DOMAIN CONTROL OF NONLINEAR NETWORKED SYSTEMS AND APPLICATIONS TO COMPLEX DISEASE NETWORKS

SUOQIN JIN

School of Mathematics and Statistics, Wuhan University
School of Mathematics and Statistics, Central China Normal University
Wuhan 430072, China

FANG-XIANG WU

Department of Mechanical Engineering, University of Saskatchewan
Saskatoon, S7N 5A9, Canada

XIUFEN ZOU*

School of Mathematics and Statistics, Wuhan University
Computational Science Hubei Key Laboratory, Wuhan University
Wuhan 430072, China

(Communicated by the associate editor name)

ABSTRACT. The control of complex nonlinear dynamical networks is an ongoing challenge in diverse contexts ranging from biology to social sciences. To explore a practical framework for controlling nonlinear dynamical networks based on meaningful physical and experimental considerations, we propose a new concept of the domain control for nonlinear dynamical networks, i.e., the control of a nonlinear network in transition from the domain of attraction of an undesired state (attractor) to the domain of attraction of a desired state. We theoretically prove the existence of a domain control. In particular, we offer an approach for identifying the driver nodes that need to be controlled and design a general form of control functions for realizing domain controllability. In addition, we demonstrate the effectiveness of our theory and approaches in three realistic disease-related networks: the epithelial-mesenchymal transition (EMT) core network, the T helper (Th) differentiation cellular network and the cancer network. Moreover, we reveal certain genes that are critical to phenotype transitions of these systems. Therefore, the approach described here not only offers a practical control scheme for nonlinear dynamical networks but also helps the development of new strategies for the prevention and treatment of complex diseases.

1. **Introduction.** The control of complex networks has attracted enormous attention in recent years because of its wide applications in a variety of natural, social, economic, and man-made systems [36, 42, 11, 46]. Despite significant advances in exploring the controllability of complex networks [36, 65, 19, 62, 45, 54, 64, 35, 10, 55], certain limitations remain. For example, most of the methods that have been developed are for the control of networks hypothetically governed by linear dynamics.

A previous study showed that according to the control theory for linear systems, it is necessary to directly control approximately 80% of the nodes to achieve complete controllability in gene regulatory networks. However, such controls are usually infeasible and unnecessary because biological and medical experimental lines of evidence show that one or few driver nodes should be enough, particularly when designing drug targets in the biological and medical fields [36, 41, 62]. Therefore, the minimum number of driver nodes in a linear dynamical system is often overestimated.

In the real world, dynamical processes on complex networks are generally nonlinear. Therefore, the control of complex networked systems using nonlinear dynamics must be considered [65, 7, 35, 10, 28, 57, 66]. Although the controllability of nonlinear problems can be formulated using Lie brackets, such a control strategy is difficult to implement if the system is large and complex [47]. Although some progresses have been made in the development of specific control methods, such as chaos control [9], pinning control [49] and optimal control [52], these methods are limited to low-dimensional systems or certain particular situations. A breakthrough in the newly developed control approach is a realistic control framework, which can be used to direct a network to a desirable state by identifying compensatory perturbations [11]. A numerical and group representational framework has also been presented to show the effect of symmetries on the controllability of nonlinear networks [61]. More recently, a highly scalable algorithm has been presented to rationally identify parameter interventions that can control the response to noise in complex nonlinear networks by modulating the heights of the barriers separating different stable steady states [60]. In addition, the concept of the attractor network, in which the nodes are the distinct attractors of the nonlinear system, has been recently introduced to investigate the transitions between different attractors by applying parameter perturbations [56]. However, a general theoretical and computational framework for the actual control of complex networks using nonlinear dynamics has not been developed yet. A general mathematical framework for controlling complex nonlinear networks is difficult to define due to the complexity of the network structure as well as the extremely diverse nonlinear dynamical behaviors of systems [30, 51]. Therefore, the problem of controlling complex nonlinear dynamical networks remains an outstanding challenge.

Indeed, many realistic networks, particularly networks representing pathogenesis of complex diseases, may not require the detailed control of all state variables during the transition from any initial state to any desired final state, which prompts us to explore the development of a practical framework for controlling nonlinear dynamical networks based on meaningful physical and experimental considerations. A common feature of nonlinear dynamical systems is the emergence of a number of distinct attractors [29, 51, 24, 25]. The performance and functions of the system are often determined by a particular attractor. Because multiple types of stable steady states (attractors) can occur, as has been observed in complex disease networks, such as cancer, tumor, and diabetes II, there are certain states that may correspond to healthy (desirable) states under normal conditions and a few other states that correspond to disease states or undesirable states [1, 5, 21, 12, 34, 33]. Potential curative interventions are those that can steer the system from a disease or pre-disease state (i.e., domain of attraction of an undesirable state) to the domain of attraction of the normal state [11, 1, 3].

Inspired by these biological problems, we propose a new concept for domain control of a nonlinear dynamical network, i.e., a method that steers a nonlinear

network from one domain of attraction to another domain of attraction. Our goal is to develop a new mathematical theory and computational approach to realize domain controllability. We then use several realistic networks to test our theory and approaches. Notably, we reveal newly meaningful strategies for controlling the fate of living cells. These results are an important step toward the realistic control of complex networked systems with nonlinear dynamics.

2. **Problem formulation.** The dynamics of complex networks can often be represented by a set of nonlinear ordinary differential equations (ODEs). Without loss of generality, we consider an n-node network whose n-dimensional dynamical state $y(t)$ is described as

$$\dot{y}(t) = F(y(t)), \tag{1}$$

where $F(y(t)) \in \mathbb{R}^n$ is a nonlinear function, which is continuously differentiable with respect to y. The corresponding nonlinear control system can be specified as

$$\dot{x}(t) = f(x(t), u(t)), \tag{2}$$

where $x(t) \in \mathbb{R}^n$ is the state vector, $u(t) \in \Omega \subseteq \mathbb{R}^m$ is the control signal, which is continuous on the interval $[t_0, t_f]$, Ω is a compact convex set and $f(x(t), u(t)) \in \mathbb{R}^n$ is continuously differentiable with respect to x and t. It should be noted that system (2) is exactly system (1) when $u = 0$.

Throughout this paper we make the following assumptions.

Assumption 1. The nonlinear system (1) exhibits at least two distinct stable steady states (attractors) $x_1{}^*$ and $x_2{}^*$ ($x_1{}^* \neq x_2{}^*$), which indicates $F(x_1^*) = F(x_2^*) = 0$. In addition, D_1 and D_2 are the domains of attraction for attractors $x_1{}^*$ and $x_2{}^*$, respectively.

Assumption 2. The function $f(x, u)$ in (2) is Lipschitz continuous with respect to x and u; thus, there exist two constants $L_1, L_2 > 0$ such that

$$\|f(x, u) - f(y, u)\| \le L_1 \|x - y\|, \forall x, y \in \mathbb{R}^n, u \in \mathbb{R}^m. \tag{3}$$

$$\|f(x, u) - f(x, v)\| \le L_2 \|u - v\|, \forall x \in \mathbb{R}^n, u, v \in \mathbb{R}^m. \tag{4}$$

Assumption 3. Nonlinear system (1) is dissipative, i.e., there exist large positive constant C and time T such that $\|x(t)\| \le C$ for any initial state $x_0 \in \mathbb{R}^n$ and $t \ge T$. In other words, any solution of nonlinear system (1) eventually is bounded, which is actually true for biological systems in practice.

Definition 2.1 (Domain controllability)**.** The nonlinear system (2) is said to be domain controllable for two different attraction domains D_1 and D_2 if for any state x_0 in the attraction domain D_1, there exists a state x_2 in the attraction domain D_2, a finite time t_f and an admissible control defined on $[t_0, t_f]$ such that for $x(t_0) = x_0$, we have $x(t_f) = x_2$.

A schematic illustration of domain control is presented in Fig. 1. Our goal is to develop a new theoretical and computational framework to quantify the domain controllability of nonlinear dynamical networks and demonstrate the application of domain controllability analysis to identify potential drug targets during disease occurrence and progression. To address these problems, we consider the following control system

$$\begin{cases} \dot{x}(t) = F(x(t)) + Bu(t), \\ x(t_0) = x_0, t \in [t_0, t_f], \end{cases} \tag{5}$$
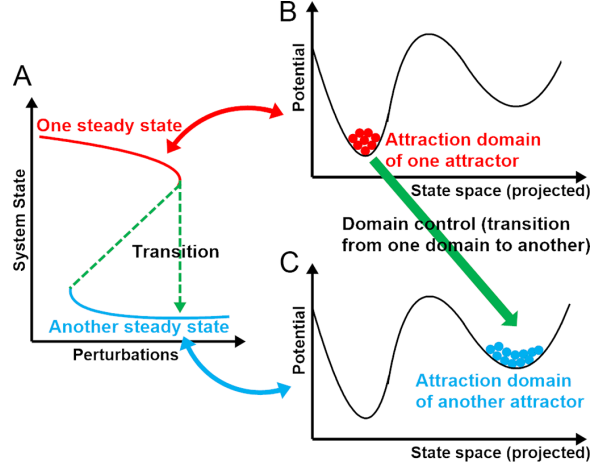
FIGURE 1. A schematic illustration of the domain control. **(A)** The nonlinear networked system can exhibit transitions between the different stable steady states. The red solid, blue solid and green dashed lines represent one stable steady state (attractor), another stable steady (another attractor) state and unstable states, respectively. The green dashed line with an arrow indicates the transition from one attractor to another attractor. **(B)** and **(C)** provide illustrations of the concept of domain control, namely the transition from the domain of attraction of one attractor to the domain of attraction of another attractor. **(B)** The domain of attraction of one attractor (corresponding to the red solid line in **(A)**). Each state in the domain of attraction is represented by a red ball. The horizontal axis is a projected state-space and the vertical axis is the potential, which indicates the relative instability of individual states. **(C)** The domain of attraction of another attractor (corresponding to the blue solid line in **(A)**). Each state in the domain of attraction is represented by a blue ball.

where $x(t) \in \mathbb{R}^n$ is the state vector, x_0 is in the attraction domain D_1 and $B \in \mathbb{R}^{n \times m}$ ($m \leq n$) is the control input matrix in which each column is a n-by-1 identity matrix. If the j-th column of matrix B is the i-th column of n-by-n identity matrix, it means that an input control signal is directly acted on variable i.

3. **Theoretical results.** Without loss of generality, we assume that the first m variables are selected and controlled, i.e., the sub-matrix consisting of the first m rows of B is an identity matrix. For convenience to prove the following Theorem 3.1, we use the following notations.

$$x(t) = \begin{bmatrix} X(t) \\ Y(t) \end{bmatrix}, F(x(t)) = \begin{bmatrix} g(x(t)) \\ h(x(t)) \end{bmatrix}, x_2^* = \begin{bmatrix} X^* \\ Y^* \end{bmatrix},$$

where

$$X(t) = (x_1(t), \cdots, x_m(t))^T \in \mathbb{R}^m, Y(t) = (x_{m+1}(t), \cdots, x_n(t))^T \in \mathbb{R}^{n-m},$$
$$g(x) = (F_1(x), \cdots, F_m(x))^T \in \mathbb{R}^m, h(x) = (F_{m+1}(x), \cdots, F_n(x))^T \in \mathbb{R}^{n-m},$$

$$X^* = (X_1^*, \cdots, X_m^*)^T \in \mathbb{R}^m, Y^* = \left(X_{m+1}^*, \cdots, X_n^*\right)^T \in \mathbb{R}^{n-m}.$$

Theorem 3.1. *Suppose that Assumptions 1, 2 and 3 hold. Then the nonlinear system (5) is domain controllable; that is, there exists a control function $u(t) \in \Omega \subseteq \mathbb{R}^m(m \le n)$ such that the corresponding solution $x(t)$ of system (5) satisfies $\lim_{t \to +\infty} \|x(t) - x_2^*\| = 0$.*

Proof. The proof includes two main steps. The first step is to determine the index set of control variables (nodes) that include m variables and another is to design the control $u(t)$ and prove the system (5) can be controlled.

First, we describe how to determine the index set of control variables (nodes) that include m variables.

Let K be an index set (possibly empty set) of variables. The purpose of K is to screen out the variables whose partial derivatives are not constantly negative. Specially, if there exists a time t and a state $x(t) \in \mathbb{R}^n$ such that the following inequality

$$\frac{\partial F_r(x(t))}{\partial x_r} \ge 0 \tag{6}$$

holds, then $r \in K$.

Let Q be an index set of variables containing q elements. We relabel the indexes of variables of the system (5) such that $Q = \{1, \cdots, q\}$, where q is the smallest positive integer satisfying

$$\begin{cases} N(p) \subseteq C, \text{ for } p = m + 1, \\ N(p) \subseteq C \bigcup \{m+1, \cdots, p-1\}, \text{ for } p \in \{m+2, \cdots, n\}, \end{cases} \tag{7}$$

where $C = Q \bigcup K = \{1, \cdots, m\}$ is the index set of control variables (nodes), $NC = \{m+1, \cdots, n\}$ is the index set of variables that do not need to be controlled and $N(p)$ is the index set of variables which are in the p-th equations of the system (5) but exclude the index p.

It is obvious that $m \le n$.

Second, we can design the control signal as follows

$$u(t) = -g(x(t)) + \lambda(X(t) - X^*), \tag{8}$$

where $\lambda < 0$. Assumption 3 implies uniform boundedness of $x(t)$ for all $t \ge t_0$. This implies uniform boundedness of $u(t)$. Thus $u(t) \in \Omega \subseteq \mathbb{R}^m$.

Recall that the first m rows of B constitutes an identity matrix. Substituting (8) into the nonlinear control system (5), we have

$$\dot{X}(t) = \lambda(X(t) - X^*) \triangleq \bar{g}(x(t)), \tag{9}$$

and

$$\dot{Y}(t) = h(x(t)). \tag{10}$$

Let

$$\bar{F}(x(t)) = \begin{bmatrix} \bar{g}(x(t)) \\ h(x(t)) \end{bmatrix} \triangleq \left(\bar{F}_1(x), \cdots, \bar{F}_n(x)\right)^T \in \mathbb{R}^n. \tag{11}$$

(9) implies

$$\lim_{t \to +\infty} x_c(t) = X_c^*, \text{ for all } c \in C, \tag{12}$$

and

$$\frac{\partial \bar{F}_l(x(t))}{\partial x_l} < 0, \ 1 \le l \le m. \tag{13}$$

Recall that Assumption 3 holds and that $F(x(t))$ is continuous. Combining (9), (11) and the definition of K, we have

$$\frac{\partial \bar{F}_l(x(t))}{\partial x_l} < 0, \ m+1 \le l \le n. \tag{14}$$

Combining (13) and (14) gives

$$\frac{\partial \bar{F}_l(x(t))}{\partial x_l} < 0, \ 1 \le l \le n. \tag{15}$$

Similar to Fiedler *et al.* [17], we use the mathematical induction method to prove

$$\lim_{t \to +\infty} x_i(t) = X_i^*, \ 1 \le i \le n. \tag{16}$$

First, the statement given in (16) obviously holds when $i = 1$. Second, we assume that (16) holds for all $i < k$. At last, we will show that (16) also holds when $i = k$. Let

$$\varphi(t) = x(t) - x_2^*,$$

where $\varphi(t) = (\varphi_1(t), \cdots, \varphi_n(t))^T \in \mathbb{R}^n$.

