Supporting Information

Liu et al. 10.1073/pnas.1215508110

SI Text

I. Algebraic Observability

The mathematical description of a control system, which responds to external inputs $\mathbf{u}(t) \in \mathbb{R}^K$ and provides specific outputs $\mathbf{y}(t) \in \mathbb{R}^M$, is best described in the state-space form

$$\begin{cases} \dot{\mathbf{x}}(t) = \mathbf{f}(t, \mathbf{x}(t), \mathbf{u}(t)) \\ \mathbf{y}(t) = \mathbf{h}(t, \mathbf{x}(t), \mathbf{u}(t)), \end{cases} \quad [S1]$$

where $x(t) \in \mathbb{R}^N$ is the state vector of the system; $\mathbf{f}(\cdot)$ and $\mathbf{h}(\cdot)$ are in general nonlinear functions.

Assume that we have no knowledge of the initial state $\mathbf{x}(0)$ of the system, but we can monitor $\mathbf{y}(t)$ perfectly in some interval so that all their time derivatives at time $t = 0$ can be calculated. The observability problem concerns the existence of relationships between the outputs $\mathbf{y}(t)$ and their time derivatives, the state vector $\mathbf{x}(t)$, and the inputs $\mathbf{u}(t)$ such that the system's initial state $\mathbf{x}(0)$ can be deduced (1–5). From the differential algebraic point of view (6–8), the observability of a rational system is determined by the dimension of the space spanned by gradients of the Lie derivatives

$$L := \frac{\partial}{\partial t} + \sum_{i=1}^{N} f_i \frac{\partial}{\partial x_i} + \sum_{j \in \mathbb{N}} \sum_{l=1}^{K} u_l^{(j+1)} \frac{\partial}{\partial u_l^{(j)}} \quad [S2]$$

of its output functions $\mathbf{h}(t, \mathbf{x}(t), \mathbf{u}(t))$. The observability problem can be further reduced to the so-called rank test: the system **S1** is algebraically observable if and only if the $NM \times N$ Jacobian matrix

$$\mathcal{J} = \begin{bmatrix} \frac{\partial L_f^0 h_1}{\partial x_1} & \frac{\partial L_f^0 h_1}{\partial x_2} & \cdots & \frac{\partial L_f^0 h_1}{\partial x_N} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial L_f^0 h_M}{\partial x_1} & \frac{\partial L_f^0 h_M}{\partial x_2} & \cdots & \frac{\partial L_f^0 h_M}{\partial x_N} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial L_f^{N-1} h_1}{\partial x_1} & \frac{\partial L_f^{N-1} h_1}{\partial x_2} & \cdots & \frac{\partial L_f^{N-1} h_1}{\partial x_N} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial L_f^{N-1} h_M}{\partial x_1} & \frac{\partial L_f^{N-1} h_M}{\partial x_2} & \cdots & \frac{\partial L_f^{N-1} h_M}{\partial x_N} \end{bmatrix} \quad [S3]$$

has full rank (6, 7),

$$\text{rank } \mathcal{J} = N. \quad [S4]$$

For example, we can perform the rank test for chemical reaction systems with mass-action kinetics. Consider a reaction system with four chemical species {A,B,C,D} involved in two reactions (Fig. S1*D*):

$$\begin{cases} \mathcal{R}_1 : \mathrm{A} + \mathrm{B} \xrightarrow{k_1} \mathrm{C} \\ \mathcal{R}_2 : \mathrm{B} + 2\mathrm{C} \xrightarrow{k_2} \mathrm{D}. \end{cases} \quad [S5]$$

Using mass-action kinetics, the balance equations for the closed system can be written as

$$\begin{cases} \dot{x}_1 = -k_1 x_1 x_2 \\ \dot{x}_2 = -k_1 x_1 x_2 - k_2 x_2 x_3^2 \\ \dot{x}_3 = k_1 x_1 x_2 - 2k_2 x_2 x_3^2 \\ \dot{x}_4 = k_2 x_2 x_3^2 \end{cases}. \quad [S6]$$

If we consider an open system, we should introduce in-flux for pure reactants (that never act as products) and out-flux for pure products (that never act as reactants) as follows:

$$\begin{cases} \dot{x}_1 = -k_1 x_1 x_2 + C_1 \\ \dot{x}_2 = -k_1 x_1 x_2 - k_2 x_2 x_3^2 + C_2 \\ \dot{x}_3 = -k_1 x_1 x_2 - 2k_2 x_2 x_3^2 \\ \dot{x}_4 = k_2 x_2 x_3^2 - C_4 x_4 \end{cases}, \quad [S7]$$

where for pure reactants A and B, we introduce constant in-flux C_1 and C_2; and for pure product D, we introduce x-dependent out-flux $C_4 x_4$. With the extra terms due to in- and out-flux, the inference diagram changes slightly—there will be self-edges for pure products. However, this will not change the prediction made by GA, because a pure product with self-edge is still a root strongly connected component (SCC) of size 1 in the inference diagram; hence, it has to be measured to yield observability. For simplicity we only consider closed systems here. We also assume that there are no external inputs $\mathbf{u}(t)$ and we consider the simplest measurement scheme, i.e., we directly measure a subset of state variables (e.g., the concentrations of some chemical species)

$$\mathbf{y}(t) = (\cdots, x_i(t), \cdots)^{\mathrm{T}}. \quad [S8]$$

Now we show that the reaction system **S5** is algebraically observable if we measure the concentration of the pure product D, i.e., $y = x_4$.

Proof: We calculate the Lie derivatives of the output function:

$$Y^{(0)} = L_f^0 y = x_4, \quad [S9]$$

$$Y^{(1)} = L_f^1 y = k_2 x_2 x_3^2, \quad [S10]$$

$$\begin{aligned} Y^{(2)} &= L_f^2 y \\ &= k_2 \dot{x}_2 x_3^2 + k_2 x_2 2 x_3 \dot{x}_3 \\ &= k_2 \left(-k_1 x_1 x_2 - k_2 x_2 x_3^2\right) x_3^2 + k_2 x_2 2 x_3 \left(k_1 x_1 x_2 - 2k_2 x_2 x_3^2\right), \end{aligned} \quad [S11]$$

$$\begin{aligned} Y^{(3)} &= L_f^3 y \\ &= k_2 \ddot{x}_2 x_3^2 + k_2 \dot{x}_2 2 x_3 \dot{x}_3 + k_2 \dot{x}_2 2 x_3 \dot{x}_3 + k_2 x_2 2 \dot{x}_3^2 + k_2 x_2 2 x_3 \ddot{x}_3 \\ &= 2k_2 x_2 \left(k_1 x_1 x_2 - 2k_2 x_2 x_3^2\right)^2 \\ &\quad + 4k_2 x_3 \left(k_1 x_1 x_2 - 2k_2 x_2 x_3^2\right)\left(-k_1 x_1 x_2 - k_2 x_2 x_3^2\right) \\ &\quad + 2k_2 x_2 x_3 \left(-k_1^2 x_1 x_2^2 - 4k_2 x_2 x_3 \left(k_1 x_1 x_2 - 2k_2 x_2 x_3^2\right)\right. \\ &\quad + k_1 x_1 \left(-k_1 x_1 x_2 - k_2 x_2 x_3^2\right) - 2k_2 x_3^2 \left(-k_1 x_1 x_2 - k_2 x_2 x_3^2\right)\Big) \\ &\quad + k_2 x_3^2 \left(k_1^2 x_1 x_2^2 - 2k_2 x_2 x_3 \left(k_1 x_1 x_2 - 2k_2 x_2 x_3^2\right)\right. \\ &\quad \left. - k_1 x_1 \left(-k_1 x_1 x_2 - k_2 x_2 x_3^2\right) - k_2 x_3^2 \left(-k_1 x_1 x_2 - k_2 x_2 x_3^2\right)\right). \end{aligned} \quad [S12]$$

