = 0\) | 486 | 324 |

Now we write the polynomials in the form of Equation 4. We’ll use superscripts for the layers to indicate the corresponding function.

- \(f_5 = x_6\).
- \(f_6 = x_4 + x_3^2 x_4 + x_2^3 x_4^2\).

In Table 1 we used the formulas in Equations 7-9 to compute the upper bounds for the number of changes upon deletion of each edge and compared with the actual number of changes. For this model, the number of states is \(p = 3\) and the number of nodes is \(n = 6\). For instance, the function \(f_1\) has a single layer with \(\ell_1 = 2\). Thus, for the deletion of the edge \(x_1 \rightarrow x_1\) the formula in Equation 7 gives us \(p^{\ell_1-1}(p - |S_{1,1}|)(|S_{1,1}|) = 3^2(2) = 162\). Notice that the \(R\) term in Equation 7 is equal to \(|S_{1,1}|\) because \(0 \notin S_{1,1}\). Similarly, for deleting the edge \(x_3 \rightarrow x_1\), we have \(p^{\ell_1-1}(p - |S_{1,1}|)(p - |S_{3,1}|) = 3^2(2)(2) = 324\).

Example 5.2 (Phage-\(\lambda\) network). Here we consider an example of a multistate network with four nodes where each variable can take on five states. This model was adapted from the model presented in [27]. This model considers the role of the genes CI, CRO, CII, and N. The original model in [27] is a mixed-state model where two of the variables are Boolean and the other two are multistate, see Equation 15. Here, we used the method in [28] to have all the variables with same number of possible states. In this case, we used the finite field with five elements: \(F = \{0, 1, 2, 3, 4\}\).

\[
\begin{align*}
CI & \in \{0, 1, 2\} \\
CRO & \in \{0, 1, 2, 3\} \\
CII & \in \{0, 1\} \\
N & \in \{0, 1\}
\end{align*}
\]

\[
\begin{align*}
x_1 & = CI \\
x_2 & = CRO \\
x_3 & = CII \\
x_4 & = N
\end{align*}
\]

The polynomial functions are listed below.
In this paper we present practical methods for quantifying the unplanned or unwanted changes in Equation 7 is equal to
\[ p^{n-\ell_1}(p - |S_{3,3}|)(p - |S_{4,3}|)(|S_{1,3}|) = 51(3)(4)(3) = 180. \]
Notice that the R term in Equation 7 is equal to |S_{1,3}| because 0 \notin S_{1,3}. Similarly, for the deletion of the edge x_2 \rightarrow x_3 the formula in Equation 7 gives 
\[ p^{n-\ell_1}(p - |S_{3,3}|)(p - |S_{4,3}|)(|S_{2,3}|) = 51(2)(4)(2) = 80. \]
Notice that the R term in Equation 7 is equal to |S_{2,3}| because 0 \notin S_{2,3}.

6 Conclusions

In this paper we present practical methods for quantifying the unplanned or unwanted changes that come along with an application of a control action in the network, which we called the side effects of the intervention. We emphasized that, while there are several methods for identifying control targets in discrete networks, there have been very few studies focusing on quantifying the changes upon the applications of controllers. This paper contributes methods for measuring...
Table 2: Upper bound versus the actual number of changes for each edge deletion in the wiring diagram of Figure 2. Rows in green indicate an accurate estimate.

| Deleted edge | Formula used | Canalizing structure | Upper bound | Exact number |
|--------------|--------------|----------------------|-------------|-------------|
| 1 $\rightarrow$ 1 | Eq. 9 | $r = 0, \ell_1 = 0$ | 500 | 480 |
| 2 $\rightarrow$ 1 | Eq. 9 | $r = 0, \ell_1 = 0$ | 500 | 100 |
| 3 $\rightarrow$ 1 | Eq. 9 | $r = 0, \ell_1 = 0$ | 500 | 400 |
| 1 $\rightarrow$ 2 | Eq. 12 | $r = 0, \ell_1 = 0$ | 300 | 225 |
| 2 $\rightarrow$ 2 | Eq. 11 | $r = 0, \ell_1 = 0$ | 500 | 425 |
| 1 $\rightarrow$ 3 | Eq. 7 | $r = 1, \ell_1 = 3$ | 180 | 180 |
| 2 $\rightarrow$ 3 | Eq. 7 | $r = 1, \ell_1 = 3$ | 80 | 80 |
| 4 $\rightarrow$ 3 | Eq. 7 | $r = 1, \ell_1 = 3$ | 120 | 120 |
| 1 $\rightarrow$ 4 | Eq. 7 | $r = 1, \ell_1 = 2$ | 200 | 200 |
| 2 $\rightarrow$ 4 | Eq. 7 | $r = 1, \ell_1 = 2$ | 75 | 75 |