According to the generalized mean value theorem, we have

$$\begin{aligned}
\dot{\varphi}_k(t) &= \bar{F}_k(x(t)) - \bar{F}_k(X_k^*) \\
&= \sum_{j=1}^{n} \left(\int_0^1 \frac{\partial \bar{F}_k}{\partial x_j} \left(x(t) - \theta\varphi(t) \right) d\theta \right) \varphi_j(t).
\end{aligned} \tag{17}$$

(17) can be rewritten as

$$\dot{\varphi}_k(t) = \alpha_k(t)\varphi_k(t) + \sum_{j \in N(k)} \beta_j(t)\varphi_j(t), \tag{18}$$

where

$$\alpha_k(t) = \int_0^1 \left(\frac{\partial \bar{F}_k}{\partial x_k} \left(x(t) - \theta\varphi(t) \right) \right) d\theta,$$

$$\beta_j(t) = \int_0^1 \frac{\partial \bar{F}_k}{\partial x_j} \left(x(t) - \theta\varphi(t) \right) d\theta, \ j \in N(k).$$

Assumption 3 implies uniform boundedness of $x(t)$, and $x(t) - \theta\varphi(t)$ for all $t \ge T$. Since $\bar{F}(x)$ is continuously differentiable with respect to x, then $\frac{\partial \bar{F}_k}{\partial x_j} \left(x(t) - \theta\varphi(t) \right)$ is uniform bounded for a given j. This implies uniform boundedness of $\beta_j(t)$ based on the Property 6 of Integrals in [50]. Recall that (15) holds, according to the Property 5 of Integrals in [50], then $\alpha_k(t) < 0$ for all $t \ge t_0$. Thus, for all $t \ge T$, there exist constant α_0 and β_0 such that

$$\begin{aligned}
\alpha_k(t) &\le \alpha_0 < 0, \\
|\beta_j(t)| &\le \beta_0.
\end{aligned} \tag{19}$$

Using the variation of constants method, the solution of (18) can be written as

$$\begin{aligned}
\varphi_k(t) = {} & \varphi_k(t_0) \exp\left(\int_{t_0}^t \alpha_k(s)\,ds \right) \\
& + \sum_{j \in N(k)} \int_{t_0}^t \left(\exp\left(\int_s^t \alpha_k(r)\,dr \right) \beta_j(s)\,\varphi_j(s) \right) ds.
\end{aligned} \tag{20}$$

Combining (19) and (20) yields

$$
\begin{aligned}
|\varphi_k(t)| \leq{}& |\varphi_k(t_0)| \exp\left(\int_{t_0}^{t} \alpha_k(s)\,ds\right) \\
&+ \sum_{j \in N(k)} \int_{t_0}^{t} \left(\exp\left(\int_{s}^{t} \alpha_k(r)\,dr\right) |\beta_j(s)|\,|\varphi_j(s)|\right) ds \\
\leq{}& |\varphi_k(t_0)| e^{\alpha_0(t-t_0)} + \sum_{j \in N(k)} \int_{t_0}^{t} \left(\beta_0 e^{\alpha_0(t-s)} |\varphi_j(s)|\right) ds.
\end{aligned}
\tag{21}
$$

Obviously, $\lim_{t\to+\infty} |\varphi_k(t_0)| e^{\alpha_0(t-t_0)} = 0$. The integral terms on the right hand of (21) can be rewritten as

$$
\int_{t_0}^{t} \left(\beta_0 e^{\alpha_0(t-s)} |\varphi_j(s)|\right) ds = \beta_0 \int_{0}^{t-t_0} e^{\alpha_0 s} |\varphi_j(t-s)|\,ds, \ j \in N(k).
\tag{22}
$$

According to the induction hypothesis, we have $\lim_{t\to+\infty} \varphi_j(t-s) = 0$ for all $j \in N(k) \subseteq C \bigcup \{m+1, \cdots, k-1\}$. Thus, $e^{\alpha_0 s} |\varphi_j(t-s)|$ is bounded for a given s. Using the Lebesgue's dominated convergence theorem [38], we have

$$
\lim_{t\to+\infty} \int_{0}^{t-t_0} e^{\alpha_0 s} |\varphi_j(t-s)|\,ds = 0.
\tag{23}
$$

Thus $\lim_{t\to+\infty} |\varphi_k(t)| = 0$, i.e., $\lim_{t\to+\infty} x_k(t) = X_k^*$. □

Remark 1. There are two fundamental questions in the control of nonlinear dynamical networks, that is, how to identify driver nodes and design the control signals to achieve the control goal for a dynamical system. In the proof of Theorem 3.1, we established a strategy for identifying the driver nodes that are theoretically required to ensure the system domain controllable. Algorithm 1 presents detailed steps for identifying the theoretically required nodes (section 4.1.1). In addition, we also designed a general form of control functions (8) that depend on the dynamics and the desired state of the system for realizing domain controllability.

Remark 2. According to our theory and approaches, it is easy to find that the number of driver nodes is much less than the number of nodes for biochemical networks, i.e., $m \ll n$. On the one hand, biochemical processes take place usually involving the degradation of molecules (e.g., genes and proteins). Thus, degradation terms (e.g., $-d_i \cdot x_i$) generally appear in the mathematical models of biochemical networks, suggesting that the partial derivative of F_i with respect to x_i is usually negative; that is, K is an empty set based on (6). On the other hand, signaling networks are characterized by their upstream and downstream signaling pathways and cascade to shape a uniquely defined response to the initial stimulus [26, 43]. Thus, based on the method of determining the set Q, we can control the dynamics of the system as long as we control these extracellular signaling molecules in signaling networks. Altogether, the number m of the control set $C = Q \bigcup K$ may be much smaller than the number n of nodes in the node set of biochemical networks.

Remark 3. In biological models, the set K has obvious biological meanings. If there are not any self-feedbacks in all variables, then the set K is empty. Otherwise, the variables that have self-feedbacks will be selected into the set K when (6) holds. For a real biological system, very few variables have self-feedbacks. In Appendix A,

we give two simple but famous biological examples to explain how to obtain the set K and show the biological significance of the selected set K.

Remark 4. The set Q can be determined according to (7). The main idea of determining the set Q is that a set of variables are selected such that the dynamical equations of all of the remaining variables can be expressed as explicit functions of these selected variables by successive evaluation of their dynamical equations. According to (7), when $p = m + 1$, $N(m + 1) \subseteq C$ means the $(m + 1)$-th equation only contains the variables in the control set C. In other words, the $(m + 1)$-th equation can be expressed as an explicit function of the controlled variables. When $p = m+2$, $N(m+2) \subseteq C \bigcup \{m+1\}$ means the $(m+2)$-th equation only contains the variables in the control set C and the variable x_{m+1}, suggesting that the $(m + 2)$-th equation can be expressed as an explicit function of the controlled variables and the variable x_{m+1}. Therefore, the $(m + 2)$-th equation can also be expressed as an explicit function of the controlled variables. By this analogy, when $p = n$, $N(n) \subseteq C \bigcup \{m + 1, \cdots, n - 1\}$ means that the (n)-th equation only contains the variables in the control set C and the variables $\{x_{m+1}, x_{m+2}, \cdots, x_{n-1}\}$. Therefore, the dynamical equations of all of the remaining variables can be expressed as explicit functions of these controlled variables. To clearly show the relationships between the controlled variables and the remaining variables, we give a schematic diagram in Fig. A3. In addition, to further illustrate the definition of the set Q, two simple examples are given in Appendix A.

Remark 5. Because Theorem 3.1 provides a sufficient condition for the domain controllability of a system, the number of the calculated driver nodes based on this Theorem may be overestimated. Therefore, we numerically investigate whether a less number of driver nodes can also achieve domain controllability (section 4.1.2). To distinguish the differences of these two approaches, we denote them as theoretically and practically required driver nodes, respectively.

4. **Algorithms for solving the domain control problem.**

4.1. **Identifying driver nodes for realizing domain controllability.**

4.1.1. *Theoretically required driver nodes.* In the proof of the Theorem 3.1, we established a strategy for identifying the driver nodes that theoretically need to be controlled. The set of driver nodes equals the union of the sets K and Q, which can be determined using (6) and (7), respectively. Here, we gave an intuitive explanation and a simple algorithm of how to select the driver nodes. Firstly, we ignored the variables that are not influenced by any other variables. The dynamics of these variables converges to the stable steady states and does not contribute to the diversity of attractors of the system. We inductively repeated this procedure until there was no such variables. Secondly, the main idea of determining the set Q is that a set of variables was selected such that the dynamical equations of all of the remaining variables can be expressed as explicit functions of these selected variables by successive evaluation of their dynamical equations. Therefore, if we appropriately control the dynamics of these selected variables, we can drive the system to the domain of attraction of the desired attractor. Thirdly, the minimal set of such selected variables is exactly the set Q in our proof. In this study, we employed an exhaustive algorithm to search the set Q. More specifically, we presented the following algorithm to identify the theoretically required driver nodes.

Algorithm 1 Identifying the theoretically required driver nodes

---

Step 1. Determine the set K according to (6).

Step 2. Ignore the variables that do not be influenced by any other variables as described above.

Step 3. Select a subset of variables, denoted by Q.

Step 4. If the dynamical equations of all of the remaining variables can be only expressed as explicit functions of the variables in Q, then stop the iteration and return $C = Q \bigcup K$ as the set of driver nodes. Otherwise, go to the next step.

Step 5. Repeat Steps 3 and 4 using the exhaustive search of the candidate set by looking first at smaller sets.

4.1.2. *Practically required driver nodes.* The Algorithm 1 for identifying the theoretically required driver nodes is based on the proposed method for finding the set $C = Q \bigcup K$ of driver nodes in the proof of Theorem 3.1. The purpose is to demonstrate the correctness of our theory. However, the purpose for finding practically required driver nodes is to investigate whether a less number of driver nodes can also achieve domain controllability because the targets are very few in biological and medical networks. Therefore, the computational method for practically required driver nodes is very simple. That is, we select one node from the set of theoretically required driver nodes and use the designed control function (8), then perform the numerical simulation to determine whether the trajectory of the control system (5) can dynamically evolve to the desirable attractor using the designed control error (24). If the system cannot be controlled by controlling one node, then we select two nodes from the set of theoretically required driver nodes and repeat the simulations, and so on. In addition, for low-dimensional problems, we also test each variable (node) in the dynamical system to observe its domain controllability.

4.2. **Determining the attractors of a nonlinear dynamical system.** In this study, we combine the Runge-Kutta method and trust-region method to determine the attractors (steady states) of an n-dimensional nonlinear dynamical system such as that given in (1). It is impossible to enumerate all of the possible attractors of a high-dimensional nonlinear dynamical system, but we can confirm the minimum number of attractors by performing many rounds of sufficient sampling. The algorithm for determining the steady states of the nonlinear dynamical system (1) was presented in Algorithm 2.

4.3. **Finding the domains of attraction.** In this study, we estimate the domains of attraction using the Latin hypercube sampling (LHS) method. As biological systems are often bounded, to measure the domains of attraction more accurately, we find the domains of attraction in the positively invariant set of a dynamical system. A positively invariant set Δ is a subset of the state space of a dynamical system with the property that, if any solution to the dynamical system with initial conditions $x(0) \in \Delta$ satisfies $x(t) \in \Delta$ when $t \geq 0$. The procedure is presented in Algorithm 3.

5. **Numerical illustrations and biological implications.** In this section, we numerically illustrate the effectiveness of our proposed theoretical and computational frameworks as well as their biological implications. Network controllability

---

Algorithm 2 Determining the attractors of the nonlinear dynamical system

---

Step 1. Let P and A be the phase space and the set of attractors of the dynamical model (1), respectively. Given the total number of iterations N, termination tolerance ε, set iteration $i = 1$ and $A = \varnothing$.

Step 2. Randomly generate an initial state $x_0^i \in P$. Integrate system in (1) over a long time using the classical fourth-order Runge-Kutta method and let $x^i(t_f)$ be the solution x_i evaluated at the terminal time point t_f.

Step 3. Start at the point $x^i(t_f)$ and again try to solve the nonlinear equations $F(x) = 0$ using the trust-region method, which return the solution of the nonlinear equation, again denoted by x^i.

Step 4. Update $A = \{x^i, A\}$ if x^i satisfies the following criterion

$$\left\| F\left(x^i\right) - 0 \right\| < \varepsilon.$$

Step 5. If $i < N$, update $i = i + 1$ and return to Step 2.

Algorithm 3 Finding the domains of attraction

Step 1. Let D be the domain of attraction of an attractor (denoted by x^*) of the dynamical model (1). Given the total number of iterations N, termination tolerance ε, set iteration $i = 1$ and $D = \varnothing$.

Step 2. Randomly generate a sample point $x_0^i \in \Delta$. Integrate system in (1) over a long time by taking x_0^i as the initial state and let $x^i(t_f)$ be the solution x_i evaluated at the terminal time point t_f.

Step 3. Start at the point $x^i(t_f)$ and again try to solve the nonlinear equations $F(x) = 0$ using the trust-region method, which return the solution of the nonlinear equation, again denoted by x^i.

Step 4. Update $D = \{x_0^i, D\}$ if x^i satisfies the following criterion

$$\left\| x^i - x^* \right\| < \varepsilon.$$

Step 5. If $i < N$, update $i = i + 1$ and return to Step 2.

---

analysis has recently emerged as powerful tools for the identification of potential drug targets of complex diseases [63, 53]. The identification of critical genes during disease progression is providing new insights into the transitions of cellular phenotypes and possible therapeutic interventions. To demonstrate effectiveness of using domain controllability analysis as a general tool to discover novel disease genes and potential drug targets, we apply our proposed approach to three realistic disease-related networks: the epithelial-mesenchymal transition (EMT) core network [40], the T helper (Th) differentiation cellular network [38] and the cancer network [32].

We use the root-mean-square error (RMSE) to measure the quality of the domain control process. The control error between the control trajectory and the desired attractor at time t is defined as

$$e(t) = \sqrt{\frac{1}{n} \sum_{i=1}^{n} \left(x_i(t) - x_{2i}^*\right)^2}, \tag{24}$$

where $x_i(t)$ and x_{2i}^* is the i-th component of the state vector of the control system (5) and the desired attractor, respectively. n is the dimension of the system (1).

5.1. **Domain control of the EMT core network.** In this paper, we first used a simple EMT core network, as shown in Fig. 2A, to illustrate the effectiveness of
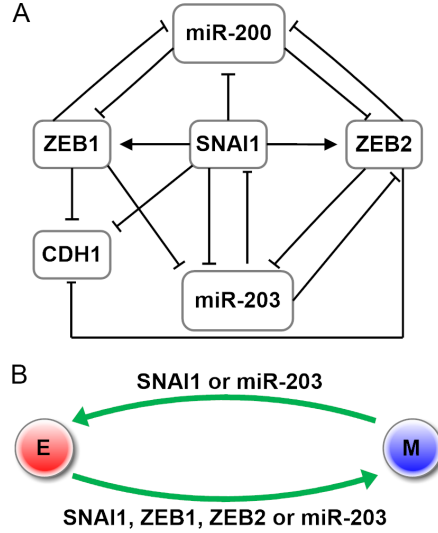
FIGURE 2. The EMT core network and domain control strategies for its phenotype transition. (**A**) Diagram (redrawn from [40]) of the EMT core network. The arrows and short bars represent activation and inhibition, respectively. For example, SNAI1 activates ZEB1 but inhibits miR-203 expression. (**B**) The practically required driver nodes (genes) for realizing domain controllability of this network. Mesenchymal cells can be induced from epithelial cells and back by any one of the four genes (SNAI1, ZEB1, ZEB2 and miR-203) and any one of the two genes (SNAI1 and miR-203), respectively.

our theory and the power of domain controllability analysis in identifying key genes that control the initiation and progression of EMT.

Based on the law of mass action and Michaelis-Menten kinetics, we built an ODE model of this network. The associated dynamics follows

$$
\begin{aligned}
\dot{x}_1 &= \frac{1}{1 + K_{12}x_2^2 + K_{13}x_3^2 + K_{14}x_4^2} - d_1 x_1, \\
\dot{x}_2 &= \frac{1}{1 + K_{26}x_6^2} - d_2 x_2, \\
\dot{x}_3 &= \frac{x_2^2}{1 + K_{32}x_2^2 + K_{35}x_5^2} - d_3 x_3, \\
\dot{x}_4 &= \frac{x_2^2}{1 + K_{42}x_2^2 + K_{45}x_5^2 + K_{46}x_6^2} - d_4 x_4, \\
\dot{x}_5 &= \frac{1}{1 + K_{52}x_2^2 + K_{53}x_3^2 + K_{54}x_4^2} - d_5 x_5, \\
\dot{x}_6 &= \frac{1}{1 + K_{62}x_2^2 + K_{63}x_3^2 + K_{64}x_4^2} - d_6 x_6,
\end{aligned}
\tag{25}
$$

where $x_i(i = 1, 2, 3, 4, 5, 6)$ represents the activity level of CDH1, SNAI1, ZEB1, ZEB2, miR-200 and miR-203, respectively. K_{ij} denotes the regulatory strength

from node j to i and d_i is the degradation rate of gene i. The parameters used to perform numerical simulations are in the following: $K_{ij} = 1, d_1 = 0.5, d_2 = 0.3, d_3 = 0.5, d_4 = 0.5, d_5 = 0.5, d_6 = 0.3$.

As a preparation, we first employed Algorithm 2 to identify the attractors of this model. As shown in Table 1, we confirmed the existence of two attractors in this network, which correspond to the epithelial and mesenchymal states. This result is consistent with previous findings [40]. Second, to find the domains of attraction of this model, we give the positively invariant set of this model, as shown in Table 2. Applying Algorithm 3 to this model, we uniformly took 10,000 samples ($N = 10,000$) from the positively invariant set and obtained the domains of attraction. Because it is impossible to show domains of attraction in high dimensional space, we projected the domains onto two-dimensional planes. The domains of attraction of this model are depicted in Fig. C1(see Appendix C).

TABLE 1. The attractors of the EMT network. The two observed attractors A and B correspond to the epithelial state and mesenchymal state [40], respectively.

	A	B		A	B
x_1 (CDH1)	1.7924	0.1111	x_4 (ZEB2)	0.0170	1.8169
x_2 (SNAI1)	0.3359	3.2229	x_5 (miR-200)	1.7924	0.1111
x_3 (ZEB1)	0.0522	1.8224	x_6 (miR-203)	2.9873	0.1851

TABLE 2. The positively invariant sets of the mathematical models of the three disease-related networks.