Then the Jacobian matrix can be calculated:

$$\mathcal{J}=\begin{bmatrix} \frac{\partial L_f^0 y}{\partial x_1} & \frac{\partial L_f^0 y}{\partial x_2} & \frac{\partial L_f^0 y}{\partial x_3} & \frac{\partial L_f^0 y}{\partial x_4} \\ \frac{\partial L_f^1 y}{\partial x_1} & \frac{\partial L_f^1 y}{\partial x_2} & \frac{\partial L_f^1 y}{\partial x_3} & \frac{\partial L_f^1 y}{\partial x_4} \\ \frac{\partial L_f^2 y}{\partial x_1} & \frac{\partial L_f^2 y}{\partial x_2} & \frac{\partial L_f^2 y}{\partial x_3} & \frac{\partial L_f^2 y}{\partial x_4} \\ \frac{\partial L_f^3 y}{\partial x_1} & \frac{\partial L_f^3 y}{\partial x_2} & \frac{\partial L_f^3 y}{\partial x_3} & \frac{\partial L_f^3 y}{\partial x_4} \end{bmatrix}=\begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & J_{22} & J_{23} & 0 \\ J_{31} & J_{32} & J_{33} & 0 \\ J_{41} & J_{42} & J_{43} & 0 \end{bmatrix},$$

with

$$J_{22}=k_2x_3^2$$

$$J_{23}=2k_2x_2x_3$$

$$J_{31}=k_1k_2x_2(2x_2-x_3)x_3$$

$$J_{32}=-k_2x_3\left(k_1x_1(-4x_2+x_3)+k_2x_3^2(8x_2+x_3)\right)$$

$$J_{33}=-2k_2x_2\left(k_1x_1(-x_2+x_3)+2k_2x_3^2(3x_2+x_3)\right)$$

$$\begin{aligned}J_{41}=&k_1k_2x_2\left(2k_2x_3^2\left(-8x_2^2+2x_2x_3+x_3^2\right)\right.\\&\left.+k_1\left(x_2x_3(-2x_2+x_3)+2x_1\left(2x_2^2-6x_2x_3+x_3^2\right)\right)\right)\end{aligned}$$

$$\begin{aligned}J_{42}=&k_2\left(2k_1k_2x_1x_3^2\left(-24x_2^2+4x_2x_3+x_3^2\right)\right.\\&+k_2^2x_3^4\left(72x_2^2+32x_2x_3+x_3^2\right)+k_1^2x_1\left(2x_2x_3(-3x_2+x_3)\right.\\&\left.\left.+x_1\left(6x_2^2-12x_2x_3+x_3^2\right)\right)\right)\end{aligned}$$

$$\begin{aligned}J_{43}=&2k_2x_2\left(2k_1k_2x_1x_3\left(-8x_2^2+3x_2x_3+2x_3^2\right)\right.\\&+k_2^2x_3^3\left(48x_2^2+40x_2x_3+3x_3^2\right)+k_1^2x_1(x_1(-3x_2+x_3)\\&\left.+x_2(-x_2+x_3))\right).\end{aligned}$$

It can be shown via symbolic calculation that $\mathcal{J}$ has full rank. Thus, the system is algebraically observable.

Note that in general we are not going to explicitly calculate the initial state from Eqs. **S9**–**S12**, which is usually very difficult, if not impossible. Observability only concerns whether such a solution exists.

II. Graphical Approach

A. Necessity. We can prove that measuring the sensor nodes predicted by graphical approach (GA) is necessary for the observability of an arbitrary dynamical system. The GA-predicted sensors are chosen from the root SCCs of the inference diagram derived from the system's dynamics. According to the definition of root SCCs, they have no incoming edges, implying that their information cannot be inferred from any other nodes in the inference diagram. Indeed, if we fail to measure all of the GA-predicted sensor nodes, then one or several columns in the Jacobian matrix will be zero (e.g., the gray columns of Fig. S1 *C*, *F*, and *I*, *Right*). For example, assume we do not measure a sensor node x_i; as the state of the sensor nodes can never be inferred from the dynamics of other nodes, the ith column of the Jacobian matrix will be zero: $\frac{\partial L_f^t y}{\partial x_i}=0$ for all $0\le t\le N-1$. Hence, rank $\mathcal{J}<N$ and the system is not observable, indicating the necessity of the GA-selected sensor nodes. Consider the simplest reaction system: A → B (as shown in Fig. S1*A*). If we just measure A's concentration as a function of time $x_1(t)$, we will never infer any information about the initial state of B, i.e., $x_2(0)$. Similarly, if we do not measure x_4 in Fig. S1*E* we can never infer it, because x_4 does not appear in any other node's balance equation.

B. Sufficiency. If all of the GA-selected sensor nodes are measured, then the Jacobian matrix does not contain any zero columns (e.g., Fig. S1 *C*, *F*, and *I*, *Left*). Because all nonzero elements in the Jacobian matrix are complicated polynomials of the state variables, the probability of having dependent columns in the Jacobian matrix will be very low, if not zero. In the following we show that measuring the GA-selected sensor nodes will very likely be sufficient to yield observability of biochemical reaction systems. We also show the exceptional cases and argue that they are rare.

We first give intuitive explanations about the sufficiency of monitoring the GA-selected sensor nodes for the observability of biochemical reaction systems. Our argument is based on structural control theory (9, 10). We linearize the right-hand side of the balance equations at an arbitrary state $\mathbf{x}$, obtaining the linearized system $\dot{\mathbf{x}}(t)=\mathbf{A}\mathbf{x}$, which can be associated with a weighted directed network $G(\mathbf{A})$ with its weighted adjacency matrix $a_{ij}=\frac{\partial f_i}{\partial x_j}|_{\mathbf{x}}$. In other words, if $a_{ij}\neq 0$ then it gives the strength of weight that node j can affect node i. Positive (or negative) values of a_{ij} suggest that the edge $(j\to i)$ is excitatory (or inhibitory). The weighted directed network is also called the system digraph in structural control theory (11–13).

For biochemical reaction systems, a simple rule to get the directed network $G(\mathbf{A})$ directly from the reactions is the following: For each reaction (*i*) draw a directed edge from each reactant to each product; (*ii*) draw a self-edge for each reactant; (*iii*) if there are more than two reactants, draw a bidirectional edge between every two reactants. For example, consider a reaction system with four chemical species {A,B,C,D} involved in two reactions (as shown in Fig. S1*G*):

$$\begin{cases}\mathcal{R}_1:\mathrm{A}\xrightarrow{k_1}\mathrm{B}+\mathrm{C}\\ \mathcal{R}_2:\mathrm{B}\underset{k_3}{\overset{k_2}{\rightleftharpoons}}\mathrm{D}\end{cases}. \qquad \textbf{[S13]}$$

Using mass-action kinetics, the balance equations for the closed system can be written as

$$\begin{cases}\dot{x}_1=-k_1x_1\\ \dot{x}_2=+k_1x_1-k_2x_2+k_3x_4\\ \dot{x}_3=+k_1x_1\\ \dot{x}_4=+k_2x_2-k_3x_4\end{cases}. \qquad \textbf{[S14]}$$

The directed network $G(\mathbf{A})$ associated with the balance equations is shown in Fig. S2*A*.

As observability and controllability represent mathematical duals (1, 2), we can map the observability problem of the network $G(\mathbf{A})$ into the controllability problem of the transposed network $G(\mathbf{A}^{\mathrm{T}})$, which is obtained by flipping the direction of each edge in the original network $G(\mathbf{A})$ (Fig. S2*B*). One can check that the transposed network $G(\mathbf{A}^{\mathrm{T}})$ is nothing but the inference diagram of the system (Fig. S1*H*). Assuming that the matrix elements a_{ij}'s are independent, we can perform the structural controllability analysis over the transposed network $G(\mathbf{A}^{\mathrm{T}})$ (9, 10). According to structural control theory, a system is structurally controllable if and only if its corresponding directed network contains neither inaccessible nodes nor dilations (9). Here, inaccessible nodes represent those state variables that cannot be reached from the inputs. The system contains a dilation if and only if there is a node subset S of the state variables

such that $|T(S)| < |S|$ where the neighborhood set $T(S)$ of a set S is defined to be the set of all nodes j that there exists a directed edge from j to a node in S, i.e., $T(S) = \{j | (j \to i) \in E(G), i \in S\}$. The input variables (also called origins, e.g., u_1 and u_2 in Fig. S2*D*) are not allowed to belong to S but may belong to $T(S)$. $|\cdot|$ denotes the cardinality of a set. We can show that by controlling the GA-predicted nodes, the directed network contains neither inaccessible nodes nor dilations. All of the nodes are accessible from inputs because we inject inputs to the nodes in the top layer of the underlying hierarchical structure. There are no dilations because each node (except pure products) has a self-edge (because their concentrations will affect the dynamics of themselves), which ensures that $|T(S)| \geq |S|$ for all of the subsets S. Hence, we only need to control those GA-predicted nodes to fully control the transposed network $G(\mathbf{A}^{\mathrm{T}})$. By invoking the duality of controllability and observability, we just need to measure those root nodes to fully observe the original network $G(\mathbf{A})$.