the number of changed transitions in the state space upon the application of an edge control in multistate networks. The approach is based on a polynomial normal form description of discrete functions that provides a way to categorize the inputs of the function and therefore to quantify their impact on the dynamics of the network. We applied our methods to two multistate models and verified that in most cases the estimates provided by our methods are accurate. Future work on characterizing the noncanalizing variables could lead to more accurate estimates for the number of changed transitions.

References

[1] Réka Albert and Hans G Othmer. The topology of the regulatory interactions predicts the expression pattern of the segment polarity genes in drosophila melanogaster. *J Theor Biol*, 223(1):1–18, Jul 2003.

[2] Colin Campbell and Réka Albert. Edgetic perturbations to eliminate fixed-point attractors in boolean regulatory networks. *Chaos*, 29(2):023130, Feb 2019.

[3] Julia Chifman, Seda Arat, Zhiyong Deng, Erica Lemler, James C Pino, Leonard A Harris, Michael A Kochen, Carlos F Lopez, Steven A Akman, Frank M Torti, Suzy V Torti, and Reinhard Laubenbacher. Activated oncogenic pathway modifies iron network in breast epithelial cells: A dynamic modeling perspective. *PLoS Comput Biol*, 13(2):e1005352, 02 2017.

[4] Minsoo Choi, Jue Shi, Sung Hoon Jung, Xi Chen, and Kwang-Hyun Cho. Attractor landscape analysis reveals feedback loops in the p53 network that control the cellular response to dna damage. *Sci. Signal.*, 5(251):ra83, 2012.

[5] Elena Dimitrova, Leslie A. Caromile, Reinhard Laubenbacher, and Linda H. Shapiro. The innate immune response to ischemic injury: a multiscale modeling perspective. *BMC Systems Biology*, 12(1), Apr 2018.

[6] Carlos Espinosa-Soto, Pablo Padilla-Longoria, and Elena R Alvarez-Buylla. A gene regulatory network model for cell-fate determination during arabidopsis thaliana flower development that is robust and recovers experimental gene expression profiles. *Plant Cell*, 16(11):2923–39, Nov 2004.

[7] Alexander J Gates and Luis M Rocha. Control of complex networks requires both structure and dynamics. *Scientific reports*, 6:24456, 2016.

[8] S. Ha, E.S. Dimitrova, S. Hoops, D. Altarawy, M. Ansariola, D. Deb, J. Glazebrook, R. Hillmer, H. Hosamadine, E. Katagiri, J. McDowell, M. McGraw, J. Setubal, B.M. Tyler, and R. Laubenbacher. Plantsimlab - a modeling and simulation web tool for plant biologists. *BMC Bioinformatics*, Accepted 2019.

[9] Qijun He and Matthew Macauley. Stratification and enumeration of boolean functions by canalizing depth. *Physica D: Nonlinear Phenomena*, 314:1–8, 2016.
S Huang. Gene expression profiling, genetic networks, and cellular states: an integrating concept for tumorigenesis and drug discovery. *J Mol Med (Berl)*, 77(6):469–80, Jun 1999.

Claus Kadelka, Yuan Li, Jack Kuipers, John O Adeyeye, and Reinhard Laubenbacher. Multi-state nested canalizing functions and their networks. *Theoretical Computer Science*, 675:1–14, 2017.

S A Kauffman. Metabolic stability and epigenesis in randomly constructed genetic nets. *J Theor Biol*, 22(3):437–67, Mar 1969.

Stuart Kauffman, Carsten Peterson, Björn Samuelsson, and Carl Troein. Random boolean network models and the yeast transcriptional network. *Proceedings of the National Academy of Sciences*, 100(25):14796–14799, 2003.