Networks	Positively invariant sets
EMT	$\left\{ (x_1, \cdots, x_6) \in \mathbb{R}^6 \left\| 0 < x_i < \frac{1}{d_i}, i = 1, \cdots, 6 \right. \right\}$
T helper	$\left\{ x \in \mathbb{R}^{23} \left\| \begin{array}{l} 0 \le x_5 < \frac{1}{d_4 d_5}, 0 \le x_6 < \frac{1}{d_1 d_6}, 0 \le x_7 < \frac{1}{d_1 d_6 d_7}, 0 \le x_{14} < \frac{1}{d_{11} d_{14}}, \\ 0 \le x_{17} < ((\frac{1}{d_3} + \frac{1}{d_{15}})\frac{1}{d_{18}} + \frac{1}{d_{22}})\frac{1}{d_{17}}, 0 \le x_{18} < (\frac{1}{d_3} + \frac{1}{d_{15}})\frac{1}{d_{18}}, \\ 0 \le x_{19} < \frac{1}{d_1 d_6 d_7 d_{19}}, 0 \le x_{20} < \frac{1}{d_9 d_{20}}, 0 \le x_{21} < \frac{1}{d_{13} d_{21}}, \\ 0 \le x_i < \frac{1}{d_i}, 0 \le x_j < 1, i = 1, 3, 4, 9, 11, 12, 13, 15, 16, j = 8, 10, 23 \end{array} \right. \right\}$
Cancer	$\left\{ (x_1, \cdots, x_{32}) \in \mathbb{R}^{32} \left\| 0 \le x_i \le 1, i = 1, \cdots, 32 \right. \right\}$

We take this network as an example to describe how to identify the driver nodes using our established strategy (section 4.1.1). First, based on the dynamical equations of this network (25), it is clear that there does not exist an index of a variable satisfying (6), i.e., K is an empty set. Second, we rewrote the dynamical equations (25) as the following form:

$$\begin{cases} \dot{x}_1 = G_1(x_2, x_3, x_4) - D_1(x_1), \\ \dot{x}_2 = G_2(x_6) - D_2(x_2), \\ \dot{x}_3 = G_3(x_2, x_5) - D_3(x_3), \\ \dot{x}_4 = G_4(x_2, x_5, x_6) - D_4(x_4), \\ \dot{x}_5 = G_5(x_2, x_3, x_4) - D_5(x_5), \\ \dot{x}_6 = G_6(x_2, x_3, x_4) - D_6(x_6), \end{cases} \qquad (26)$$

where $G_i(x_j)$ and $D_i(x_i)$ indicate the terms related to other variables ($j \neq i$) and itself (x_i) in the i-th equation, respectively. By successive evaluation of the dynamical equations, we obtained the following equivalent system:

$$
\begin{cases}
\dot{x}_1 = G_1(G_2(x_6), G_3(G_2(x_6), x_5), G_4(G_2(x_6), x_5, x_6)) - D_1(x_1), \\
\dot{x}_2 = G_2(x_6) - D_2(x_2), \\
\dot{x}_3 = G_3(G_2(x_6), x_5) - D_3(x_3), \\
\dot{x}_4 = G_4(G_2(x_6), x_5, x_6) - D_4(x_4), \\
\dot{x}_5 = G_5(x_2, x_3, x_4) - D_5(x_5), \\
\dot{x}_6 = G_6(x_2, x_3, x_4) - D_6(x_6),
\end{cases}
\tag{27}
$$

As shown in (27), the dynamical equations of the remaining variables (i.e., x_1, x_2, x_3, x_4) are explicit functions of selected variables (i.e., x_5, x_6). Therefore, the dynamics of the remaining variables is determined uniquely by that of the selected variables, suggesting that the control of x_5 and x_6 can indeed steer the system (25) to evolve to the desired attractor. In addition, according to (26), it can easily be shown that $N(2) = \{6\} \subseteq \{5,\ 6\}$, $N(3) = \{2,5\} \subseteq \{5,\ 6\} \bigcup \{2\}$, $N(4) = \{2,5,6\} \subseteq \{5,\ 6\} \bigcup \{2\}$, $N(1) = \{2,3,4\} \subseteq \{5,\ 6\} \bigcup \{2,3,4\}$. Recall that K is an empty set. Therefore, we can set $Q = \{5,\ 6\}$. Accordingly, $C = Q \bigcup K = \{5,\ 6\}$. Finally, we relabel the variables in accordance with the order $\{x_5, x_6, x_2, x_3, x_4, x_1\}$, then (7) holds.

The above analysis indicates that we can drive epithelial cells to transition into mesenchymal cells and back by controlling two genes, i.e., miR-200 (x_5) and miR-203 (x_6). Based on the designed control function (8), we can obtain the control function for the theoretically required driver nodes (i.e., miR-200 and miR-203), which is presented in Table 3. Based on this control function, we simulated the dynamics while controlling these two genes. The node state trajectories $x(t)$, the control function $u(t)$ and the control error $e(t)$ are depicted in Fig. C2 (see Appendix C), which indicates that the control errors remained very small ($e(t) < 10^{-3}$) with the time evolution, suggesting that our approach accurately controls the network dynamics. Furthermore, we randomly selected 10,000 points from the domain of attraction of the undesirable state and found that 100% control eventually caused the network to evolve to the desired target state using our method.

To show how the parameter λ in our designed control function influences the control process, we perform numerical simulations by changing the value of λ. Fig. C3 (see Appendix C) shows that a higher absolute value of λ indicates a less time the system takes to evolve to the desirable attractor. However, there is no significant influence when the absolute value of λ is larger than a certain value. When we perform numerical simulations for the three disease-related networks, the parameter λ is set to be -0.2.

To identify the practically required driver nodes (section 4.1.2), we tested each of the six nodes and found that controlling any one of the four genes (SNAI1, ZEB1, ZEB2 and miR-203) can drive the system to switch from the epithelial state to the mesenchymal state, and controlling any one of the two genes (SNAI1 and miR-203) can drive the system to transition back. Fig. $2B$ summarizes the practically required driver nodes (genes) for inducing transitions between the epithelial and mesenchymal states. Given an initial point in the domain of attraction of the epithelial state, we simulated the dynamics while controlling a single node. Based

TABLE 3. The theoretically required driver nodes and designed control functions for realizing domain control of the three disease-related networks. $x_{2,i}^*$ is the i-th component of the state vector of the desired attractor.

Networks	Driver nodes	Control functions ($u(t)$)
EMT	miR-200 and miR-203	$-\left(\dfrac{1}{1+K_{52}x_2^2+K_{53}x_3^2+K_{54}x_4^2}-d_5x_5\right)+\lambda\left(x_5-x_{2,5}^*\right)$ $-\left(\dfrac{1}{1+K_{62}x_2^2+K_{63}x_3^2+K_{64}x_4^2}-d_6x_6\right)+\lambda\left(x_6-x_{2,6}^*\right)$
T helper	GATA-3 and STAT1	$-\left(\dfrac{k_{1-1}x_1+k_{1-21}x_{21}}{1+k_{1-1}x_1+k_{1-21}x_{21}+k_{1-22}x_{22}}-d_1x_1\right)+\lambda\left(x_1-x_{2,1}^*\right)$ $-(k_{18-3}x_3+k_{18-15}x_{15}-d_{18}x_{18})+\lambda\left(x_{18}-x_{2,18}^*\right)$
Cancer	P53, RB, AKT, EGFR, HIF1, CDK2, CDK1, BCL2 and NFκB	$-\left(\dfrac{a(s_1x_1^n+s_9x_9^n+s_{14}x_{14}^n)}{3S^n+s_1x_1^n+s_9x_9^n+s_{14}x_{14}^n}+\dfrac{bS^n}{S^n+1/2s_{16}x_{16}^n+1/2s_{31}x_{31}^n}-kx_2\right)+\lambda\left(x_2-x_{2,2}^*\right)$ $-\left(\dfrac{2bS^n}{S^n+1/2s_{17}x_{17}^n+1/2s_{18}x_{18}^n}-kx_6\right)+\lambda\left(x_6-x_{2,6}^*\right)$ $-\left(\dfrac{a(s_{11}x_{11}^n+s_{13}x_{13}^n)}{S^n+1/3s_{11}x_{11}^n+1/3s_{12}x_{12}^n+1/3s_{13}x_{13}^n)/3}+\dfrac{bS^n}{S^n+s_4x_4^n}-kx_{10}\right)+\lambda\left(x_{10}-x_{2,10}^*\right)$ $-\left(\dfrac{a(s_9x_9^n+s_{10}x_{10}^n+s_{11}x_{11}^n+s_{13}x_{13}^n+s_{14}x_{14}^n+s_{15}x_{15}^n)}{6S^n+s_9x_9^n+s_{10}x_{10}^n+s_{11}x_{11}^n+s_{13}x_{13}^n+s_{14}x_{14}^n+s_{15}x_{15}^n}+\dfrac{bS^n}{S^n+s_2x_2^n}-kx_{12}\right)+\lambda\left(x_{12}-x_{2,12}^*\right)$ $-\left(\dfrac{a(s_{10}x_{10}^n+s_{13}x_{13}^n+s_{25}x_{25}^n+s_{26}x_{26}^n+s_{28}x_{28}^n)}{5S^n+s_{10}x_{10}^n+s_{13}x_{13}^n+s_{25}x_{25}^n+s_{26}x_{26}^n+s_{28}x_{28}^n}+\dfrac{bS^n}{S^n+s_2x_2^n}-kx_{14}\right)+\lambda\left(x_{14}-x_{2,14}^*\right)$ $-\left(\dfrac{a(s_9x_9^n+s_{10}x_{10}^n+s_{20}x_{20}^n)}{3S^n+s_9x_9^n+s_{10}x_{10}^n+s_{20}x_{20}^n}+\dfrac{bS^n}{S^n+1/2s_3x_3^n+1/2s_6x_6^n}-kx_{17}\right)+\lambda\left(x_{17}-x_{2,17}^*\right)$ $-\left(\dfrac{2bS^n}{S^n+1/3s_3x_3^n+1/3s_5x_5^n+1/3s_{30}x_{30}^n}-kx_{19}\right)+\lambda\left(x_{19}-x_{2,19}^*\right)$ $-\left(\dfrac{a(s_{22}x_{22}^n+s_{23}x_{23}^n)/2}{S^n+1/2s_{22}x_{22}^n+1/2s_{23}x_{23}^n}+\dfrac{bS^n}{S^n+1/2s_{10}x_{10}^n+1/2s_{24}x_{24}^n}-kx_{21}\right)+\lambda\left(x_{21}-x_{2,21}^*\right)$ $-\left(\dfrac{a(s_{10}x_{10}^n+s_{13}x_{13}^n)/2}{S^n+1/2s_{10}x_{10}^n+1/2s_{13}x_{13}^n}+\dfrac{bS^n}{S^n+s_4x_4^n}-kx_{25}\right)+\lambda\left(x_{25}-x_{2,25}^*\right)$

TABLE 4. An example of the practically required driver nodes and designed control functions for realizing domain control of the three disease-related networks. $x_{2,i}^*$ is the i-th component of the state vector of the desired attractor.

| Transitions	Drivers	Control functions ($u(t)$)
E↔M	miR-203	$-\left(\frac{1}{1+K_{62}x_2^2+K_{63}x_3^2+K_{64}x_4^2} - d_6x_6\right) + \lambda\left(x_6 - x_{2,6}^*\right)$
Th0↔Th1	T-bet	$-\left(\frac{k_{22-18}x_{18}+k_{22-22}x_{22}}{1+k_{22-1}x_1+k_{22-22}x_{22}+k_{22-18}x_{18}} - d_{22}x_{22}\right) + \lambda\left(x_{22} - x_{2,22}^*\right)$
Th0↔Th2	GATA-3	$-\left(\frac{k_{1-1}x_1+k_{1-21}x_{21}}{1+k_{1-1}x_1+k_{1-21}x_{21}+k_{1-22}x_{22}} - d_1x_1\right) + \lambda\left(x_1 - x_{2,1}^*\right)$
C↔A	AKT	$-\left(\frac{a(s_{11}x_{11}{}^n+s_{12}x_{12}{}^n+s_{13}x_{13}{}^n)}{3S^n+s_{11}x_{11}{}^n+s_{12}x_{12}{}^n+s_{13}x_{13}{}^n} + \frac{bS^n}{S^n+s_4x_4{}^n} - kx_{10}\right) + \lambda\left(x_{10} - x_{2,10}^*\right)$
C↔N	RB	$-\left(\frac{4bS^n}{2S^n+s_{17}x_{17}{}^n+s_{18}x_{18}{}^n} - kx_6\right) + \lambda\left(x_6 - x_{2,6}^*\right)$
A↔N	NFκB	$-\left(\frac{a(s_{10}x_{10}{}^n+s_{13}x_{13}{}^n)}{2S^n+s_{10}x_{10}{}^n+s_{13}x_{13}{}^n} + \frac{bS^n}{S^n+s_4x_4{}^n} - kx_{25}\right) + \lambda\left(x_{25} - x_{2,25}^*\right)$
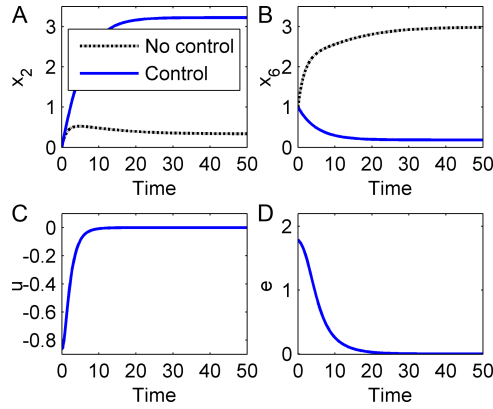
FIGURE 3. Domain control of the EMT core network transition from the epithelial state to the mesenchymal state. The representative node state trajectories $x(t)$ (**A-B**), control function $u(t)$ (**C**) and control error $e(t)$ (**D**) are depicted when the practically required driver node, namely miR-203, is controlled. The symbols x_2 and x_6 indicate the activities of SNAI1 and miR-203, respectively. The black dashed and blue solid lines represent the initial epithelial state (i.e., the undesired state with no control) and desired mesenchymal state, respectively.

on the designed control function (8), we can also obtain the control function for the practically required driver nodes (e.g., miR-203), which is presented in Table 4. Based on this designed control function, Fig. 3 shows the results obtained while controlling node miR-203 and the corresponding time evolution of the genes and the control function. The control error results show that we were able to accurately steer the system from the epithelial state to the mesenchymal state. We observed that we could drive the network to the desired state by controlling any one of the three genes SNAI1, ZEB1 and miR-203, which is in line with the previous findings that these genes are important regulators in the EMT [40, 31], suggesting a potential application of our approach to uncover key genes in the EMT program.
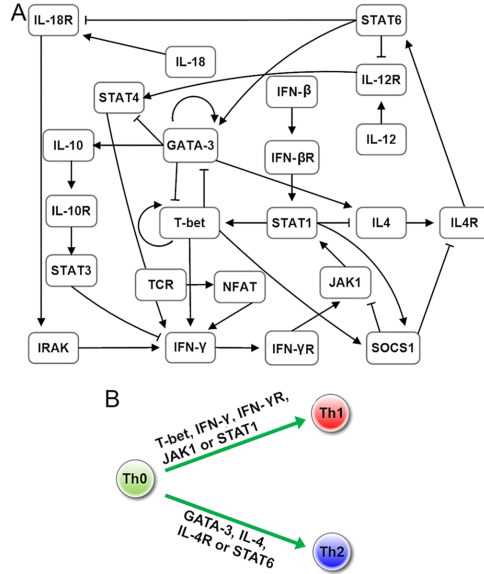
FIGURE 4. The Th differentiation cellular network and domain
control strategies for its phenotype transition. (**A**) Diagram (re-
drawn from the reference [38]) of the Th differentiation cellular
network. (**B**) The practically required driver nodes (genes) for re-
alizing domain controllability of this network. Th1 and Th2 cells
can be induced from Th0 cells by any one of the five genes (T-bet,
IFN-γ, IFN-γR, JAK1 and STAT1) and any one of the four genes
(GATA-3, IL-4, IL-4R and STAT6), respectively.

5.2. **Domain control of the Th differentiation cellular network.** Next, we
applied our proposed theoretical and computational frameworks to a core regula-
tory network controlling the differentiation of Th cells, which consists of 23 nodes
(Fig. 4). Starting with the discrete model presented in a previous study [16], we
constructed an ODE model (see (B1) in Appendix B). Applying Algorithm 2 to
this model, numerical simulations show that there are three attractors, which cor-
respond to the Th0, Th1 and Th2 phenotypes (Table 5). The patterns of activity
of these attractors are qualitatively consistent with those identified in a previous
study [16]. In addition, we obtained the corresponding domains of attraction by
using Algorithm 3, which are shown in Fig. C4 (see Appendix C).