The linearized systems are usually not structured systems, because $\mathbf{A}$ is not a structured matrix—its elements (a_{ij}) are typically not independent from each other. For example, we have

$$\mathbf{A} = \begin{bmatrix} -k_1x_2 & -k_1x_1 & 0 & 0 \\ -k_1x_2 & (-k_1x_1 - k_2x_3^2) & (-2k_2x_2x_3) & 0 \\ k_1x_2 & (k_1x_1 - 2k_2x_3^2) & (-4k_2x_2x_3) & 0 \\ 0 & k_2x_3^2 & 2k_2x_2x_3 & 0 \end{bmatrix} \quad \textbf{[S15]}$$

for the system of Fig. S1*D*, with many elements depending on the same variable, hence being correlated with each other. Therefore, the structural observability analysis, and thence GA could in principle underestimate the number of sensor nodes that we need to measure. We therefore need to test GA's validity for fully nonlinear systems.

For this we randomly generated chemical reaction networks (*Materials and Methods*). We start from a few randomly generated compounds, and create mass-balanced reactions with randomly assigned rate constants k_i. Due to the arithmetic complexity involved in the generic rank calculation (14), the largest reaction system we were able to test contains 221 species involved in 121 mass-balanced reactions (15). We generate 1,000 such reaction systems and use GA to identify the sensor nodes for those systems. By using Sedoglavic's algorithm (14), we perform the rank test of the Jacobian matrix and confirm that monitoring the GA-predicted sensor nodes indeed yields full observability for almost all of the connected reaction systems we generated.

We find that exceptions occur when there are reversible reactions, e.g.,

$$\mathrm{A} + \mathrm{B} \underset{k_2}{\overset{k_1}{\rightleftharpoons}} \mathrm{C} + \mathrm{D}, \quad \textbf{[S16]}$$

which are isolated from the rest of the reaction system, forming isolated root SCCs in the inference diagram. For example, using mass-action kinetics the balance equations for this closed subsystem **S16** can be written as

$$\begin{cases} \dot{x}_1 = -k_1x_1x_2 + k_2x_3x_4 \\ \dot{x}_2 = -k_1x_1x_2 + k_2x_3x_4 \\ \dot{x}_3 = +k_1x_1x_2 - k_2x_3x_4 \\ \dot{x}_4 = +k_1x_1x_2 - k_2x_3x_4. \end{cases} \quad \textbf{[S17]}$$

One can show that measuring any single species, e.g., $y = x_1$, will not yield observability of the whole system. This is due to the existence of symmetries in the state variables leaving the output and its time derivatives invariant (14). For such a reaction, one has to measure at least two species, e.g., $\mathbf{y} = (x_1, x_2)^{\mathrm{T}}$, to achieve observability.

In principle there could be more complicated exceptional cases for which the GA-predicted sensors are not sufficient for observability, e.g., $\mathrm{A} + \mathrm{B} + \mathrm{C} \underset{k_2}{\overset{k_1}{\rightleftharpoons}} \mathrm{D} + \mathrm{E}$, $\mathrm{A} + \mathrm{B} + \mathrm{C} \underset{k_2}{\overset{k_1}{\rightleftharpoons}} 2\mathrm{D} + \mathrm{E}$, $\mathrm{A} + \mathrm{B} + \mathrm{C} \underset{k_2}{\overset{k_1}{\rightleftharpoons}} \mathrm{D} + \mathrm{E} + \mathrm{F}$, $2\mathrm{A} + \mathrm{B} + \mathrm{C} \underset{k_2}{\overset{k_1}{\rightleftharpoons}} \mathrm{D} + \mathrm{E} + \mathrm{F}$, $2\mathrm{A} + \mathrm{B} + \mathrm{C} \underset{k_2}{\overset{k_1}{\rightleftharpoons}} 2\mathrm{D} + \mathrm{E} + \mathrm{F}$. Systematic search for those exceptional cases, or equivalently the existence of nontrivial Lie subalgebra of models' symmetries letting the inputs and the outputs be invariant (14), is beyond the scope of the current work. Here, we emphasize that those exceptional cases containing isolated reactions are extremely rare. Consider N species involved in M reactions and assume each reaction contain exactly four species. The probability that four species A, B, C, and D involve in an isolated reaction, i.e., none of them involves in any other reactions, is given by $\frac{M}{\binom{N}{4}}\left[\frac{\binom{N-4}{4}}{\binom{N}{4}}\right]^{M-1}$, which decays to zero rapidly as N and M grow. For $N = M = 10$, such a probability is 2.305×10^{-12}.

Moreover, the symmetries of state variables in those exceptional cases are easily broken due to either different stoichiometry coefficients or additional reactions. For example, the reversible reaction

$$\mathrm{A} + \mathrm{B} \underset{k_2}{\overset{k_1}{\rightleftharpoons}} 2\mathrm{C} + \mathrm{D} \quad \textbf{[S18]}$$

contains no symmetry any longer and any single species can be chosen as the sensor to observe the whole system. Similarly, the reaction system

$$\begin{cases} \mathrm{A} + \mathrm{B} \underset{k_2}{\overset{k_1}{\rightleftharpoons}} \mathrm{C} + \mathrm{D} \\ \mathrm{C} \underset{k_4}{\overset{k_3}{\rightleftharpoons}} \mathrm{E} \end{cases} \quad \textbf{[S19]}$$

contains no symmetry and any single species can be chosen as the sensor to observe the whole system. Hence those exceptional cases, e.g., **S16**, are exceptionally rare in large chemical reaction systems, possibly never occurring in real systems. Indeed, we did not find any occurrence of such configurations in the metabolic networks of three well-studied model organisms, *Escherichia coli*, *Saccharomyces cerevisiae*, and *Homo sapiens*, starting from their complete metabolic reconstruction (16).

III. Real Biochemical Reaction Systems

We also tested GA on several real biochemical reaction systems, e.g., the simplified glycolytic reaction map, the model for ligand binding, and the model for cell cycle control, confirming that these systems are all observable via monitoring the minimum set of sensor nodes predicted by GA. We further apply GA to the genome-scale metabolic networks of three well-studied model organisms, *E. coli*, *S. cerevisiae*, and *H. sapiens*. We find that the fraction of sensor nodes determined by GA is not very sensitive to the assigned reversibility of the reactions in the genome-scale metabolic reconstructions.

A. Simplified Glycolytic Reaction Map. We first consider the simplified glycolytic reaction map. The system consists of $N = 10$ chemical species [glucose (Gluc); ADP; glucose 6-phosphate (G6P); ATP; glucose 1-phosphate (G1P); AMP; fructose 6-phosphate (F6P); fructose 2,6-biphosphate (F2-6BP); triose phosphate (TP); pyruvate (Pyr)] involved in $R = 9$ reactions (main text, Fig. 3*A*). The balance equations are given by

$$\begin{cases} \dot{x}_1 = -k_1x_1x_4 + C_1 \\ \dot{x}_2 = k_1x_1x_4 + 2k_4x_4x_6 - 2k_5x_2^2 + k_7x_4x_7 + k_9x_4x_7 + k_{10}x_4 - 2k_{11}x_2^2x_9 \\ \dot{x}_3 = k_1x_1x_4 - k_2x_3 + k_3x_5 - k_6x_3 \\ \dot{x}_4 = -k_1x_1x_4 - k_4x_4x_6 + k_5x_2^2 - k_7x_4x_7 - k_9x_4x_7 - k_{10}x_4 + 2k_{11}x_2^2x_9 \\ \dot{x}_5 = k_2x_3 - k_3x_5 \\ \dot{x}_6 = -k_4x_4x_6 + k_5x_2^2 \\ \dot{x}_7 = k_6x_3 - k_7x_4x_7 + k_8x_8 - k_9x_4x_7 \\ \dot{x}_8 = k_7x_4x_7 - k_8x_8 \\ \dot{x}_9 = 2k_9x_4x_7 - k_{11}x_2^2x_9 \\ \dot{x}_{10} = k_{11}x_2^2x_9 - C_2x_{10} \end{cases}. \quad \textbf{[S20]}$$

It is easy to check that pure product pyruvate (Pyr, or x_{10}) is the only root node of the inference diagram (main text, Fig. 3*B*). To check the observability of the system with Pyr ($y = x_{10}$) as the only output, we use Sedoglavic's algorithm and find that the system is indeed observable.