Stuart Kauffman, Carsten Peterson, Björn Samuelsson, and Carl Troein. Genetic networks with canalyzing boolean rules are always stable. *Proc Natl Acad Sci U S A*, 101(49):17102–7, Dec 2004.

Steffen Klamt, Julio Saez-Rodriguez, Jonathan A Lindquist, Luca Simeoni, and Ernst D Gilles. A methodology for the structural and functional analysis of signaling and regulatory networks. *BMC Bioinformatics*, 7:56, Feb 2006.

Fangting Li, Tao Long, Ying Lu, Qi Ouyang, and Chao Tang. The yeast cell-cycle network is robustly designed. *Proc Natl Acad Sci U S A*, 101(14):4781–6, Apr 2004.

Rui Li, Meng Yang, and Tianguang Chu. Controllability and observability of boolean networks arising from biology. *Chaos*, 25(2):023104, Feb 2015.

David Murrugarra and Elena S Dimitrova. Molecular network control through boolean canalization. *EURASIP J Bioinform Syst Biol*, 2015(1):9, Dec 2015.

David Murrugarra and Reinhard Laubenbacher. Regulatory patterns in molecular interaction networks. *Journal of Theoretical Biology*, 288(0):66–72, 11 2011.

David Murrugarra and Reinhard Laubenbacher. The number of multistate nested canalizing functions. *Physica D: Nonlinear Phenomena*, 241(10):929–938, 5 2012.

David Murrugarra, Alan Veliz-Cuba, Boris Aguilar, and Reinhard Laubenbacher. Identification of control targets in boolean molecular network models via computational algebra. *BMC Syst Biol*, 10(1):94, Sep 2016.

A Naldi, D Berenguier, A Fauré, F Lopez, D Thieffry, and C Chaouiya. Logical modelling of regulatory networks with ginsim 2.3. *Biosystems*, 97(2):134–9, Aug 2009.

Arnaud Poret and Jean-Pierre Boissel. An in silico target identification using boolean network attractors: Avoiding pathological phenotypes. *Comptes rendus biologies*, 337(12):661—678, December 2014.

Yushan Qiu, Takeyuki Tamura, Wai-Ki Ching, and Tatsuya Akutsu. On control of singleton attractors in multiple boolean networks: integer programming-based method. *BMC Syst Biol*, 8 Suppl 1:S7, 2014.

E. Remy, S. Rebouissou, C. Chaouiya, A. Zinovyev, F. Radvanyi, and L. Calzone. A modeling approach to explain mutually exclusive and co-occurring genetic alterations in bladder tumorigenesis. *Cancer Research*, 75(19):4042–4052, Aug 2015.

Ilya Shmulevich and Edward R. Dougherty. *Probabilistic Boolean Networks - The Modeling and Control of Gene Regulatory Networks*. SIAM, 2010.

D Thieffry and R Thomas. Dynamical behaviour of biological regulatory networks–ii. immunity control in bacteriophage lambda. *Bull Math Biol*, 57(2):277–97, Mar 1995.

Alan Veliz-Cuba, Abdul Salam Jarrah, and Reinhard Laubenbacher. Polynomial algebra of discrete models in systems biology. *Bioinformatics*, 26(13):1637–43, Jul 2010.
[29] Paola Vera-Licona, Eric Bonnet, Emmanuel Barillot, and Andrei Zinovyev. Ocsana: optimal combinations of interventions from network analysis. *Bioinformatics*, 29(12):1571–3, Jun 2013.

[30] C. H Waddington. *The strategy of the genes: a discussion of some aspects of theoretical biology*. Allen & Unwin, London, 1957.

[31] Jorge G T Zañudo and Réka Albert. Cell fate reprogramming by control of intracellular network dynamics. *PLoS Comput Biol*, 11(4):e1004193, Apr 2015.

[32] Jorge Gómez Tejeda Zañudo, Maurizio Scaltriti, and Réka Albert. A network modeling approach to elucidate drug resistance mechanisms and predict combinatorial drug treatments in breast cancer. *Cancer convergence*, 1(1):5, 2017.

[33] Jorge Gomez Tejeda Zañudo, Gang Yang, and Réka Albert. Structure-based control of complex networks with nonlinear dynamics. *Proc Natl Acad Sci U S A*, 114(28):7234–7239, 07 2017.