Applying our theory to the Th differentiation cellular network leads to the find-
ings that controlling two genes (GATA-3 and STAT1) can force the system to per-
form transitions among the three phenotypes in the Th differentiation process. The
designed control function for these two genes was shown in Table 3. Furthermore,
a single driver node (T-bet, IFN-γ, IFN-γR, JAK1 or STAT1) was observed when
attempting to identify the practically required driver nodes of this network for the
transition from Th0 to Th1, and another driver node (GATA-3, IL-4, IL-4R or S-
TAT6) was identified for the transitions from Th0 to Th2 and Th1 to Th2. Fig. 4B
summarizes the practically required driver nodes needed to achieve domain control
of the transitions from Th0 state to the Th1 and Th2 states. Using the designed
control function (Table 4), Fig. 5 shows that the control of one of the two driver

TABLE 5. The attractors of the Th differentiation cellular network. The three observed attractors (A, B and C) correspond to the Th0, Th1 and Th2 phenotypes [16], respectively.

|	A	B	C		A	B	C
x_1 (GATA-3)	0.0000	0.0000	2.0238	x_{13} (IL-4R)	0.0000	0.0000	1.3358
x_2 (IFN-β)	0.0000	0.0000	0.0000	x_{14} (IRAK)	0.0000	0.0000	0.0000
x_3 (IFN-βR)	0.0000	0.0000	0.0000	x_{15} (JAK1)	0.0000	0.4406	0.0000
x_4 (IFN-γ)	0.0000	2.1012	0.0000	x_{16} (NFAT)	0.0000	0.0000	0.0000
x_5 (IFN-γR)	0.0000	2.1012	0.0000	x_{17} (SOCS1)	0.0000	2.1459	0.0000
x_6 (IL-10)	0.0000	0.0000	2.0238	x_{18} (STAT1)	0.0000	0.4406	0.0000
x_7 (IL-10R)	0.0000	0.0000	2.0238	x_{19} (STAT3)	0.0000	0.0000	2.0238
x_8 (IL-12)	0.0000	0.0000	0.0000	x_{20} (STAT4)	0.0000	0.0000	0.0000
x_9 (IL-12R)	0.0000	0.0000	0.0000	x_{21} (STAT6)	0.0000	0.0000	2.2264
x_{10} (IL-18)	0.0000	0.0000	0.0000	x_{22} (T-bet)	0.0000	1.7053	0.0000
x_{11} (IL-18R)	0.0000	0.0000	0.0000	x_{23} (TCR)	0.0000	0.0000	0.0000
x_{12} (IL-4)	0.0000	0.0000	2.0094				
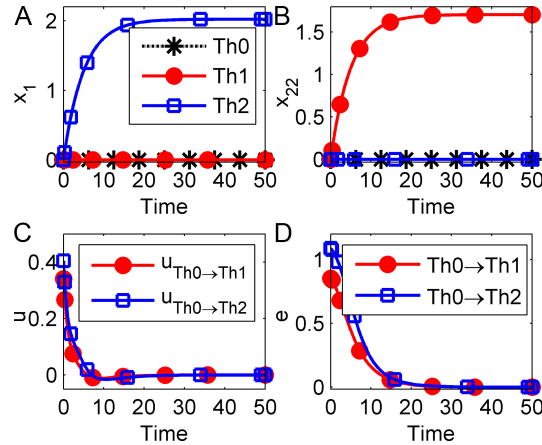
FIGURE 5. Domain control of the Th differentiation cellular network transitions from Th0 to Th1 and Th2. The representative node state trajectories $x(t)$ (**A, B**), control function $u(t)$ (**C**) and control error $e(t)$ (**D**) are depicted. The symbols x_1 and x_{22} indicate the activities of GATA-3 and T-bet, respectively. The black dashed line indicates the Th0 phenotype, i.e., the initial state with no control. The red line with circle markers and blue line with square markers represent transitions from Th0 to Th1 and Th0 to Th2, respectively. The driver node is T-bet when Th0 is steered to the Th1 phenotype (Th0→Th1) and GATA-3 when Th0 is driven to the Th2 phenotype (Th0→ Th2).

nodes, T-bet and GATA-3, can indeed steer a cell's differentiation from Th0 to either Th1 or Th2. The simulation results for the transition from Th1 to Th2 when controlling GATA-3 are shown in Fig. C5 (see Appendix C). These findings are consistent with the experimental results that stimulating the system with either T-bet or IFN-γ results in differentiations from Th0 to Th1, and with either GATA-3
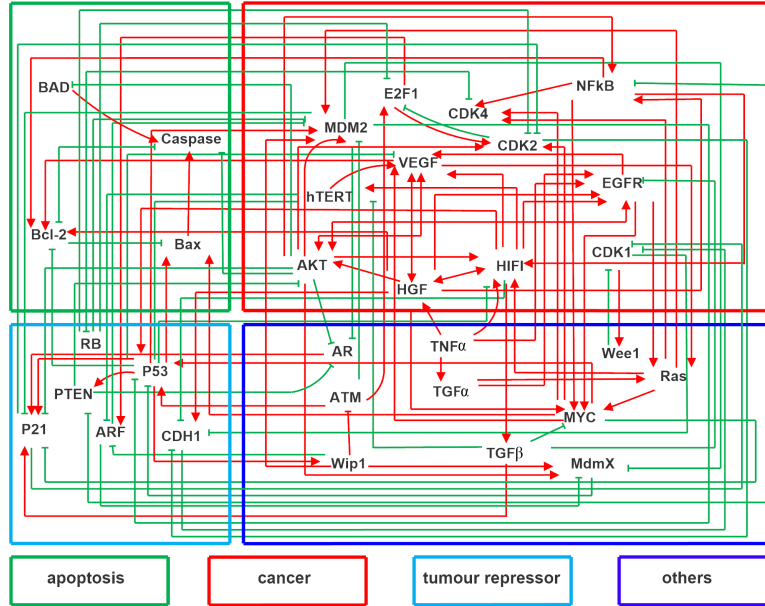
FIGURE 6. Diagram of the cancer network. This network is re-drawn from the reference [32], including 32 nodes (genes) and 111 edges (66 activation interactions and 45 repression interactions). The network mainly includes four types of marker genes: apoptosis marker genes (green rectangles), cancer marker genes (red rectangles), tumor repressor genes (light blue rectangles) and other genes (blue rectangles).The red arrows represent activation and the green short bars represent repression.

or IL-4 leads to differentiations from Th0 to Th2 [38, 22, 23]. These results suggest that the driver genes revealed by domain controllability analysis may play a vital role in the transitions among these three phenotypes during the Th differentiation process.

5.3. **Domain control of the cancer network.** Finally, we applied our proposed theoretical and computational frameworks to a cancer gene regulatory network, as shown in Fig. 6. The corresponding ODEs describing the dynamics of the underlying system were constructed in a previous study [32]. This system describes the expression level of 32 genes (nodes) and 111 regulations (edges) ((B2) in Appendix B). Applying algorithm 2 to the constructed mathematical model, we observed three attractors characterizing three important biological states: normal, cancer and apoptosis (Table 6). In addition, we obtained the corresponding domains of attraction by using Algorithm 3, which are depicted in Fig. C6 (see Appendix C).

Table 6: The attractors of the cancer network. The three observed attractors (A, B and C) correspond to the apoptosis, normal and cancer states [32], respectively.

	A	B	C		A	B	C

x_1 (ATM)	0.4165	0.4277	0.4712	x_{17} (CDK2)	0.1336	0.4613	0.8022
x_2 (P53)	0.4668	0.4642	0.4545	x_{18} (CDK4)	0.1342	0.4273	0.8620
x_3 (P21)	0.5700	0.4511	0.4378	x_{19} (CDK1)	0.5788	0.4853	0.5550
x_4 (PTEN)	0.7438	0.2973	0.2836	x_{20} (E2F1)	0.1922	0.2620	0.3441
x_5 (CDH1)	0.3215	0.6263	0.5298	x_{21} (Caspase)	0.8766	0.0688	0.0621
x_6 (RB)	0.9970	0.7105	0.1921	x_{22} (BAX)	0.7159	0.2641	0.2498
x_7 (ARF)	0.2756	0.2645	0.3087	x_{23} (BAD)	0.8486	0.1024	0.0923
x_8 (AR)	0.4134	0.2439	0.1955	x_{24} (BCL2)	0.1740	0.7533	0.7705
x_9 (MYC)	0.6647	0.4760	0.4758	x_{25} (NFκB)	0.1433	0.8853	0.9007
x_{10} (AKT)	0.3044	0.8058	0.8294	x_{26} (RAS)	0.4089	0.5158	0.5427
x_{11} (EGFR)	0.5262	0.4606	0.4636	x_{27} (TGFα)	0.0000	0.0000	0.0000
x_{12} (VEGF)	0.4145	0.6239	0.6464	x_{28} (TNFα)	0.0000	0.0000	0.0000
x_{13} (HGF)	0.1484	0.5908	0.6214	x_{29} (TGFβ)	0.0806	0.6772	0.7016
x_{14} (HIF1)	0.2998	0.6632	0.6823	x_{30} (Wee1)	0.6282	0.4550	0.5882
x_{15} (hTERT)	0.3765	0.4680	0.4702	x_{31} (MdmX)	0.6915	0.8148	0.6413
x_{16} (MDM2)	0.2471	0.4911	0.7550	x_{32} (Wip1)	0.4933	0.4877	0.4666

By applying the strategy for identifying the theoretically required driver nodes (Algorithm 1) to the cancer network, we revealed that control of the nine genes (e.g., P53, RB, AKT, EGFR, HIF1, CDK2, CDK1, BCL2 and NFκB) is sufficient to induce transitions among the cancer, normal and apoptosis states. The designed control function for these genes was shown in Table 3.

Furthermore, we tried to identify the practically required driver nodes for inducing transition from one domain of attraction to another in this network (section 4.1.2). Strikingly, a single driver node, such as RB (a tumor suppressor gene) and CDK2 (an oncogene), was detected by the application of our designed control function ((8) and Table 4) to this network for induction of the transition from the cancer to normal states, and another driver node, such as AKT (an oncogenes), PTEN (a tumor suppressor gene) and NFκB, was observed to cause the transitions from the cancer to apoptosis states. Given an initial point in the domain of attraction of the cancer state, we simulated the dynamics when one of the two genes AKT and RB was controlled. Fig. 7 shows the time evolution of the key genes and the control functions. The control errors show that we can accurately steer the system to switch from the cancer state to the apoptosis state and the normal state by controlling one of these genes, respectively.

Moreover, our study also revealed that only a single node, namely MDM2, RB, P53, CDK2 or CDK4, is needed to drive the normal state to the cancer state and that another driver node, namely AKT, PTEN or NFκB, is sufficient to switch the system from the normal state to the apoptosis state (see Fig. C7 in Appendix C). We took the driver nodes RB and NFκB as examples to show the designed control functions for achieving the transitions from the normal state to the cancer and apoptosis states, respectively (4). In addition, controlling any one of the six genes, including NFκB, AKT, PTEN, VEGF, HGF and HIF1, can cause the system to switch from the apoptosis state to the normal state and controlling AKT can induce an apoptosis-to-cancer transition (see Fig. C8 in Appendix C). Taken together, our results revealed that AKT, RB, PTEN, NFκB, CDK2 and VEGF play critical roles in the transitions among the cancer, normal and apoptosis states and may be possible targets for preventing cancer occurrence or transforming the early cancer state back to the normal state, which is consistent with findings that the repression
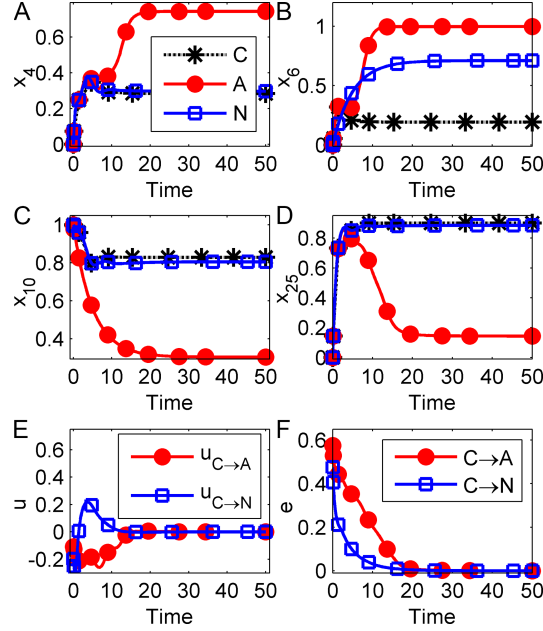
FIGURE 7. Domain control of the cancer network transitions from the cancer state to the apoptosis state and normal state. The representative node state trajectories $x(t)$ **(A-D)**, control function $u(t)$ **(E)** and control error $e(t)$ **(F)** are depicted when AKT and RB are controlled, respectively. The symbols x_4, x_6, x_{10} and x_{25} indicate the activities of PTEN, RB, AKT and NF-κB, respectively. The black dashed lines with star markers indicate the cancer state, i.e., the undesired state with no control. The blue solid lines with squares and red solid lines with filled circles represent transitions from cancer (C) to apoptosis (A) (denoted by C→A) and cancer (C) to normal (N) (denoted by C→N), respectively. The driver node is AKT when cancer is steered to the apoptosis and RB when cancer is driven to the normal state.

of key oncogene AKT will inhibit the formation of cancer state [20, 58] and that PTEN and NFκB play an integral role in preventing the onset and progression of numerous cancers [48, 8], implying that control of these genes can attenuate cancer by inducing a cancer-to-normal transition or cell apoptosis. Fig. 8 summarizes the practically required driver nodes needed to achieve domain control of the transitions among the cancer, normal and apoptosis states found using our approach.

5.4. **Biological implications.**

5.4.1. *New systems biology approaches to reveal the underlying mechanisms.* Network based systems biology approaches have recently emerged as powerful tools for the study of complex diseases [6] because these promising approaches allow us to take a more systems-level view of a biological process and thereby identify multiple factors (such as genes, proteins and their regulations) responsible for maintaining different cellular phenotypes and driving different cell fate decisions. In this study, using a novel network control framework, we identified the key proteins and their
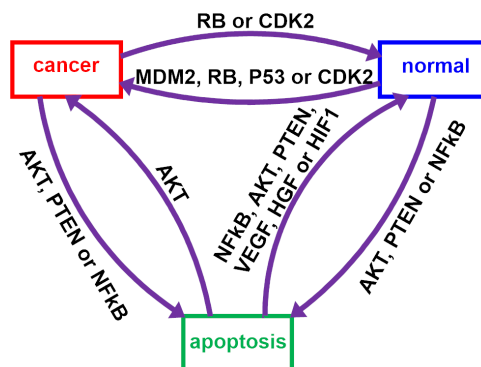
FIGURE 8. Domain control strategies for phenotype transitions of
the cancer network. The practically required driver nodes (genes)
for realizing domain controllability of this network were presented.
For example, the normal and apoptosis states can be induced from
the cancer state by any one of the two genes (RB and CDK2) and
any one of the three genes (AKT, PTEN and NFκB), respectively.

interactions in biological networks that determine the transitions between normal
and disease states. For example, in the EMT network that is established during
developmental and disease processes, we identified that any one of the four genes
(SNAI1, ZEB1, ZEB2 and miR-203) can induce the system to switch between the
epithelial and mesenchymal states (Figs. 2B and 3). Interestingly, results from
biological experiments have demonstrated that SNAI1 and ZEB1 can promotes tu-
morigenicity and miR-203 functions as a tumor suppressor in basal cell carcinoma
[59, 14, 27], suggesting that these genes are crucial for the EMT program and dys-
regulation of them may result in cancerization. In the Th differentiation cellular
network, we found that T-bet is needed for the transition from Th0 to Th1 and
that GATA-3 is needed for the transition from Th0 to Th2 (Figs. 4B and 5), which
agrees with previous reports [22, 23]. Importantly, we determined that one of three
genes, namely IFN-γR, JAK1 and STAT1, is required for the transition from Th0 to
Th1 and that one of the two genes IL-4R and STAT6 is required for the transitions
from Th0 to Th2 and Th1 to Th2, respectively, which are findings that have not
previously been reported. In the cancer network, our results uncovered that activa-
tion of RB(a tumor repressor gene) activity and AKT(an oncogenes) activity can
drive the system from the normal state to the cancer state and the apoptosis state
(Fig. C7), respectively, providing possible mechanisms for cancer formation. Taken
together, these results predict that disease-associated genes and proteins. There-
fore, our control framework provides a novel approach for uncovering the underlying
mechanisms through which proteins and their interactions lead to diseases.