B. Model for Ligand Binding. Six species are incorporated in the ligand binding model (17): erythropoietin (Epo); Epo receptor (EpoR); Epo_EpoR complex; internalized complex Epo_EpoR_i; degraded internalized ligand dEpo_i, and degraded extracellular ligand dEpo_e (main text, Fig. 3*C*). Using mass-action kinetics, the reaction fluxes are given by

$$\begin{cases} v_1 = k_{on}\cdot[\text{Epo}]\cdot[\text{EpoR}] \\ v_2 = k_{on}\cdot k_D\cdot[\text{Epo\_EpoR}] \\ v_3 = k_t\cdot B_{max} \\ v_4 = k_t\cdot[\text{EpoR}] \\ v_5 = k_e\cdot[\text{Epo\_EpoR}] \\ v_6 = k_{ex}\cdot[\text{Epo\_EpoR\_i}] \\ v_7 = k_{di}\cdot[\text{Epo\_EpoR\_i}] \\ v_8 = k_{de}\cdot[\text{Epo\_EpoR\_i}] \end{cases} \quad \textbf{[S21]}$$

and the balance equations are given by

$$\begin{cases} \dfrac{d[\text{Epo}]}{dt} = -v_1 + v_2 + v_6 \\ \dfrac{d[\text{EpoR}]}{dt} = -v_1 + v_2 + v_3 - v_4 + v_6 \\ \dfrac{d[\text{Epo\_EpoR}]}{dt} = v_1 - v_2 - v_5 \\ \dfrac{d[\text{Epo\_EpoR\_i}]}{dt} = v_5 - v_6 - v_7 - v_8 \\ \dfrac{d[\text{dEpo\_i}]}{dt} = v_7 \\ \dfrac{d[\text{dEpo\_e}]}{dt} = v_8 \end{cases} \quad \textbf{[S22]}$$

One can easily check that the sensor node set consists of two pure products (dEpo_i and dEpo_e) (main text, Fig. 3*D*). To check the observability of the system with dEpo_i and dEpo_e as the outputs (y_1 = [dEpo_i], y_2 = [dEpo_e]), we use Sedoglavic's algorithm and find that the system is indeed observable.

C. Model for Cell Cycle Control in Fission Yeast. Novak and Tyson proposed a mathematical model for cell cycle control in fission yeast (18). Using standard principles of biochemical kinetics, they obtained a set of differential equations describing how the concentrations of the major state variables in their model change with time:

$$\frac{d[\text{Cdc25}]}{dt} = -\frac{k_{cr}[\text{Cdc25}]}{K_{mcr} + [\text{Cdc25}]} + \frac{k_c[\text{Cdc25C}]([\text{G2K}] + \beta[\text{PG2}])}{K_{mc} + 1 - [\text{Cdc25}]},$$

$$\frac{d[\text{G1K}]}{dt} = k_5 + (k_4 + k_{8r})[\text{G1R}] - k_8[\text{G1K}]R - [\text{G1K}]\big(V_{6p}(1 - [\text{UbE2}]) + V_6[\text{UbE2}]\big),$$

$$\frac{d[\text{G1R}]}{dt} = -k_4[\text{G1R}] - k_{6p}[\text{G1R}] - k_{8r}[\text{G1R}] + k_8[\text{G1K}]R,$$

$$\begin{aligned}\frac{d[\text{G2K}]}{dt} = {} & k_1 + (k_4 + k_{7r})[\text{G2R}] + \big(V_{25p}(1 - [\text{Cdc25}]) \\ & + V_{25}[\text{Cdc25}]\big)[\text{PG2}] - k_k[\text{G2K}][R] \\ & - [\text{G2K}]\big(V_{2p}(1 - [\text{UbE}]) + V_2[\text{UbE}]\big) \\ & - [\text{G2K}]\big(V_{wp}(1 - [\text{Wee1}]) + V_w[\text{Wee1}]\big),\end{aligned}$$

$$\begin{aligned}\frac{d[\text{G2R}]}{dt} = {} & -k_4[\text{G2R}] - k_{7r}[\text{G2R}] + k_7[\text{G2K}][R] \\ & - [\text{G2R}]\big(k_{2p} + V_{2p}(1 - [\text{UbE}]) + V_2[\text{UbE}]\big),\end{aligned}$$

$$\frac{d[\text{IE}]}{dt} = -\frac{k_{ir}[\text{IE}]}{K_{mir} + [\text{IE}]} + \frac{k_i[\text{IEC}]([\text{G2K}] + \beta[\text{PG2}])}{K_{mi} + [\text{IEC}]},$$

$$\frac{d[\text{mass}]}{dt} = \mu[\text{mass}],$$

$$\begin{aligned}\frac{d[\text{PG2}]}{dt} = {} & -\big(V_{25p}(1 - [\text{Cdc25}]) + V_{25}[\text{Cdc25}]\big)[\text{PG2}] \\ & + k_4[\text{PG2R}] + k_{7r}[\text{PG2R}] - k_7[\text{PG2}][R] \\ & - [\text{PG2}]\big(V_{2p}(1 - [\text{UbE}]) + V_2[\text{UbE}]\big) \\ & + [\text{G2K}]\big(V_{wp}(1 - [\text{Wee1}]) + V_w[\text{Wee1}]\big),\end{aligned}$$

$$\begin{aligned}\frac{d[\text{PG2R}]}{dt} = {} & -k_4[\text{PG2R}] - k_{7r}[\text{PG2R}] + k_7[\text{PG2}][R] \\ & - [\text{PG2R}]\big(k_{2p} + V_{2p}(1 - [\text{UbE}]) + V_2[\text{UbE}]\big),\end{aligned}$$

$$\begin{aligned}\frac{d[R]}{dt} = {} & k_3 + k_{6p}[\text{G1R}] + k_{8r}[\text{G1R}] + k_{7r}[\text{G2R}] + k_{7r}[\text{PG2R}] \\ & - k_4[R] - k_8[\text{G1K}][R] - k_7[\text{G2K}][R] - k_7[\text{PG2}][R] \\ & - \frac{k_p[\text{mass}]([\text{Cig1}] + \alpha[\text{G1K}] + [\text{G2K}] + \beta[\text{PG2}])[R]}{K_{mp} + [R]} \\ & + [\text{G2R}]\big(k_{2p} + V_{2p}(1 - [\text{UbE}]) + V_2[\text{UbE}]\big) \\ & + [\text{PG2R}]\big(k_{2p} + V_{2p}(1 - [\text{UbE}]) + V_2[\text{UbE}]\big),\end{aligned}$$

$$\frac{d[\text{UbE}]}{dt} = -\frac{k_{ur}[\text{UbE}]}{K_{mur} + [\text{UbE}]} + \frac{k_u[\text{IE}](1 - [\text{UbE}])}{K_{mu} + 1 - [\text{UbE}]},$$

$$\frac{d[\text{UbE2}]}{dt} = -\frac{k_{ur2}[\text{UbE2}]}{K_{mur2} + [\text{UbE2}]} + \frac{k_{u2}([\text{G2K}] + \beta[\text{PG2}])(1 - [\text{UbE2}])}{K_{mu2} + 1 - [\text{UbE2}]},$$

$$\frac{d[\text{Wee1}]}{dt} = -\frac{k_w([\text{G2K}] + \beta[\text{PG2}])[\text{Wee1}]}{K_{mw} + [\text{Wee1}]} + \frac{k_{wr}(1 - [\text{Wee1}])}{K_{mwr} + 1 - [\text{Wee1}]},$$

with rate constants $k_1 = 0.015$, $k_3 = 0.09375$, $k_{2'} = 0.05$, $k_4 = 0.1875$, $k_5 = 0.00175$, $k_7 = 100$, $k_{7r} = 0.1$, $k_{6'} = 0$, $k_8 = 10$, $k_{8r} = 0.1$, $k_p = 3.25$, $k_i = 0.4$, $k_{ir} = 0.1$, $k_{ur} = 0.1$, $k_u = 0.2$, $k_{ur2} = 0.3$, $k_{u2} = 1$, $k_{wr} = 0.25$, $k_w = 1$, $k_{cr} = 0.25$, $k_c = 1$, $V_2 = 0.25$, $V_{2'} = 0.0075$, $V_{25} = 0.5$, $V_{25'} = 0.025$, $V_6 = 7.5$, $V_{6'} = 0.0375$, $V_w = 0.35$, $V_{w'} = 0.035$, Michaelis constants $K_{mc} = K_{mcr} = 0.1$, $K_{mi} = K_{mir} = 0.01$, $K_{mp} = 0.001$, $K_{mu} = K_{mur} = 0.01$, $K_{mu2} = K_{mur2} = 0.05$, $K_{mw} = K_{mwr} = 0.1$, and miscellaneous constants $\alpha = 0.25$, $\beta = 0.05$, $\mu = 0.00495$, [Cig1] = 0.