5.4.2. *Network-based drug targets for complex diseases.* Complex diseases are often
regulated by the underlying biological networks, which can exhibit complex non-
linear dynamical behaviors. With the increase of network structure and dynamics
being uncovered, the design of network-based drugs may be feasible and thereby
provide an attractive strategy to potentially treat complex diseases such as can-
cer [2, 13, 44]. In our analyses, we looked at diseases as an undesirable network
state and reframed the goal of drug therapy toward shifting the network back to

a normal state. In the cancer network, we demonstrated that NFκB is sufficient to switch the system from the apoptosis state to the normal state and that AKT is sufficient to induce the transition of the system from the apoptosis state to the cancer state (Fig. C8). More importantly, our results provide certain predictions regarding which genes in the cancer network are critical to the transitions from the cancer to the normal and apoptosis states. We revealed that a single driver node, such as RB or CDK2 (an oncogene), can induce the transition from the cancer state to the normal state and that another driver node, such as AKT, PTEN (a tumor suppressor gene) or NFκB (a critical link between inflammation and cancer [8]) can induce cancer cell death (Figs. 7 and 8). These results may provide two rational drug designs of anti-cancer strategies, namely returning the early cancer state back to the normal state or inducing the death of cancer cells, for the prevention of cancer initiation and progression by perturbation of the activities of these genes through experimental techniques. Taken together, the key nodes and their dynamics identified in this study can be used to guide the design of anti-cancer strategies that involve targeting key genes or their regulations.

In summary, the approach described here will not only provide a new systems biology approach for revealing novel pathogenic mechanisms of complex diseases through which genes or proteins lead to disease but also aid the development of innovative intervention strategies for the prevention and treatment of complex diseases that involve the manipulation of the biochemical state of genes or the dynamics of signal transduction.

6. **Discussions and Conclusions.** In this study, we present a new theoretical approach to describe the controllability of networks using a novel concept of domain control and demonstrate how to drive a complex networked system transition from one domain of attraction to another. Such problems are of practical relevance and importance, particularly in the identification of potential disease genes and the development of therapeutic strategies for complex diseases. Based on our theory, we have proposed an algorithm of identifying the driver nodes for realizing domain controllability. More importantly, we have designed a general form of control functions, which can be used to uncover the practically required driver nodes for inducing transition from one domain of attraction to another. The successful application of domain controllability analysis to three real disease-related networks demonstrates the effectiveness of our approach, which provides a powerful tool for finding the critical nodes (genes) that contribute to phenotype transitions of systems. Recent life sciences are aiming at the control of biological systems for medical purposes. Such problems may be solved if we can identify the key driver genes and control the activity of these genes. Overall, these results indicate that our proposed domain controllability framework not only offers a practical control scheme but also identifies potential disease genes and drug targets, providing innovative therapeutic approaches, as well as prognostic or diagnostic markers for complex diseases.

To demonstrate the robustness of the control effectiveness, we performed the perturbation analysis for the EMT core network presented in section 5.1. The parameter values in (25) are perturbed within 10% variations for 100 times. Figure C9 shows the evolution of the control function $u(t)$ and the control error $e(t)$ of each perturbation across 100 disturbed parameter sets. As shown in Fig. C9, there are almost no changes of $u(t)$ and $e(t)$ with the perturbations of the parameters. In

other words, the control effectiveness is robust to perturbations of the parameter values.

Despite these exciting results regarding the control of the nonlinear dynamics of complex networks, domain control frameworks can still meet some difficulties. For example, estimating the domains of attraction is an important and challenging problem in higher-dimensional nonlinear systems [68]. Although many techniques have been proposed to address this problem, these are mainly based on the elements of Lyapunov stability theory [4, 37]. Unfortunately, Lyapunov functions are difficult to determine using computational and analytical methods [29]. In this study, we employed a sampling method to estimate the domains of attraction, which was proven to efficiently address the domain controllability problem, as demonstrated through its applications to several complex systems.

To identify the driver nodes for achieving domain controllability, we employed the method presented in the proof of Theorem 3.1 to obtain theoretically required driver nodes, including two index sets K and Q. The set Q is similar to a minimal feedback vertex set of a directed graph in graph theory, which has been studied in previous literatures [17, 39]. However, one of the remarkable aspects is that we designed a general form of control functions for realizing domain controllability. It is of great importance because it tells us the proper control input that we should impose on the driver nodes in practice. Therefore, we also performed simulations to obtain practically required driver nodes. Undoubtedly, these methods may take too much time to determine the set Q for the high-dimension problems. Because this study focused on the theory of domain controllability and its validation using the numerical simulations, designing effective algorithms for high-dimension problems would be our further work. In addition, nonlinear systems often display different types of attractors, such as stable steady states (fixed points), periodic attractors and strange attractors. In this paper, we mainly studied the transition between different stable steady states which have been observed in complex disease systems. However, the proposed approaches should also be extended to include other types of attractors which need for further examination.

In summary, we developed a practical and efficient framework for the domain controllability of nonlinear dynamical networks based on theoretical analysis and successfully applied to three disease-related complex networks. The principles outlined here can be used to guide pharmacological design and optimize drug combination strategies targeting complex diseases. Furthermore, the phenomenon of multistability has observed in many important biological processes such as cell cycle and cell fate decision. Therefore, our approaches are of wide applications in the biological and medical fields, as explained in section 5.4. In addition, the approaches presented here can also be extended to the control of related problems in other fields such as engineering and economy only if the dynamical system can exhibit multistable phenomenon. Our methodological framework may lead to new insights into how the behavior of complex systems can be altered in a desired manner.

Appendix A. **Three biological examples for illustrating the definitions of the sets K and Q.** In this section, we first give two simple but famous biological examples to explain how to obtain the set K and show the biological significance of the selected set K.

The first example is a system without any self-feedbacks. The well-known bistable gene-regulatory network, i.e., the genetic toggle switch in Escherichia coli, has been

constructed and verified experimentally [18]. The regulatory network is shown in Fig. A1.
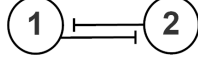

FIGURE A1. The gene-regulatory network of the genetic toggle switch.

The dynamical model of this system is presented as follows.

$$\begin{aligned} \frac{dx_1}{dt} &= \frac{\alpha_1}{1+x_2^\beta} - x_1, \\ \frac{dx_2}{dt} &= \frac{\alpha_2}{1+x_1^\gamma} - x_2, \end{aligned} \tag{A1}$$

where x_1, x_2 represent the concentrations of repressors 1 and 2, respectively. α_1 and α_2 indicate the effective rate of synthesis of repressors 1 and 2, respectively. β and γ are the cooperativity of repression of promoters 2 and 1, respectively.

From the system (A1), it is easy to obtain that $\frac{\partial F_1}{\partial x_1} = -1 < 0, \frac{\partial F_2}{\partial x_2} = -1 < 0$. According to the definition of the set K, the set K for this system is empty.

The second example is a gene regulatory circuit with a self-feedback, which includes two genes. The regulatory network is shown in Fig. A2.
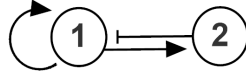
FIGURE A2. A two-gene regulatory circuit with a self-feedback.

The dynamical model of this system is presented as follows (Please see the details of the model in the previous study [67]).

$$\begin{aligned} \frac{dx_1}{dt} &= \alpha_1 \frac{\beta_1 + x_1^n}{(K_1^n + x_1^n)(1 + (x_2/K_2)^n)} - \lambda_1 x_1, \\ \frac{dx_2}{dt} &= \alpha_2 \frac{\beta_2 + x_1^n}{K_3^n + x_1^n} - \lambda_2 x_2, \end{aligned} \tag{A2}$$

where x_1, x_2 represent the expression levels of genes 1 and 2, respectively. α_1 and α_2 are the maximum expression rate. β_1 and β_2 indicate the inverse of the fold-change of enhanced transcription rates when the promoter is occupied by the activator. K_1 and K_2 are Michaelis constants and n is the Hill coefficient. λ_1 and λ_2 are the degradation rates.

According to the system (A2), we have $\frac{\partial F_2}{\partial x_2} = -\lambda_2, \frac{\partial F_1}{\partial x_1} = \frac{\alpha_1}{1+(x_2/K_2)^n} \frac{nx_1^{n-1}(K_1^n - \beta_1)}{(K_1^n + x_1^n)^2} - \lambda_1$. Because λ_i are degradation rates, thus $\lambda_i > 0 (i = 1, 2)$. Based on the biological significance of the parameters, we have $\beta_i = \rho_i K_i^n (\rho_i \in [10^{-3}, 10^{-2}], i = 1, 2)$ [67]. Thus, it is easy to see that $\frac{\partial F_1}{\partial x_1} \geq 0$ is not constantly negative. Thus $K = \{1\}$.

The main idea of determining the set Q is that a set of variables are selected such that the dynamical equations of all of the remaining variables can be expressed as explicit functions of these selected variables by successive evaluation of their dynamical equations. To clearly show the relationships between the controlled variables and the remaining variables, we give a schematic diagram in Fig. A3.
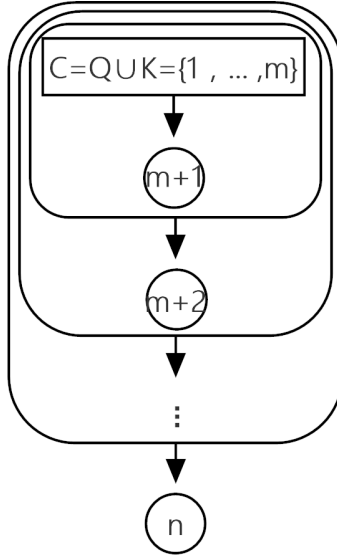
FIGURE A3. A schematic diagram of the relationships between the controlled variables (nodes) and the remaining variables. We relabel the indexes of variables of the system such that the control set $C = \{1, ..., m\}$. Then the $(m+1)$-th equation only contains the variables in the control set C. The $(m+2)$-th equation only contains the variables in the union of the control set C and $(m+1)$-th node. By this analogy, the (n)-th equation contains the variables in the set $C \bigcup \{m+1, \cdots, n-1\}$.

To further illustrate the definition of the set Q, we give two simple examples.

Example 1. We also consider the system of the genetic toggle switch and its corresponding dynamical equations in (A1). According to (A1), the second equation only includes the first variable x_1 except for itself x_2, namely $N(2) = \{1\}$, where $N(i)$ is the index set of variables which are in the i-th equations of the system but exclude the index i. In other words, the second equation is an explicit function of x_1. Obviously, $N(2) \subseteq \{1\}$. Because $K = \emptyset$, then we can choose the set $Q = \{1\}$. Accordingly, $C = Q \bigcup K = \{1\}$. Therefore, (7) holds. Actually, the set Q is not unique. Since $N(1) = \{2\}$, i.e., the first equation is also an explicit function of x_2, Similar to the above method, the set Q can also be set as $\{2\}$.

Example 2. The famous Goodwin model [15]. The end-product of a metabolic pathway inhibits the expression of a gene coding for an enzyme that catalyzes a reaction step in the pathway, which is shown in Fig. A4.

Its dynamics is governed by three ordinary differential equations.

$$\begin{aligned}
\dot{x}_1 &= k_1 r(x_3) - d_1 x_1, \\
\dot{x}_2 &= k_2 x_1 - d_2 x_2, \\
\dot{x}_3 &= k_3 x_2 - d_3 x_3,
\end{aligned} \tag{A3}$$

where x_1, x_2, x_3 represent the concentrations of a gene mRNA, an enzyme and a metabolite, respectively. k_i is production constant, d_i is degradation constant, and $r : \mathbb{R} \to \mathbb{R}$ is a decreasing, nonlinear function.
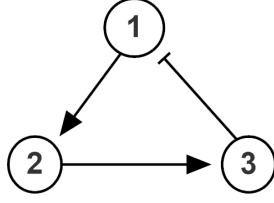
FIGURE A4. The Goodwin model for a genetic regulatory system involving end-product inhibition.

According to ($A3$), the second equation only includes the first variable x_1 except for itself x_2, that is, $N(2) = \{1\}$. Similarly, we have $N(3) = \{2\}$. By successive evaluation of the dynamical equations, the third equation is also an explicit function of x_1. Thus, we can select the set Q to be $\{1\}$ and $C = Q \bigcup K = \{1\}$. It can easily be shown that $N(2) \subseteq C, N(3) \subseteq C \bigcup \{2\}$. In fact, the set Q is also not unique. It is easy to verify that the set Q can also be set as $\{2\}$ or $\{3\}$.

Appendix B. Mathematical models of the Th differentiation cellular network and the cancer network.

B.1. **Mathematical model of the Th differentiation cellular network.** In the literature [16], the steady states of the Th differentiation cellular network were calculated using the SQUAD program, a software for the dynamic simulation of signaling networks. SQUAD converts the network into a discrete dynamical model and then creates a continuous dynamical model to identify all the steady states of the system. However, the automatically transformed continuous model is extremely complicated. In our study, we constructed a simplified ODE model by combining these existed models and the law of mass action. The dynamics of the Th differentiation cellular network is quantified by the following differential equations.

$$\dot{x}_1 = \frac{k_{1-1}x_1 + k_{1-21}x_{21}}{1 + k_{1-1}x_1 + k_{1-21}x_{21} + k_{1-22}x_{22}} - d_1 x_1,$$

$$\dot{x}_2 = -d_2 x_2,$$

$$\dot{x}_3 = k_{3-2}x_2 - d_3 x_3,$$

$$\dot{x}_4 = \frac{k_{4-14}x_{14} + k_{4-16}x_{16} + k_{4-20}x_{20} + k_{4-22}x_{22}}{1 + k_{4-14}x_{14} + k_{4-16}x_{16} + k_{4-20}x_{20} + k_{4-22}x_{22} + k_{4-19}x_{19}} - d_4 x_4,$$

$$\dot{x}_5 = k_{5-4}x_4 - d_5 x_5,$$

$$\dot{x}_6 = k_{6-1}x_1 - d_6 x_6,$$

$$\dot{x}_7 = k_{7-6}x_6 - d_7 x_7,$$

$$\dot{x}_8 = -d_8 x_8, \tag{B1}$$

$$\dot{x}_9 = \frac{k_{9-8}x_8{}^2}{1 + k_{9-8}x_8{}^2 + k_{9-21}x_{21}{}^2} - d_9 x_9,$$

$$\dot{x}_{10} = -d_{10}x_{10},$$

$$\dot{x}_{11} = \frac{k_{11-10}x_{10}{}^2}{1 + k_{11-10}x_{10}{}^2 + k_{11-21}x_{21}{}^2} - d_{11}x_{11},$$

$$\dot{x}_{12} = \frac{k_{12-1}x_1{}^2}{1 + k_{12-1}x_1{}^2 + k_{12-18}x_{18}{}^2} - d_{12}x_{12},$$

$$\dot{x}_{13} = \frac{k_{13-12}x_{12}{}^2}{1 + k_{13-12}x_{12}{}^2 + k_{13-17}x_{17}{}^2} - d_{13}x_{13},$$

$$\dot{x}_{14} = k_{14-11}x_{11} - d_{14}x_{14},$$

$$\dot{x}_{15} = \frac{k_{15-5}x_5{}^2}{1 + k_{15-5}x_5{}^2 + k_{15-17}x_{17}{}^2} - d_{15}x_{15},$$

$$\dot{x}_{16} = k_{16-23}x_{23} - d_{16}x_{16},$$

$$\dot{x}_{17} = k_{17-18}x_{18} + k_{17-22}x_{22} - d_{17}x_{17},$$

$$\dot{x}_{18} = k_{18-3}x_3 + k_{18-15}x_{15} - d_{18}x_{18},$$

$$\dot{x}_{19} = k_{19-7}x_7 - d_{19}x_{19},$$

$$\dot{x}_{20} = k_{20-9}x_9 - d_{20}x_{20},$$

$$\dot{x}_{21} = k_{21-13}x_{13} - d_{21}x_{21},$$

$$\dot{x}_{22} = \frac{k_{22-18}x_{18} + k_{22-22}x_{22}}{1 + k_{22-1}x_1 + k_{22-22}x_{22} + k_{22-18}x_{18}} - d_{22}x_{22},$$

$$\dot{x}_{23} = -d_{23}x_{23},$$

where $x_i(i = 1, 2, \cdots, 23)$ represents the activity level of GATA-3, IFN-β, IFN-βR, IFN-γ, IFN-γR, IL-10, IL-10R, IL-12, IL-12R, IL-18, IL-18R, IL-4, IL-4R, IRAK, JAK1, NFAT, SOCS1, STAT1, STAT3, STAT4, STAT6, T-bet and TCR, respectively. The parameters used to perform numerical simulations are: $d_1 = 0.4, d_2 = 1, d_3 = 1, d_4 = 0.3, d_5 = 1, d_6 = 1, d_7 = 1, d_8 = 1, d_9 = 1, d_{10} = 1, d_{11} = 1, d_{12} = 0.4, d_{13} = 0.6, d_{14} = 1, d_{15} = 1, d_{16} = 1, d_{17} = 1, d_{18} = 1, d_{19} = 1, d_{20} = 1, d_{21} = 0.6, d_{22} = 0.4, d_{23} = 1, k_{i-j} = 1 (i, j = 1, 2, \cdots, 22).$

B.2. **Mathematical model of the cancer network.** The ODEs describing the dynamics of the cancer network are formulated as follows [32].

$$\dot{x}_1 = \frac{2bS^n}{S^n + s_{32}x_{32}{}^n} - kx_1,$$

$$\dot{x}_2 = \frac{a\left(s_1x_1{}^n + s_9x_9{}^n + s_{14}x_{14}{}^n\right)/3}{S^n + 1/3s_1x_1{}^n + 1/3s_9x_9{}^n + 1/3s_{14}x_{14}{}^n} + \frac{2bS^n}{2S^n + s_{16}x_{16}{}^n + s_{31}x_{31}{}^n} - kx_2,$$

$$\dot{x}_3 = \frac{a\left(s_2x_2{}^n + s_8x_8{}^n + s_{29}x_{29}{}^n\right)/3}{S^n + 1/3s_2x_2{}^n + 1/3s_8x_8{}^n + 1/3s_{29}x_{29}{}^n} + \frac{3bS^n}{3S^n + s_9x_9{}^n + s_{10}x_{10}{}^n + s_{16}x_{16}{}^n} - kx_3,$$

$$\dot{x}_4 = \frac{as_2x_2{}^n}{S^n + s_2x_2{}^n} + \frac{bS^n}{S^n + s_{25}x_{25}{}^n} - kx_4,$$

$$\dot{x}_5 = \frac{as_{13}x_{13}{}^n}{S^n + s_{13}x_{13}{}^n} + \frac{bS^n}{S^n + 1/3s_{14}x_{14}{}^n + 1/3s_{17}x_{17}{}^n + 1/3s_{19}x_{19}{}^n} - kx_5,$$

$$\dot{x}_6 = \frac{2bS^n}{S^n + 1/2s_{17}x_{17}{}^n + 1/2s_{18}x_{18}{}^n} - kx_6,$$

$$\dot{x}_7 = \frac{as_{20}x_{20}{}^n}{S^n + s_{20}x_{20}{}^n} + \frac{bS^n}{S^n + 1/2s_{15}x_{15}{}^n + 1/2s_{32}x_{32}{}^n} - kx_7,$$

$$\dot{x}_8 = \frac{2bS^n}{S^n + 1/3s_4x_4{}^n + 1/3s_{10}x_{10}{}^n + 1/3s_{16}x_{16}{}^n} - kx_8,$$

$$\dot{x}_9 = \frac{a\left(s_{11}x_{11}{}^n + s_{13}x_{13}{}^n + s_{25}x_{25}{}^n + s_{26}x_{26}{}^n\right)/4}{S^n + 1/4s_{11}x_{11}{}^n + 1/4s_{13}x_{13}{}^n + 1/4s_{25}x_{25}{}^n + 1/4s_{26}x_{26}{}^n} + \frac{bS^n}{S^n + s_{29}x_{29}{}^n} - kx_9,$$

$$\dot{x}_{10} = \frac{a\left(s_{11}x_{11}{}^n + s_{12}x_{12}{}^n + s_{13}x_{13}{}^n\right)/3}{S^n + 1/3s_{11}x_{11}{}^n + 1/3s_{12}x_{12}{}^n + 1/3s_{13}x_{13}{}^n} + \frac{bS^n}{S^n + s_4x_4{}^n} - kx_{10},$$

$$\dot{x}_{11} = \frac{a\left(s_{10}x_{10}{}^n + s_{13}x_{13}{}^n + s_{14}x_{14}{}^n + s_{27}x_{27}{}^n + s_{28}x_{28}{}^n\right)}{5S^n + s_{10}x_{10}{}^n + s_{13}x_{13}{}^n + s_{14}x_{14}{}^n + s_{27}x_{27}{}^n + s_{28}x_{28}{}^n} + \frac{bS^n}{S^n + s_{29}x_{29}{}^n} - kx_{11},$$

$$\dot{x}_{12} = \frac{a\left(s_9x_9{}^n + s_{10}x_{10}{}^n + s_{11}x_{11}{}^n + s_{13}x_{13}{}^n + s_{14}x_{14}{}^n + s_{15}x_{15}{}^n\right)}{6S^n + s_9x_9{}^n + s_{10}x_{10}{}^n + s_{11}x_{11}{}^n + s_{13}x_{13}{}^n + s_{14}x_{14}{}^n + s_{15}x_{15}{}^n} + \frac{bS^n}{S^n + s_2x_2{}^n}$$
$$- kx_{12},$$

(B2)

$$\dot{x}_{13} = \frac{2a\left(s_{12}x_{12}{}^n + s_{14}x_{14}{}^n + s_{28}x_{28}{}^n\right)/3}{S^n + 1/3s_{12}x_{12}{}^n + 1/3s_{14}x_{14}{}^n + 1/3s_{28}x_{28}{}^n} - kx_{13},$$

$$\dot{x}_{14} = \frac{a\left(s_{10}x_{10}{}^n + s_{13}x_{13}{}^n + s_{25}x_{25}{}^n + s_{26}x_{26}{}^n + s_{28}x_{28}{}^n\right)}{5S^n + s_{10}x_{10}{}^n + s_{13}x_{13}{}^n + s_{25}x_{25}{}^n + s_{26}x_{26}{}^n + s_{28}x_{28}{}^n} + \frac{bS^n}{S^n + s_2x_2{}^n} - kx_{14},$$

$$\dot{x}_{15} = \frac{as_{14}x_{14}{}^n}{S^n + s_{14}x_{14}{}^n} + \frac{bS^n}{S^n + 1/2s_2x_2{}^n + 1/2s_{29}x_{29}{}^n} - kx_{15},$$

$$\dot{x}_{16} = \frac{a\left(s_2x_2{}^n + s_{10}x_{10}{}^n + s_{26}x_{26}{}^n + s_{32}x_{32}{}^n\right)}{4S^n + s_2x_2{}^n + s_{10}x_{10}{}^n + s_{26}x_{26}{}^n + s_{32}x_{32}{}^n} + \frac{3bS^n}{3S^n + s_1x_1{}^n + s_6x_6{}^n + s_7x_7{}^n} - kx_{16},$$

$$\dot{x}_{17} = \frac{a\left(s_9x_9{}^n + s_{10}x_{10}{}^n + s_{20}x_{20}{}^n\right)}{3S^n + s_9x_9{}^n + s_{10}x_{10}{}^n + s_{20}x_{20}{}^n} + \frac{2bS^n}{2S^n + s_3x_3{}^n + s_6x_6{}^n} - kx_{17},$$

$$\dot{x}_{18} = \frac{a\left(s_9x_9{}^n + s_{25}x_{25}{}^n + s_{26}x_{26}{}^n\right)/3}{S^n + 1/3s_9x_9{}^n + 1/3s_{25}x_{25}{}^n + 1/3s_{26}x_{26}{}^n} + \frac{bS^n}{S^n + s_6x_6{}^n} - kx_{18},$$

$$\dot{x}_{19} = \frac{2bS^n}{S^n + 1/3s_3x_3{}^n + 1/3s_5x_5{}^n + 1/3s_{30}x_{30}{}^n} - kx_{19},$$

$$\dot{x}_{20} = \frac{as_1x_1{}^n}{S^n + s_1x_1{}^n} + \frac{bS^n}{S^n + 1/2s_6x_6{}^n + 1/2s_{17}x_{17}{}^n} - kx_{20},$$

$$\dot{x}_{21} = \frac{a\left(s_{22}x_{22}{}^n + s_{23}x_{23}{}^n\right)/2}{S^n + 1/2s_{22}x_{22}{}^n + 1/2s_{23}x_{23}{}^n} + \frac{bS^n}{S^n + 1/2s_{10}x_{10}{}^n + 1/2s_{24}x_{24}{}^n} - kx_{21},$$

$$\dot{x}_{22} = \frac{a\left(s_2x_2{}^n + s_9x_9{}^n\right)/2}{S^n + 1/2s_2x_2{}^n + 1/2s_9x_9{}^n} + \frac{bS^n}{S^n + s_{24}x_{24}{}^n} - kx_{22},$$

$$\dot{x}_{23} = \frac{2bS^n}{S^n + s_{10}x_{10}{}^n} - kx_{23},$$

$$\dot{x}_{24} = \frac{a\left(s_{12}x_{12}{}^n + s_{13}x_{13}{}^n + s_{25}x_{25}{}^n\right)/3}{S^n + 1/3s_{12}x_{12}{}^n + 1/3s_{13}x_{13}{}^n + 1/3s_{25}x_{25}{}^n} + \frac{bS^n}{S^n + 1/2s_2x_2{}^n + 1/2s_{21}x_{21}{}^n} - kx_{24},$$

$$\dot{x}_{25} = \frac{a\left(s_{10}x_{10}{}^n + s_{13}x_{13}{}^n\right)/2}{S^n + 1/2s_{10}x_{10}{}^n + 1/2s_{13}x_{13}{}^n} + \frac{bS^n}{S^n + s_4x_4{}^n} - kx_{25},$$

$$\dot{x}_{26} = \frac{2a\left(s_{11}x_{11}{}^n + s_{12}x_{12}{}^n + s_{27}x_{27}{}^n\right)/3}{S^n + 1/3s_{11}x_{11}{}^n + 1/3s_{12}x_{12}{}^n + 1/3s_{27}x_{27}{}^n} - kx_{26},$$

$$\dot{x}_{27} = \frac{2as_{28}x_{28}{}^n}{S^n + s_{28}x_{28}{}^n} - kx_{27},$$

$$\dot{x}_{28} = -kx_{28},$$

$$\dot{x}_{29} = \frac{2as_{14}x_{14}{}^n}{S^n + s_{14}x_{14}{}^n} - kx_{29},$$

$$\dot{x}_{30} = \frac{2as_{19}x_{19}{}^n}{S^n + s_{19}x_{19}{}^n} - kx_{30},$$

$$\dot{x}_{31} = \frac{a\left(s_{10}x_{10}{}^n + s_{32}x_{32}{}^n\right)/2}{S^n + 1/2s_{10}x_{10}{}^n + 1/2s_{32}x_{32}{}^n} + \frac{bS^n}{S^n + 1/2s_7x_7{}^n + 1/2s_{16}x_{16}{}^n} - kx_{31},$$

$$\dot{x}_{32} = \frac{2as_2x_2{}^n}{S^n + s_2x_2{}^n} - kx_{32}.$$

In the above equations, S represents the threshold (inflection point) of the explicitly sigmoid functions, i.e. the strength of the regulatory interaction, and n is the Hill coefficient which determines the steepness of the sigmoid function. In addition, k is self-degradation constant, b is repression constant and a is activation constant. s_i indicates the activation or inhibition strength ($i = 1, 2, \cdots, 32$). The default parameter values are: $S = 0.5, n = 4, a = 0.5, b = 0.5, k = 1, s_1 = 1.028, s_2 = 1.281, s_3 = 0.598, s_4 = 0.827, s_5 = 1.02, s_6 = 2.244, s_7 = 1.037, s_8 = 1.507, s_9 = 0.203, s_{10} = 1.299, s_{11} = 1.337, s_{12} = 0.921, s_{13} = 2.044, s_{14} = 0.678, s_{15} = 0.924, s_{16} = 0.467, s_{17} = 0.947, s_{18} = 0.242, s_{19} = 0.941, s_{20} = 0.48, s_{21} = 0.666, s_{22} = 0.232, s_{23} =$

$1.156, s_{24} = 0.951, s_{25} = 0.849, s_{26} = 0.698, s_{27} = 1.288, s_{28} = 0.966, s_{29} = 1.419, s_{30} = 0.401, s_{31} = 0.293, s_{32} = 1.479.$