The associated inference diagram contains two SCCs: {[mass]} and {[Cdc25], [G1K], [G1R], [G2K], [G2R], [IE], [PG2], [PG2], [PG2R], [R], [UbE], [UbE2], [Wee1]} (Fig. S3). The latter is a root SCC. We verified, via Sedoglavic's algorithm, that by monitoring any node in the root SCC, the system is observable.

D. Genome-Scale Metabolic Networks. We applied GA to the genome-scale metabolic networks of three well-studied model organisms, *E. coli*, *S. cerevisiae*, and *H. sapiens* (16). During the genome-scale metabolic network reconstruction, the reversibility and directionality of reactions have to be carefully assigned. To achieve this, the thermodynamic consistency analysis has been introduced in the metabolic reconstruction process (19, 20). The detailed verification and error diagnostics of the assigned reversibility and directionality of the reactions in the genome-scale metabolic reconstructions are beyond the scope of our current research. However, to determine the sensitivity of our result on the assignment of the reaction reversibility, we performed the following test. For the reaction list of each model organism, we randomly selected a p fraction of irreversible reactions, changed them to be reversible, and recalculated the fraction of sensors (n_s) from the modified inference diagram. This random selection is repeated 10 times to get statistical error. We then plot n_s as a function of p (Fig. S4). We find that n_s decreases slowly as p increases. For example, for the genome-scale metabolic network of *E. coli* (iAF1260), $n_s(p=0) \approx 0.058$ and $n_s(p=0.9) \approx 0.028$, which means that if 90% of the original irreversible reactions were actually reversible, the fraction of sensors will decrease by 52%. This calculation suggests that our result is not very sensitive to the assigned reversibility of the reactions in the genome-scale metabolic reconstructions.

IV. Linear Observer

Observability only concerns our ability to reconstruct the internal state of a system from its outputs. GA described in this work helps us identify the nodes through which we can observe a complex system— it does not tell us how to do it. To achieve actual observability, we need to explicitly build an observer, i.e., a dynamic device that models a real system through which we uncover, from the available outputs, the rest of the (unmeasured) variables. For the sake of completion, we illustrate this procedure for a small linear reaction system—for nonlinear systems the observer construction is rather involved and still an open and active area of research (21).

Consider the chemical reaction system shown in Fig. S1*G*, which has linear balance equations. GA predicts a minimum sensor set contains two nodes (x_3 and x_4), and monitoring them should in principle yield full observability. Because this is a linear dynamic system, one can easily design a Luenberger observer.

A general linear time-invariant system is described by

$$\begin{cases} \dot{\mathbf{x}}(t) = \mathbf{A}\,\mathbf{x}(t) + \mathbf{B}\,\mathbf{u}(t) \\ \mathbf{y}(t) = \mathbf{C}\,\mathbf{x}(t). \end{cases} \quad \textbf{[S23]}$$

The Luenberger observer has the following form:

$$\dot{\mathbf{z}}(t) = \mathbf{A}\,\mathbf{z}(t) + \mathbf{L}[\mathbf{y}(t) - \mathbf{C}\,\mathbf{z}(t)] + \mathbf{B}\,\mathbf{u}(t), \quad \textbf{[S24]}$$

where the $N \times K$ matrix $\mathbf{L}$ is to be specified later. Note that if the observer is initiated with $\mathbf{z}(0) = \mathbf{x}(0)$, then it follows that $\mathbf{z}(t) = \mathbf{x}(t)$ exactly for all $t > 0$. Because $\mathbf{x}(0)$ is usually unavailable (which is the very reason we need an observer), we have $\mathbf{z}(0) \neq \mathbf{x}(0)$; we hope $\mathbf{z}(t)$ will asymptotically converge to $\mathbf{x}(t)$, i.e., the state of the observer tracks the state of the original system. This can be guaranteed by a suitable choice of the $\mathbf{L}$ matrix. Consider the time evolution of the error vector $\mathbf{e}(t) = \mathbf{z}(t) - \mathbf{x}(t)$. From Eqs. **S23** and **S24**, one has

$$\begin{aligned} \dot{\mathbf{e}}(t) &= \dot{\mathbf{z}}(t) - \dot{\mathbf{x}}(t) \\ &= [\mathbf{A} - \mathbf{LC}][\mathbf{z}(t) - \mathbf{x}(t)] \\ &= [\mathbf{A} - \mathbf{LC}]\mathbf{e}(t). \end{aligned} \quad \textbf{[S25]}$$

If the matrix $[\mathbf{A} - \mathbf{LC}]$ is asymptotically stable, the error vector will converge to zero with rate determined by the largest eigenvalue of $[\mathbf{A} - \mathbf{LC}]$.

For the linear reaction system shown in Fig. S1*G*, we have

$$\begin{cases} \dot{x}_1 = -k_1 x_1 \\ \dot{x}_2 = +k_1 x_1 - k_2 x_2 + k_3 x_4 \\ \dot{x}_3 = +k_1 x_1 \\ \dot{x}_4 = +k_2 x_2 - k_3 x_4. \end{cases} \quad \textbf{[S26]}$$

The output vector is given by $\mathbf{y} = (x_3, x_4)^{\mathrm{T}}$ and for simplicity we have assumed there are no external inputs $\mathbf{u}(t)$. Then, the Luenberger observer is given by

$$\begin{cases} \dot{z}_1 = -k_1 z_1 \\ \dot{z}_2 = +k_1 z_1 - k_2 z_2 + k_3 z_4 \\ \dot{z}_3 = +k_1 z_1 + K(x_3 - z_3) \\ \dot{z}_4 = +k_2 z_2 - k_3 z_4 + K(x_4 - z_4) \end{cases} \quad \textbf{[S27]}$$

denoted as observer 1, where we choose

$$\mathbf{L} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ K & 0 \\ 0 & K \end{pmatrix} \quad \textbf{[S28]}$$

and K is a constant (called observer gain) (2). We show that the observer can be used to monitor the dynamics of the system and the state of the observer indeed tracks the state of the original system (Fig. S5).

However, if we do not measure x_4 and just measure x_3, from the inference diagram we can easily tell that x_2 and x_4 can never be inferred. This can also be seen from the simulation of the following observer:

$$\begin{cases} \dot{z}_1 = -k_1 z_1 \\ \dot{z}_2 = +k_1 z_1 - k_2 z_2 + k_3 z_4 \\ \dot{z}_3 = +k_1 z_1 + K(x_3 - z_3) \\ \dot{z}_4 = +k_2 z_2 - k_3 z_4 \end{cases} \quad \textbf{[S29]}$$

denoted as observer 2. One sees that indeed this observer will not track x_2 and x_4 at all (Fig. S6).

V. Other Dynamic Systems

A. Ecological Systems. We consider an ecological community consisting of N species, in which population dynamics is driven by interspecific interactions. We assume a Holling type I functional response, and the population dynamics of species i is given by

$$\dot{x}_i = x_i\left(r_i - s_i x_i + \sum_{j=1, j\neq i}^{N} \gamma_{ij} x_j\right), \qquad \text{[S30]}$$

where r_i is the intrinsic population and mortality rate of species i, s_i is density-dependent self-regulation, and γ_{ij} represents the interaction coefficient between two species i and j. Two species i and j interact with probability C. We consider three different types of interactions: (i) random, where γ_{ij} and γ_{ji} are uncorrelated and they do not have to appear in pair; (ii) predator–prey, where γ_{ij} and γ_{ji} always appear in pair and have opposite signs; (iii)mixture of competition and mutualism, where γ_{ij} and γ_{ji} always appear in pair and have the same sign, either "+" or "−" representing mutualistic or competitive interaction, respectively.

We generate 100 ecological systems with $N = 50$ species, interaction probability $C = 0.05$, and randomly assigned parameter values. We then identify the sensor species using GA. By using Sedoglavic's algorithm, we find those systems are indeed observable by monitoring the GA-selected sensor species for all of the three interaction types.

B. Neuron Systems. As a simplification of the classical Hodgkin–Huxley neuron model (22), the Hindmarsh–Rose model (23) aims to model the spiking–bursting behavior of a single neuron. This model contains three state variables: $x(t), y(t)$, and $z(t)$, representing the membrane potential, the transport rate of sodium and potassium ions through fast ion channels (spiking variable), and the transport rate of other ions thorough slow channels (bursting variable), respectively. Here, we consider a diffusion-coupled directed network with N identical Hindmarsh–Rose neurons. The dynamic equation of each neuron is given by

$$\begin{cases} \dot{x}_i = y_i + \phi(x_i) - z_i + I_a + \sum_{j\in\partial^+ i}(v_j - v_i) \\ \dot{y}_i = \psi(x_i) - y_i \\ \dot{z}_i = r[s(x_i - x_{\mathrm{R}}) - z_i], \end{cases} \qquad \text{[S31]}$$

with $\phi(x) = ax^2 - x^3$, $\psi(x) = 1 - bx^2$, and $\partial^+ i$ representing all nodes pointed by node i.

We generate 100 neuron systems with up to $N = 50$ neurons randomly connected with each other with probability $C = 0.1$. We fix the system parameters to be $s = 4, x_{\mathrm{R}} = -8/5, a = 3, b = 5$, and $r = 0.001$. We consider the current I_a that enters the neuron as an input and assume it is the same for all of the neurons (24). We then identify the sensor neurons using GA. By using Sedoglavic's algorithm, we verify that those systems are indeed observable by monitoring the GA-predicted sensor neurons.

1. Kalman RE (1963) Mathematical description of linear dynamical systems. *J Soc Indus Appl Math Ser A* 1(2):152–192.
2. Luenberger DG (1979) *Introduction to Dynamic Systems: Theory, Models, & Applications* (Wiley, New York).
3. Chui CK, Chen G (1989) *Linear Systems and Optimal Control* (Springer, New York).
4. Slotine JJ, Li W (1991) *Applied Nonlinear Control* (Prentice-Hall, Englewood Cliffs, NJ).
5. Isidori A (1995) *Nonlinear Control Systems* (Springer, Berlin).
6. Diop S, Fliess M (1991) On nonlinear observability. *Proceedings of ECC'91* (Hermès, Paris), Vol 1, pp 152–157.
7. Diop S, Fliess M (1991) Nonlinear observability, identifiability, and persistent trajectories. *Proceedings of the 30th IEEE Conference on Decision and Control* (IEEE Press, New York), Vol 1, pp 714–719.
8. Anguelova M April (2004) PhD thesis (Chalmers University of Technology and Göteborg University, Göteborg, Sweden).
9. Lin CT (1974) Structural controllability. *IEEE Trans Automat Control* 19(3):201–208.
10. Liu YY, Slotine JJ, Barabási AL (2011) Controllability of complex networks. *Nature* 473(7346):167–173.
11. Šiljak DD (1978) *Large-scale Dynamic Systems: Stability and Structure* (North-Holland, New York).
12. Khan UA, Moura JMF (2008) Distributing the Kalman filter for large-scale systems. *IEEE Trans Signal Process* 56(10):4919–4935.
13. Doostmohammadian M, Khan UA (2011) Communication strategies to ensure generic networked observability in multi-agent systems. *Conference Record of the Forty Fifth Asilomar Conference on Signals, Systems and Computers (ASILOMAR)* (IEEE Press, New York), pp 1865–1868.
14. Sedoglavic A (2002) A probabilistic algorithm to test local algebraic observability in polynomial time. *J Symb Comput* 33(5):735–755.
15. Basler G, Nikoloski Z (2011) JMassBalance: Mass-balanced randomization and analysis of metabolic networks. *Bioinformatics* 27(19):2761–2762.
16. Schellenberger J, Park JO, Conrad TM, Palsson BØ (2010) BiGG: A Biochemical Genetic and Genomic knowledgebase of large scale metabolic reconstructions. *BMC Bioinf* 11:213.
17. Raue A, Becker V, Klingmüller U, Timmer J (2010) Identifiability and observability analysis for experimental design in nonlinear dynamical models. *Chaos* 20(4):045105.
18. Novak B, Tyson JJ (1997) Modeling the control of DNA replication in fission yeast. *Proc Natl Acad Sci USA* 94(17):9147–9152.
19. Kümmel A, Panke S, Heinemann M (2006) Systematic assignment of thermodynamic constraints in metabolic network models. *BMC Bioinformatics* 7:512.
20. Feist AM, et al. (2007) A genome-scale metabolic reconstruction for Escherichia coli K-12 MG1655 that accounts for 1260 ORFs and thermodynamic information. *Mol Syst Biol* 3:121.
21. Chaves M, Sontag ED (2002) State-estimators for chemical reaction networks of Feinberg-Horn-Jackson zero deficiency type. *Eur J Control* 8(4):343–359.
22. Hodgkin AL, Huxley AF (1952) A model of the nerve impulse using two first-order differential equations. *J Physiol* 117:500–554.
23. Hindmarsh JL, Rose RM (1982) A model of the nerve impulse using two first-order differential equations. *Nature* 296(5853):162–164.
24. Tabareau N, Slotine JJ, Pham QC (2010) How synchronization protects from noise. *PLOS Comput Biol* 6(1):e1000637.