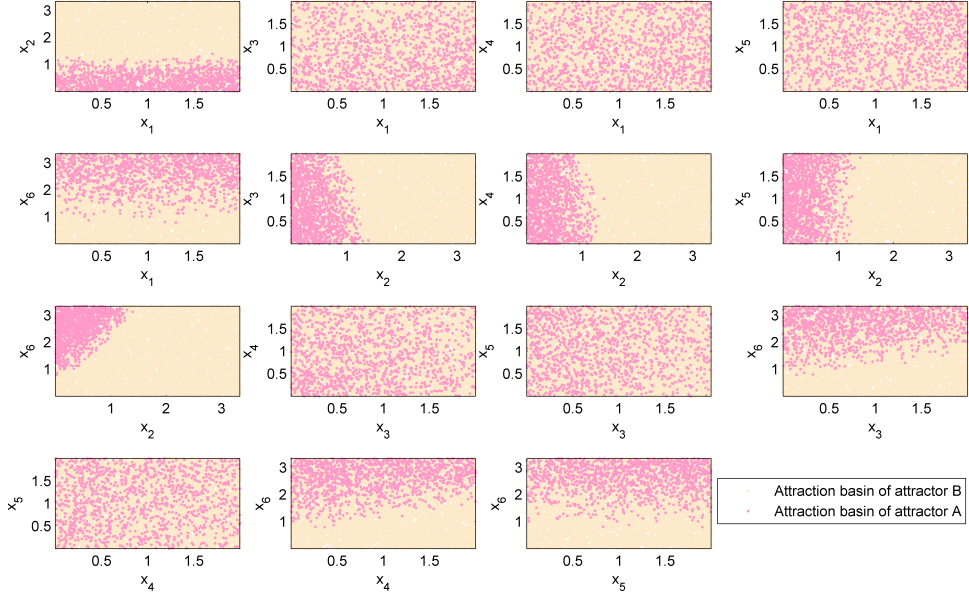**Appendix** C. **Appendix figures.**

FIGURE C1. Distributions of the domains of attraction of the EMT core network model. We project the domains into a two-dimensional plane. The light red and light yellow regions represent the domains of attraction of the epithelial attractor and mesenchymal attractor, respectively.
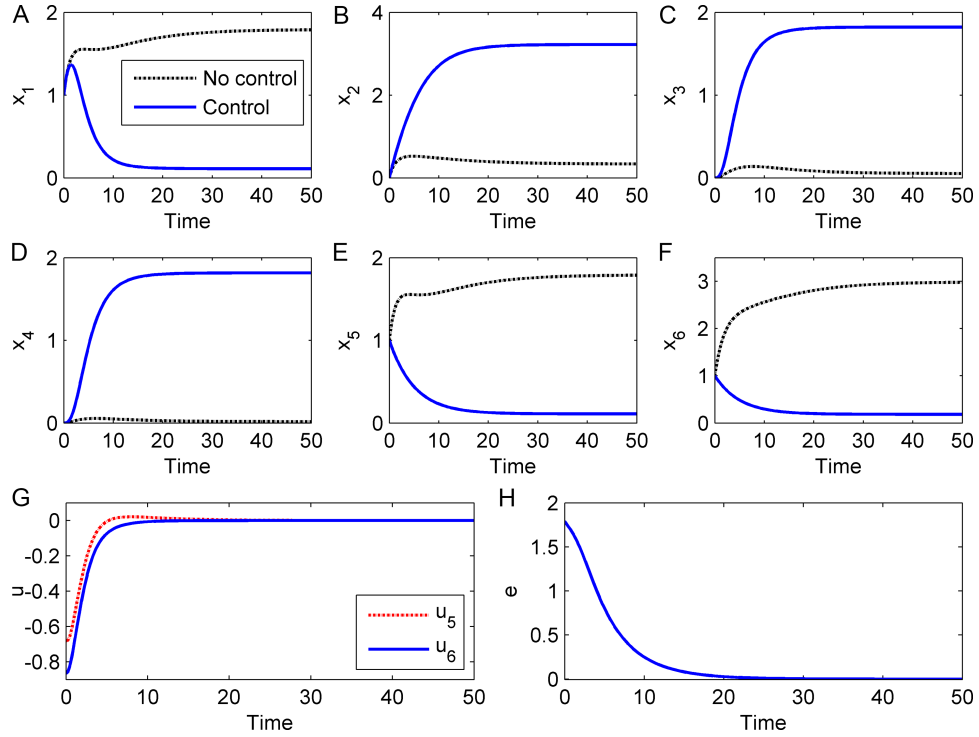
FIGURE C2. Domain control of the EMT core network transition from the epithelial state to the mesenchymal state by the theoretically required driver nodes. The node state trajectories $x(t)$ (**A-F**), control function $u(t)$ (**G**) and control error $e(t)$ (**H**) are depicted when the theoretically required driver nodes (i.e., miR-200 and miR-203) are controlled. The symbols x_1, x_2, x_3, x_4, x_5 and x_6 indicate the activities of CDH1, SNAI1, ZEB1, ZEB2, miR-200 and miR-203, respectively. u_5 and u_6 indicate the control functions for miR-200 and miR-203, respectively. The black dashed and blue solid lines represent the initial epithelial state (i.e., the undesired state with no control) and desired mesenchymal state, respectively.
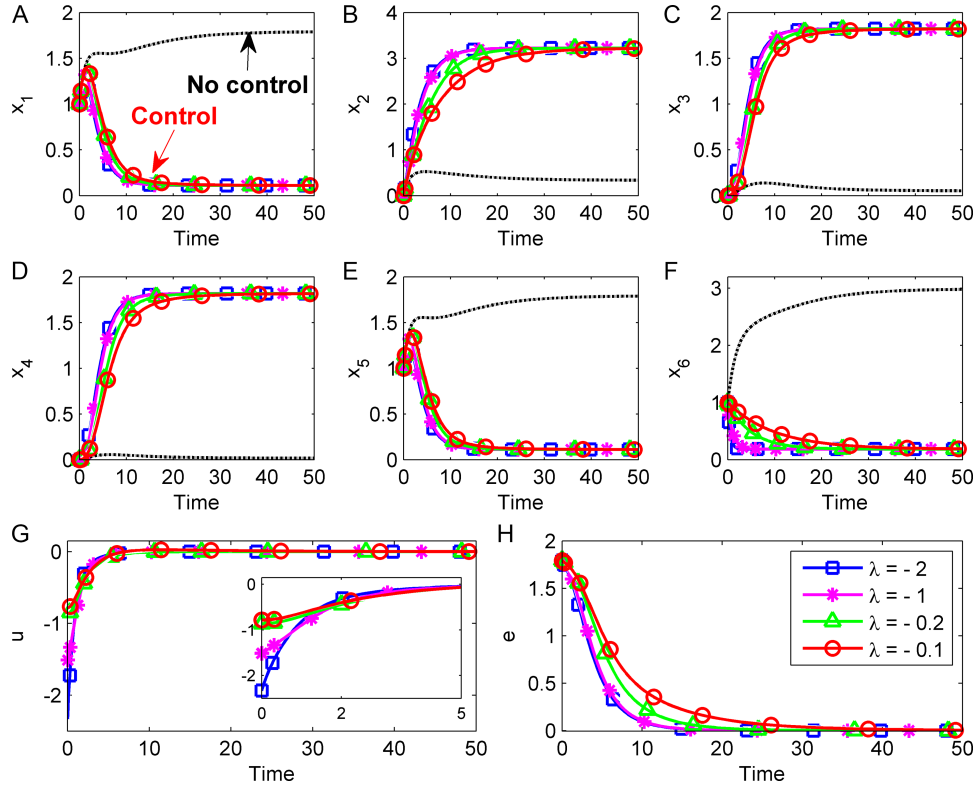
FIGURE C3. Effect of the parameter λ in the designed control function on the domain control of the EMT core network. The node state trajectories $x(t)$ (**A-F**), control function $u(t)$ (**G**) and control error $e(t)$ (**H**) are depicted when the practically required driver node, namely miR-203, is controlled. The symbols x_1, x_2, x_3, x_4, x_5 and x_6 indicate the activities of CDH1, SNAI1, ZEB1, ZEB2, miR-200 and miR-203, respectively. The black dashed and colored solid lines represent the initial epithelial state (i.e., the undesired state with no control) and desired mesenchymal state, respectively. Here, we take the EMT core network as an example to show how the parameter λ in our designed control function influences the control process. As shown in (H), to a certain extent, a higher absolute value of indicates a less time the system takes to evolve to the desirable attractor. However, there is no significant influence when the absolute value of λ is larger than 1.
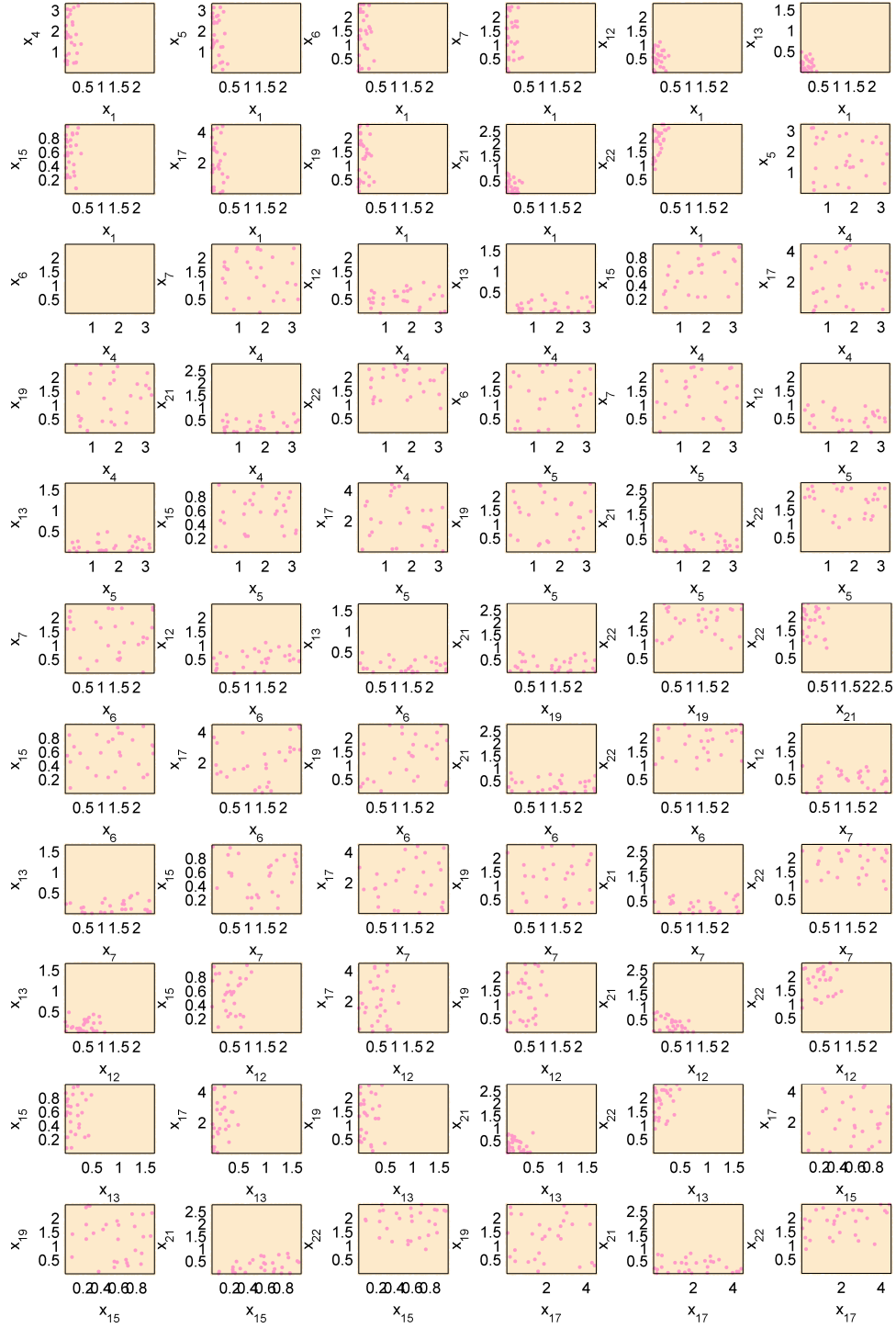
FIGURE C4. Distributions of the domains of attraction of the Th differentiation cellular network model. We project the domains into a two-dimensional plane. The light red and light yellow regions represent the domains of attraction of the Th1 attractor and Th2 attractor, respectively.
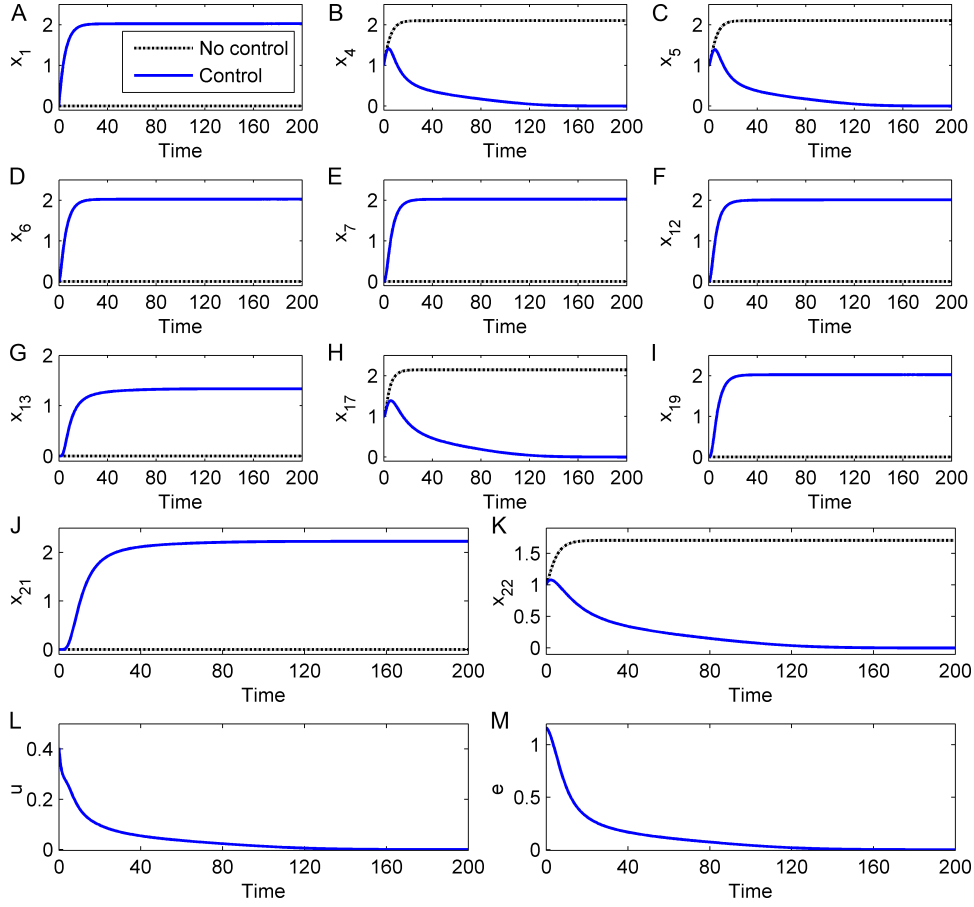
FIGURE C5. Domain control of the T helper differentiation cellular network transition from Th1 to Th2. The representative node state trajectories $x(t)$ (**A-K**), control function $u(t)$ (**L**) and control error $e(t)$ (**M**) are depicted. The symbols $x_1, x_4, x_5, x_6, x_7, x_{12}, x_{13}, x_{17}, x_{19}, x_{21}$ and x_{22} indicate the activities of GATA-3, IFN-γ, IFN-γR, IL-10, IL-10R, IL-4, IL-4R, SOCS1, STAT3, STAT6 and T-bet, respectively. The dashed black line indicates the Th1 phenotype, i.e. the initial state with no control. The solid blue line represents the Th2 phenotype. The driver node is GATA-3 when Th1 is driven to the Th2 phenotype.

FIGURE C6. Distributions of the domains of attraction of the cancer network. We project the domains into a two-dimensional plane. The light red and light yellow regions represent the domains of attraction of the cancer attractor and normal attractor, respectively.
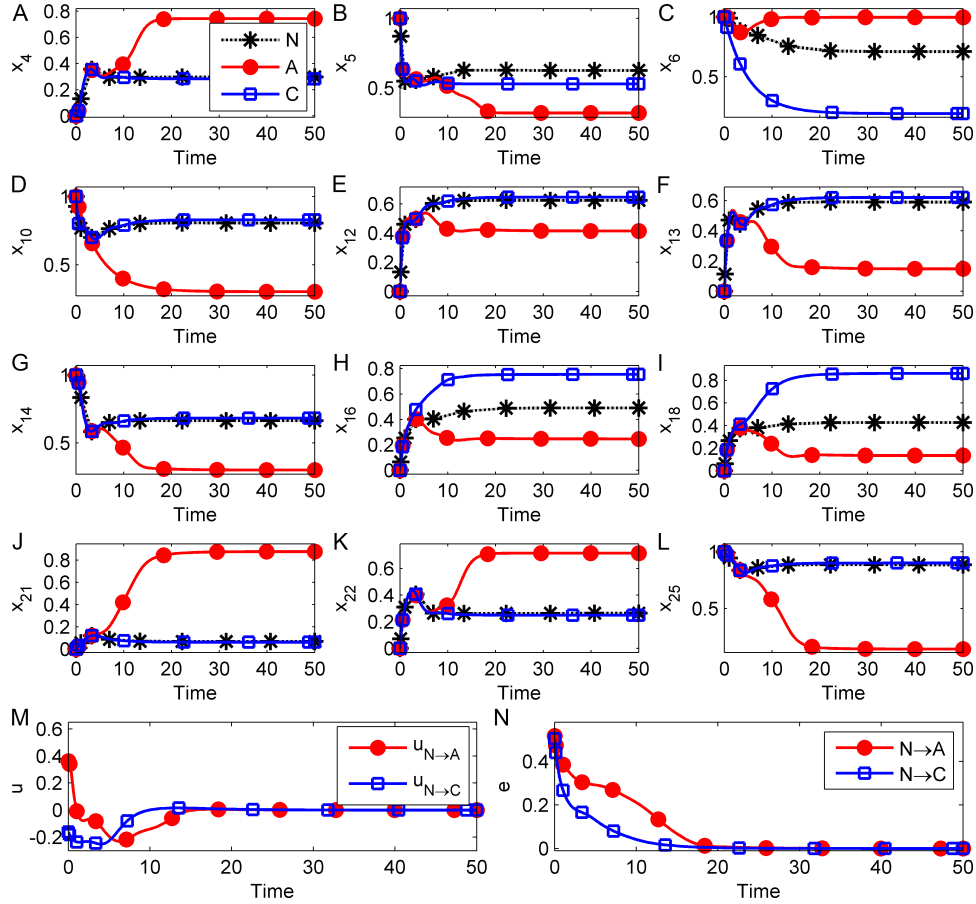
FIGURE C7. Domain control of the cancer network transitions from the normal state to the apoptosis state and cancer state. The representative node state trajectories $x(t)$ **(A-L)**, control function $u(t)$ **(M)** and control error $e(t)$ **(N)** are depicted when AKT $(u_{N\to A})$ and RB $(u_{N\to C})$ are controlled, respectively. The symbols $x_4, x_5, x_6, x_{10}, x_{12}, x_{13}, x_{14}, x_{16}, x_{18}, x_{21}, x_{22}$ and x_{25} indicate the activities of PTEN, CDH1, RB, AKT, VEGF, HGF, HIF1, MDM2, CDK4, Caspase, BAX and NFκB, respectively. The black dashed lines with star markers indicate the normal state, i.e., the initial state with no control. The red solid lines with filled circles and blue solid lines with squares represent transitions from normal (N) to apoptosis (A) (denoted by N$\to$A) and normal (A) to cancer (C) (denoted by N$\to$C), respectively. The driver node is AKT when normal is steered to the apoptosis and RB when normal is driven to the cancer state.
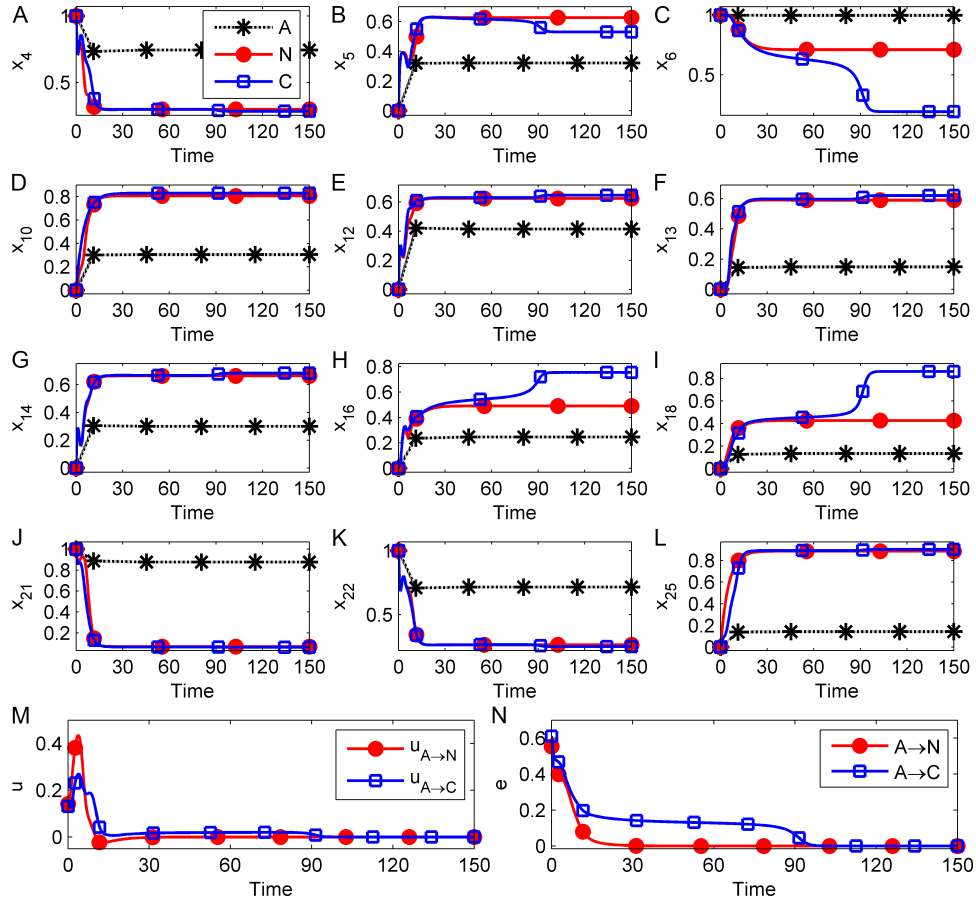
FIGURE C8. Domain control of the cancer network transitions from the apoptosis state to the normal state and cancer state. The black dashed lines with star markers indicate the apoptosis state, i.e., the undesired state with no control. The red solid lines with filled circles and blue solid lines with squares represent transitions from the apoptosis state (A) to normal state (N) (denoted by A→N) and apoptosis state (A) to cancer (C) (denoted by A→C), respectively. The driver node is NFκB when the apoptosis state is steered to the normal state and AKT when apoptosis state is driven to the cancer state.
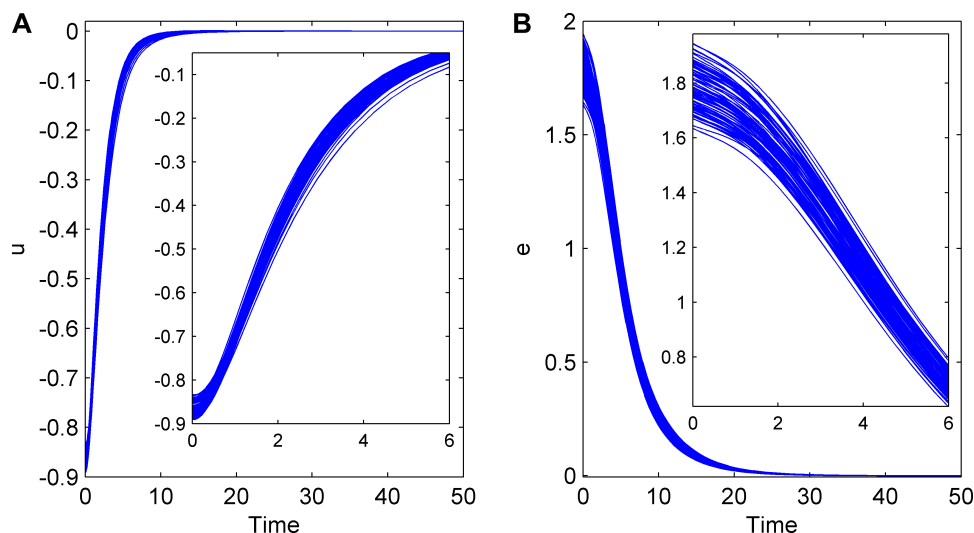
FIGURE C9. Robustness analysis of the control effectiveness a-
gainst parameter perturbation. The evolution of the control func-
tion $u(t)$ and control error $e(t)$ are depicted when miR-203 is con-
trolled. Each curve represents one simulation for one perturbation
of the parameter values. In our numerical simulations, we random-
ly generate 100 parameter sets in which every parameter of the
system is perturbed within a range of $\pm 10\%$. The initial state is
the epithelial state, which is set to be $(1, 0, 0, 0, 1, 1)^T$ in all the
numerical experiments. We zoom into the curves in a small win-
dow.