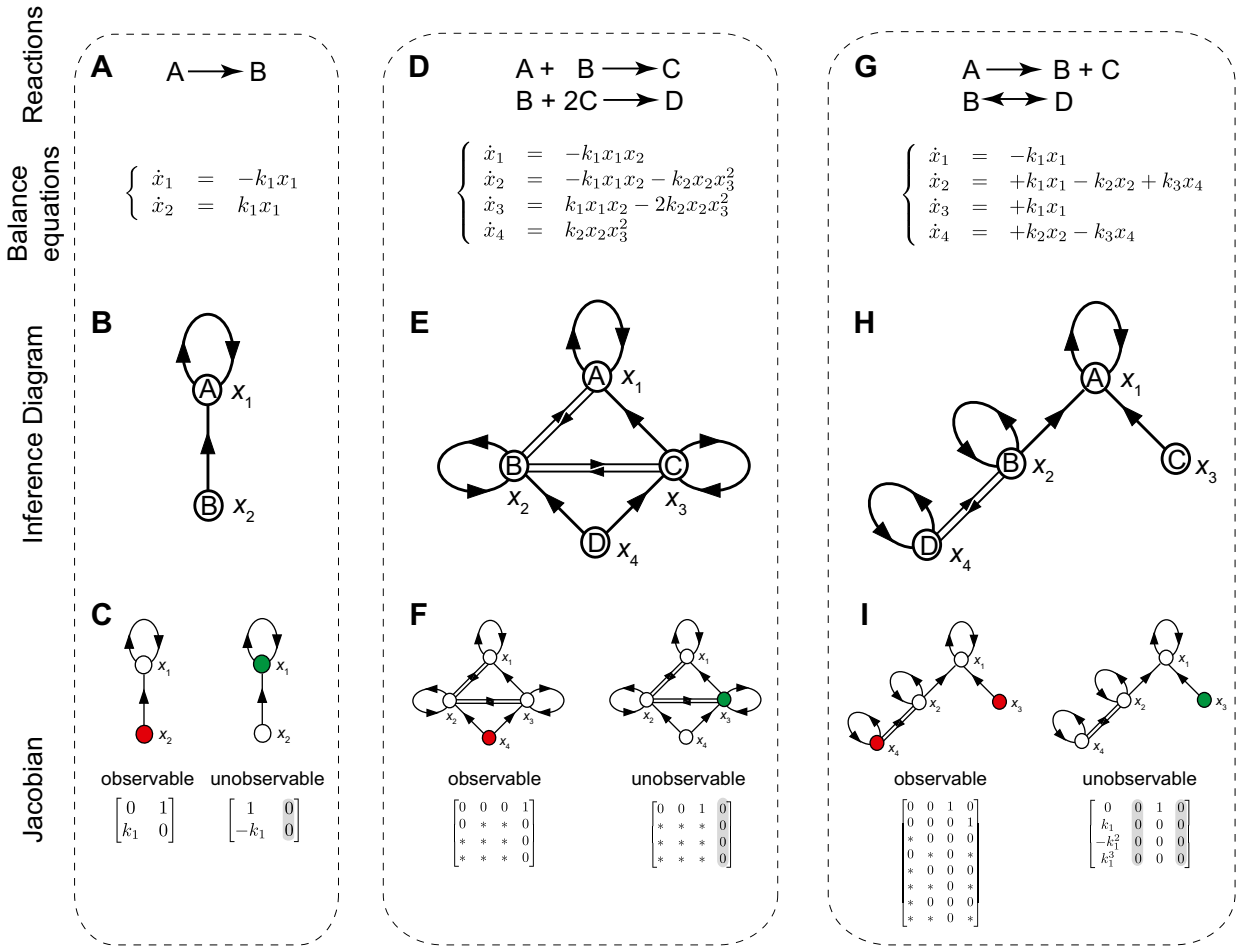

Fig. S1. Observability of simple chemical reaction systems. For each reaction system, we show its reaction(s), balance equations, inference diagram, and the Jacobian matrices corresponding to an observable case by measuring the GA-selected sensor node(s) (shown in red); and an unobservable case (with sensors shown in green). (*A*) A simple chemical reaction with two species A and B. The balance equations associated with this reaction are linear. (*B*) Inference diagram derived from the balance equations shown in *A*. (*C*) Jacobian matrices corresponding to an observable case (*Left*) and a nonobservable case (*Right*). (*D*) A chemical reaction system with four species (A, B, C, and D) involved in two reactions. The associated balance equations are nonlinear. (*E*) Inference diagram derived from the balance equations shown in *D*. (*F*) Jacobian matrices corresponding to an observable case (*Left*) and a nonobservable case (*Right*). (*G*) A chemical reaction system with four species (A, B, C, and D) involved in two reactions (one is reversible). The associated balance equations are linear. (*H*) Inference diagram derived from the balance equations shown in *G*. (*I*) Jacobian matrices corresponding to an observable case (*Left*) and a nonobservable case (*Right*).

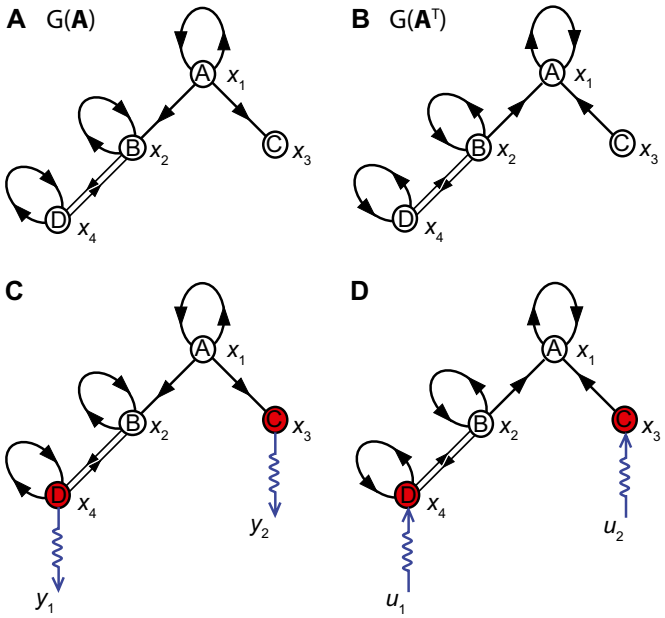

Fig. S2. Duality between observability and controllability. (*A*) Balance equations of a chemical reaction system can be associated with a directed weighted network $G(\mathbf{A})$, called the system digraph. (*B*) By flipping the direction of each edge in $G(\mathbf{A})$, we get a transposed network $G(\mathbf{A}^T)$, which is just the inference diagram introduced in the main text. (*C*) To ensure the observability, we need to measure nodes x_3 and x_4 as outputs y_1 and y_2. (*D*) To ensure the controllability, we need to drive nodes x_3 and x_4 with inputs u_1 and u_2.

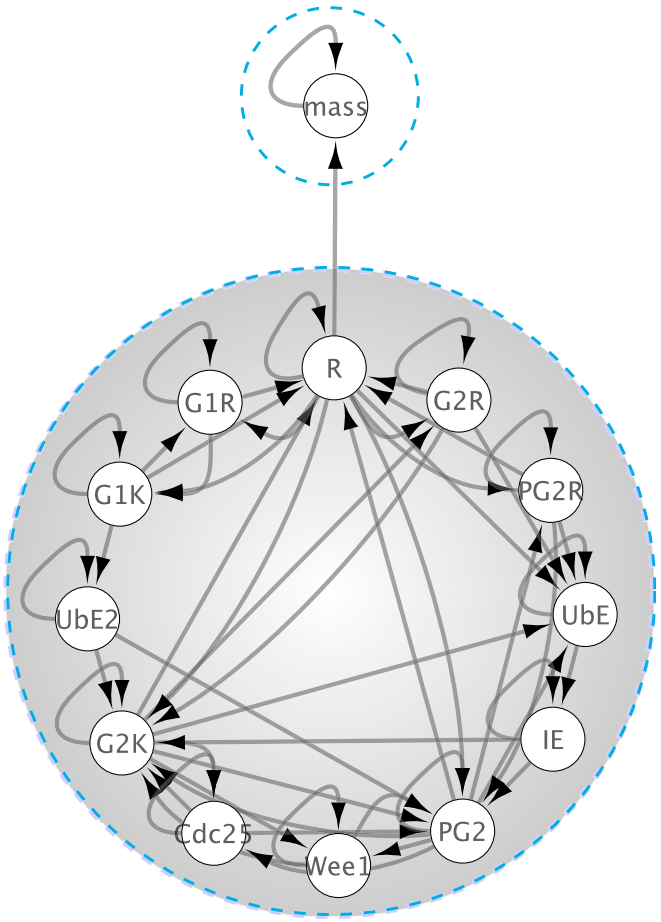

Fig. S3. Inference diagram of the cell cycle control model in fission yeast (18). Strongly connected components (SCCs) are marked with dashed circles. The root SCC, which has no incoming links, is shaded in gray.

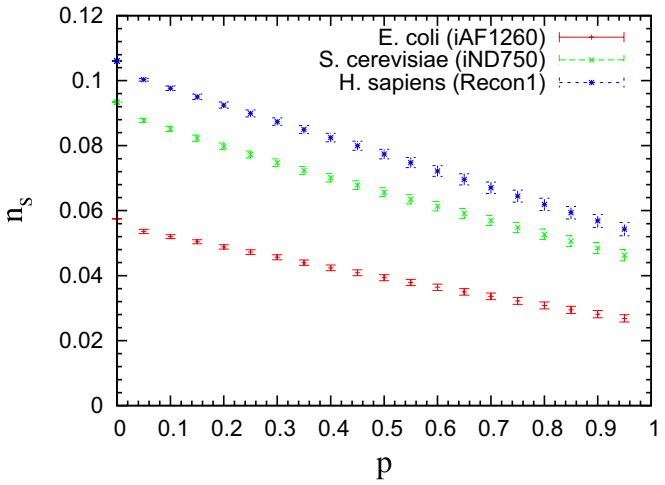

Fig. S4. Fraction of sensor nodes as a function of p, with p the fraction of irreversible reactions in the original genome-scale metabolic networks of three well-studied model organisms, *Escherichia coli* (red "+"), *Saccharomyces cerevisiae* (green "x"), and *Homo sapiens* (blue "*"). This p fraction of irreversible reactions is forced to be reversible, and then we recalculate the fraction of sensor nodes of the perturbed metabolic networks. The results are averaged over 10 random selections of p fraction of irreversible reactions, with error bars defined as SEM.