REFERENCES

[1] P. Ao, D. Galas, L. Hood and X. M. Zhu, *Cancer as robust intrinsic state of endogenous molecular-cellular network shaped by evolution*, Med. Hypotheses, **70** (2008), 678–684.

[2] R. P. Araujo, L. A. Liotta and E. F. Petricoin, *Proteins, drug targets and the mechanisms they control: the simple truth about complex networks*, Nat. Rev. Drug Discov., **6** (2007), 871–880.

[3] K. Axelrod, A. Sanchez and J. Gore, *Phenotypic states become increasingly sensitive to per-turbations near a bifurcation in a synthetic gene network*, eLife, **4** (2015), e07935.

[4] S. Balint, E. Kaslik, A. M. Balint and A. Grigis, *Methods for determination and approxima-tion of the domain of attraction in the case of autonomous discrete dynamical systems*, Adv. Differ. Equ., **2006** (2006), 1–15.

[5] Y. Bar-Yam, D. Harmon and B. de Bivort, *Attractors and democratic dynamics*, Science, **323** (2009), 1016–1017.

[6] A.-L. Barabási, N. Gulbahce and J. Loscalzo, *Network medicine: a network-based approach to human disease*, Nat. Rev. Genet., **12** (2011), 56–68.

[7] B. Barzel and A.-L. Barabási, *Universality in network dynamics*, *Nat. Phys.*, **9** (2013), 673–681.

[8] Y. Ben-Neriah and M. Karin, *Inflammation meets cancer, with nf-κb as the matchmaker*, *Nat.Immunol.*, **12** (2011), 715–723.

[9] G. Chen and X. Yu, Chaos Control: Theory and Applications, Springer-Verlag Berlin Heidelberg, Berlin, 2003.

[10] Y.-Z. Chen, L. Wang, W. Wang and Y.-C. Lai, The paradox of controlling complex networks: control inputs versus energy requirement, preprint, `arXiv:1509.03196`.

[11] S. P. Cornelius, W. L. Kath and A. E. Motter, *Realistic control of network dynamics*, *Nat. Commun.*, **4** (2013), 1942.

[12] P. Creixell, E. M. Schoof, J. T. Erler and R. Linding, *Navigating cancer network attractors for tumor-specific therapy*, *Nat. Biotechnol.*, **30** (2012), 842–8.

[13] P. Csermely, T. Korcsmáros, H. J. Kiss, G. London and R. Nussinov, *Structure and dynamics of molecular networks: a novel paradigm of drug discovery: a comprehensive review*, *Pharmacol. & Therapeut.*, **138** (2013), 333–408.

[14] B. De Craene and G. Berx, *Regulatory networks defining EMT during cancer initiation and progression*, *Nat. Rev. Cancer*, **13** (2013), 97–110.

[15] H. De Jong, *Modeling and simulation of genetic regulatory systems: a literature review*, *J. Comput. Biol.*, **9** (2002), 67–103.

[16] A. Di Cara, A. Garg, G. De Micheli, I. Xenarios and L. Mendoza, *Dynamic simulation of regulatory networks using SQUAD*, *BMC Bioinformatics*, **8** (2007), 462.

[17] B. Fiedler, A. Mochizuki, G. Kurosawa and D. Saito, *Dynamics and control at feedback vertex sets. I: informative and determining nodes in regulatory networks*, *J. Dyn. Differ. Equ.*, **25** (2013), 563–604.

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

[19] J. Gao, Y.-Y. Liu, R. M. D'Souza and A.-L. Barabási, *Target control of complex networks*, *Nat. Commun.*, **5** (2014), 5415.

[20] B. T. Hennessy, D. L. Smith, P. T. Ram, Y. Lu and G. B. Mills, *Exploiting the PI3K/AKT pathway for cancer drug discovery*, *Nat. Rev. Drug Discov.*, **4** (2005), 988–1004.

[21] S. Huang, I. Ernberg and S. Kauffman, *Cancer attractors: a systems view of tumors from a gene network dynamics and developmental perspective*, *Semin. Cell Dev. Biol.*, **20** (2009), 869–76.

[22] E. S. Hwang, S. J. Szabo, P. L. Schwartzberg and L. H. Glimcher, *T helper cell fate specified by kinase-mediated interaction of T-bet with GATA-3*, *Science*, **307** (2005), 430–433.

[23] R. G. Jenner, M. J. Townsend, I. Jackson, K. Sun, R. D. Bouwman, R. A. Young, L. H. Glimcher and G. M. Lord, *The transcription factors T-bet and GATA-3 control alternative pathways of T-cell differentiation through a shared set of target genes*, *Proc. Natl. Acad. Sci. U. S. A.*, **106** (2009), 17876–17881.

[24] S. Jin, Y. Li, R. Pan and X. Zou, *Characterizing and controlling the inflammatory network during influenza a virus infection*, *Sci. Rep.*, **4** (2014), 3799.

[25] S. Jin, L. Niu, G. Wang and X. Zou, *Mathematical modeling and nonlinear dynamical analysis of cell growth in response to antibiotics*, *Int. J. Bifurcat. Chaos*, **25** (2015), 1540007.

[26] J. D. Jordan, E. M. Landau and R. Iyengar, *Signaling networks: the origins of cellular multitasking*, *Cell*, **103** (2000), 193–200.

[27] R. Kalluri and R. A. Weinberg, *The basics of epithelial-mesenchymal transition*, *J. Clin. Invest.*, **119** (2009), 1420–1428.

[28] S. Karl and T. Dandekar, *Convergence behaviour and control in non-linear biological networks*, *Sci. Rep.*, **5** (2015), 9746.

[29] H. K. Khalil, Nonlinear Systems, 3rd edition, Prentice Hall, New Jersey, 2002.

[30] Y. C. Lai, *Controlling complex, non-linear dynamical networks*, *National Science Review*, **1** (2014), 339–341.

[31] S. Lamouille, J. Xu and R. Derynck, *Molecular mechanisms of epithelial–mesenchymal transition*, *Nat. Rev. Mol. Cell Biol.*, **15** (2014), 178.

[32] C. Li and J. Wang, *Quantifying the underlying landscape and paths of cancer*, *J. R. Soc. Interface*, **11** (2014), 20140774.

[33] Y. Li, S. Jin, L. Lei, Z. Pan and X. Zou, *Deciphering deterioration mechanisms of complex diseases based on the construction of dynamic networks and systems analysis*, *Sci. Rep.*, **5** (2015), 9283.

[34] Y. Li, M. Yi and X. Zou, *The linear interplay of intrinsic and extrinsic noises ensures a high accuracy of cell fate selection in budding yeast*, *Sci. Rep.*, **4** (2014), 5764.

[35] Y. Y. Liu and A. L. Barabasi, Control principles of complex networks, preprint, `arXiv:1508.05384v2`.

[36] Y. Y. Liu, J. J. Slotine and A. L. Barabasi, *Controllability of complex networks*, *Nature*, **473** (2011), 167–173.

[37] L. G. Matallana, A. M. Blanco and J. A. Bandoni, *Estimation of domains of attraction: A global optimization approach*, *Math. Comput. Model.*, **52** (2010), 574–585.

[38] L. Mendoza, *A network model for the control of the differentiation process in Th cells*, *Biosystems*, **84** (2006), 101–114.

[39] A. Mochizuki, B. Fiedler, G. Kurosawa and D. Saito, *Dynamics and control at feedback vertex sets. II: A faithful monitor to determine the diversity of molecular activities in regulatory networks*, *J. Theor. Biol.*, **335** (2013), 130–146.

[40] M. Moes, A. Le Béchec, I. Crespo, C. Laurini, A. Halavatyi, G. Vetter, A. Del Sol and E. Friederich, *A novel network integrating a miRNA-203/SNAI1 feedback loop which regulates epithelial to mesenchymal transition*, *PloS One*, **7** (2012), e35440.

[41] F.-J. Müller and A. Schuppert, *Few inputs can reprogram biological networks*, *Nature*, **478** (2011), E4–E4.

[42] T. Nepusz and T. Vicsek, *Controlling edge dynamics in complex networks*, *Nat. Phys.*, **8** (2012), 568–573.

[43] J. A. Papin, T. Hunter, B. O. Palsson and S. Subramaniam, *Reconstruction of cellular signalling networks and analysis of their properties*, *Nat. Rev. Mol. Cell Biol.*, **6** (2005), 99–111.

[44] J. Pei, N. Yin, X. Ma and L. Lai, *Systems biology brings new dimensions for structure-based drug design*, *J. Am. Chem. Soc.*, **136** (2014), 11556–11565.

[45] M. Pósfai and P. Hövel, *Structural controllability of temporal networks*, *New J. Phys.*, **16** (2014), 123055.

[46] J. Ruths and D. Ruths, *Control profiles of complex networks*, *Science*, **343** (2014), 1373–1376.

[47] J.-J. E. Slotine and W. Li, Applied Nonlinear Control, Prentice-Hall, New Jersey, 1991.

[48] M. S. Song, L. Salmena and P. P. Pandolfi, *The functions and regulation of the pten tumour suppressor*, *Nat. Rev. Mol. Cell Biol.*, **13** (2012), 283–296.

[49] F. Sorrentino, M. di Bernardo, F. Garofalo and G. R. Chen, *Controllability of complex networks via pinning*, *Phys. Rev. E*, **75** (2007), 046103.

[50] G. Strang, Calculus, Wellesley-Cambridge Press, Massachusetts, 1991.

[51] J. Tan and X. Zou, *Complex dynamical analysis of a coupled network from innate immune responses*, *Int. J. Bifurcat. Chaos*, **23** (2013), 1350180.

[52] J. Tan and X. Zou, *Optimal control strategy for abnormal innate immune response*, *Comput. Math. Method M.*, **2015** (2015), 386235.

[53] A. Vinayagam, T. E. Gibson, H.-J. Lee, B. Yilmazel, C. Roesel, Y. Hu, Y. Kwon, A. Sharma, Y.-Y. Liu, N. Perrimon and A.-L. Barabsi, *Controllability analysis of the directed human protein interaction network identifies disease genes and drug targets*, *Proc. Natl. Acad. Sci. U. S. A.*, **113** (2016), 4976–4981.

[54] D. Wang, S. Jin, F. X. Wu and X. Zou, *Estimation of control energy and control strategies for complex networks*, *Adv. Complex Syst.*, **19** (2015), 1550018.

[55] D. Wang, S. Jin and X. Zou, *Crosstalk between pathways enhances the controllability of signalling networks*, *IET Syst. Biol.*, **10** (2016), 2–9.

[56] L. Z. Wang, R. Q. Su, Z. G. Huang, X. Wang, W. X. Wang, C. Grebogi and Y. C. Lai, *A geometrical approach to control and controllability of nonlinear dynamical networks*, *Nat. Commun.*, **7** (2016), 11323.

[57] Y. Wang, J. Tan, F. Sadre-Marandi, J. Liu and X. Zou, *Mathematical modeling for intracellular transport and binding of HIV-1 gag proteins*, *Math. Biosci.*, **262** (2015), 198–205.

[58] R. Weinberg, *The Biology of Cancer*, 2nd edition, Garland Science, New York, 2013.

[59] U. Wellner, J. Schubert, U. C. Burk, O. Schmalhofer, F. Zhu, A. Sonntag, B. Waldvogel, C. Vannier, D. Darling, A. zur Hausen et al., *The EMT-activator ZEB1 promotes tumorigenicity by repressing stemness-inhibiting micrornas*, *Nat. Cell Biol.*, **11** (2009), 1487–1495.

[60] D. K. Wells, W. L. Kath and A. E. Motter, *Control of stochastic and induced switching in biophysical networks*, *Phys. Rev. X*, **5** (2015), 031036.

[61] A. J. Whalen, S. N. Brennan, T. D. Sauer and S. J. Schiff, *Observability and controllability of nonlinear networks: The role of symmetry*, *Phys. Rev. X*, **5** (2015), 011005.

[62] F.-X. Wu, L. Wu, J. Wang, J. Liu and L. Chen, *Transittability of complex networks and its applications to regulatory biomolecular networks*, *Sci. Rep.*, **4** (2014), 4819.

[63] S. Wuchty, *Controllability in protein interaction networks*, *Proc. Natl. Acad. Sci. U. S. A.*, **111** (2014), 7156–7160.

[64] G. Yan, G. Tsekenis, B. Barzel, J.-J. Slotine, Y.-Y. Liu and A.-L. Barabási, *Spectrum of controlling and observing complex networks*, *Nat. Phys.*, **11** (2015), 779–786.

[65] Z. Yuan, C. Zhao, Z. Di, W.-X. Wang and Y.-C. Lai, *Exact controllability of complex networks*, *Nat. Commun.*, **4** (2013), 2447.

[66] J. G. Zañudo, G. Yang and R. Albert, Structure-based control of complex networks with nonlinear dynamics, `arXiv:1605.08415v2`.

[67] J. Zhang, Z. Yuan, H.X. Li and T. Zhou, *Architecture-dependent robustness and bistability in a class of genetic circuits*, *Biophys. J.*, **99** (2010), 1034–1042.

[68] N. Zhong, *Computational unsolvability of domains of attraction of nonlinear systems*, *Proc. Amer. Math. Soc.*, **137** (2009), 2773–2783.

Received xxxx 20xx; revised xxxx 20xx.

E-mail address: `sqjin@whu.edu.cn`
E-mail address: `faw341@mail.usask.ca`
E-mail address: `xfzou@whu.edu.cn`