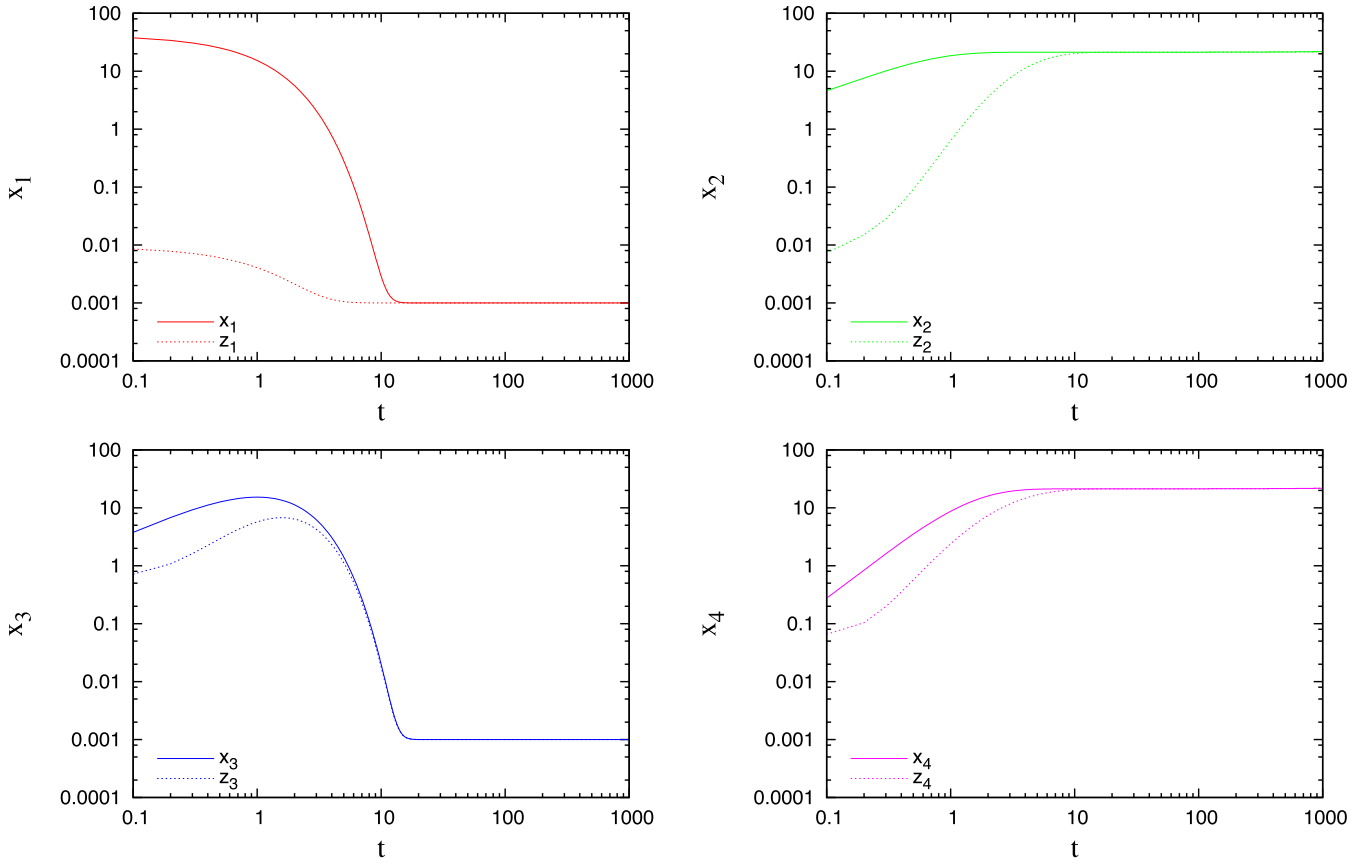

Fig. S5. Observer 1 works: the estimates $z_i(t)$ (shown in dotted lines) will converge to the original state variables $x_{i(t)}$ (shown in solid line) for large t. Here, x_1, x_2, x_3, and x_4 are shown in red, green, blue, and magenta, respectively.

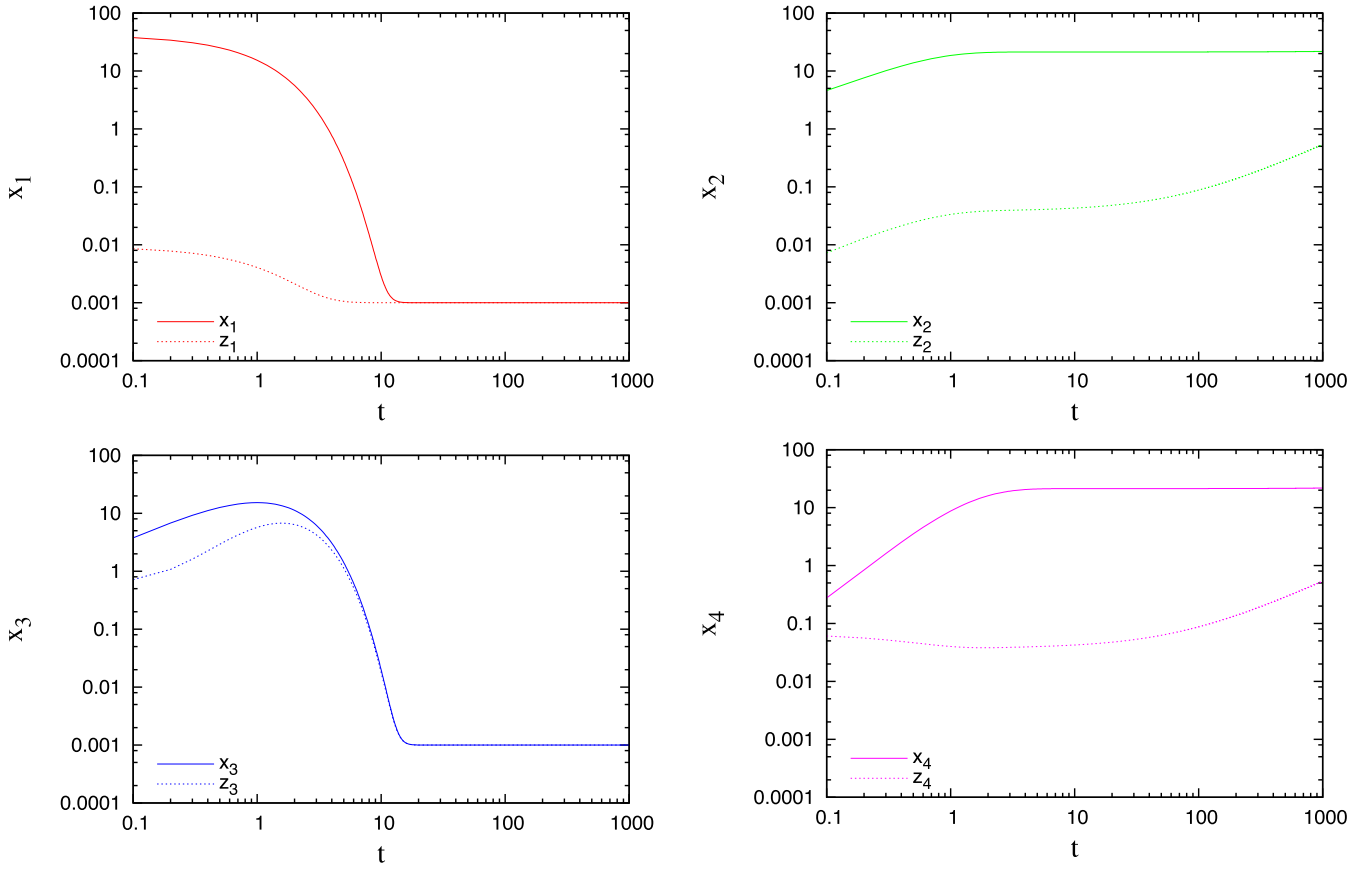

Fig. S6. Observer 2 does not work: only the estimates z_1 (t) and z_3 (t) will converge to the original state variables x_1 (t) and x_3 (t), respectively, at large t. The estimates z_2 (t) and z_4 (t) will not converge to x_2 (t) and x_4 (t) Here, x_1, x_2, x_3, and x_4 are shown in red, green, blue, and magenta, respectively.