# 3 Signal Transduction in Biological Systems and its Possible Uses in Computation and Communication Systems

**LIU Jian-Qin and PEPER Ferdinand**

Cellular signal transduction is crucial for cell communication and is described by signaling pathway networks that sustain the biological functions of living cells. The robustness of the molecular mechanism of cellular signal transduction forms an important inspiration in the design of future communication networks based on the information processing mechanisms of cellular signal transduction. This paper discusses some important aspects of the computational issues on cellular signal transduction: (1) How to formally represent kernel information of cellular signal transduction; (2) How to get a fixed point from a pathway network with feedbacks; and (3) How to encode information in signal transduction pathways by error-correcting codes, such as to increase the fault-tolerance of the system, while at the same time conform to the unstructured nature of such pathways. The above results provide a basis for innovative future communications networks, with biological signaling pathway networks acting as references for systems with improved performance in factors like robustness.

***Keywords***
Cell communication, Signal transduction, Signaling pathway networks, Error Correcting codes

## 1 Introduction

Signal transduction plays a crucial role in the complex dynamics of living cells to the extent that it is considered a fundamental information processing mechanism in living systems. The recent availability of data on signal transduction has the potential for the creation of artificial systems conducting computation and communication using its inherent mechanisms. It also promises to give inspiration on building computation and communication systems in our world that are based on novel principles only used to date in biological organisms.

It is yet unknown what advantages can be gained from using biologically-inspired mechanisms in the application to information processing systems, but given the high efficiency by which biological organisms function, it makes sense to study them, especially in the framework of an information processing paradigm. The boundary conditions of the processes in biological systems tend to be quite different from what we are used to in our daily lives. Noise, for example, is large compared to signal levels. Mechanisms to cope with it in traditional systems include error correcting codes, and it is an interesting issue to investigate whether and to what extent such techniques are applicable in biological systems. To investigate issues like these it is important to have a formal model describing biological signal transduction. The most commonly used model is the network, which has topological features such as hubs in a scale-

free network[1]. This suggest the exploration of efficient ways to systematically understand the robustness of networks in terms of graphs, where the building block of the signal transduction networks that are treated as complex systems called *network motif* in[2] are defined. Based on the technology of networks, we can model the dynamics of signal transduction networks and find a quantitative description of its signaling mechanism that sustains the robustness of the corresponding cellular signaling processes, which have been widely reported in the signal transduction networks for chemotaxis, heat shock response, ultrasensitivity, and cell cycle control[3].

In this paper, we formulate a model of signal transduction in terms of graph theory in order to increase our understanding of its information processing role in biological systems. A graph is a set of nodes, the so-called *vertexes*, with relations between them called the *edges*. In the framework of a communication network, for example, the edges represent the communication channels between the nodes between which communication takes place. Applied to biological systems, we obtain a formulation of the structure of signaling pathways of signal transduction. We illustrate our concepts through a particular protein called MAPK (Mitogen-Activating Protein Kinase), which plays an important role in intra-cellular communication processes. We then add concepts related to error correcting codes to study how robustness of the above processes can be improved, as well as a new concept of *fixed points* in biological systems, which serve to restore signals through nonlinear dynamics with feedback.

This paper is focused on computational aspects of signal transduction in cells. In Section **2**, based on the biochemical features of signal transduction processes, the data structure of a graph is presented to formulate the signaling pathway of signal transduction, in which MAPK is discussed as an instance of pathways for describing the dynamical processes of signal transduction networks. In Section **3**, a fixed point phenomenon is studied as well as the robustness factor for developing molecular communication systems. In Section **4**, molecular codes for error correction are described from the network level of signal transduction.

## 2 Cellular signal transduction networks and their formal model

Signal transduction in cells is a biochemical process that is of fundamental importance for their functioning. In living cells signal transduction is carried out by series of biochemical reactions that are regulated by genetic factors. The signals used in signal transduction are usually quantified by concentrations of the corresponding chemicals.

### 2.1 Some preliminaries of the biochemistry of signal transduction

Cellular signal transduction is defined as a phenomenon, process, or mechanism that realizes a series of biochemical reactions in cells in response to stimuli of chemical signals outside the cells; this function includes so-called cell communication.

Cell communication is the term used for the communication processes in cells that take the form of chemical signals and that is realized by the biochemical reactions in cells through cellular signal transduction. Cell communication can be distinguished into inter-cellular and intra-cellular communications.

*Inter-cell communication* describes how cells interact with each other. An important mechanism in inter-cell communication is formed by signaling molecules, which are also known as *first messengers*. Inter-cell communication has four types[4]: (1) contact-dependent signaling, in which cells have direct membrane-to-membrane contact to exchange signals, (2) paracrine signaling, in which cells release signals into the extracellular space to act locally on neighboring cells, (3) synaptic signaling, in which neuronal cells transmit signals electrically along their axons and release

neurotransmitters at synapses, and (4) endocrine signaling, in which hormonal signals are secreted into the blood stream to be distributed on a wide scale throughout an organism's body.

*Intra-cell communication* concerns the communication within cells. When an incoming first messenger molecule reaches a cell, it cannot directly pass the cell membrane, but is bound by specific receptors that effectuate the activation of certain signaling molecule proteins within the cell. Referred to as *secondary messengers*, these signals are relayed by a chemical reaction process through a signaling cascade, which relays the signals to the nucleus of the cell. In this paper we will mainly write about intra-cell communication. Important signaling molecules in cells are proteins. We will be especially interested in proteins that can bind to a phosphate molecule. Such proteins are called *phospho-proteins*. When a phosphate is attached to a protein it becomes *phosphorylized* through a *phosphorylation* process; when it becomes detached it becomes *dephosphorylized* through a *dephosphorylation* process. To switch between the two states, special enzymes are required. The enzyme that realizes phosphorylation is called *kinase*, the enzyme that realizes dephosphorylation is called *phosphatase*.

The phosphorylation / dephosphorylation state of a protein will be used in the following to encode the binary state of a variable. This state can be detected by *immunofluorescence analysis*, which provides us with a possible tool to read out such a variable.

The signaling pathway in cells is a series of biochemical reactions, which have specific biological functions.

## 2.2 Graph representation for signal transduction

The reactions in a signaling pathway will be described by a directed graph with input and output. By the graph representation, we can get the information form of the signal transduction network, from which we can investigate the structure, encoding, and networks of signal transduction in cells.

To study the structural relations concerning the transduction of signals, we define transduction in a spatial form as a graph. Let a molecule be represented by a vertex (node) in a graph and let a biochemical reaction be represented by an edge (link), then we obtain a graph

$$G = <V, E>,$$

where the vertex set is defined as the set $V = \{V_1, V_2, .., V_n\}$ and the edge set is defined as $E = E\,(V_i, V_j)\ (V_i, V_j \in V)$. Any parameters of a biochemical reaction represented by an edge are depicted as labels to that edge.

The direction of a biochemical reaction is represented in this formalism by a directed edge in the graph, which is graphically depicted as an arrow from one vertex to another vertex. In case a biochemical reaction is bidirectional, the corresponding edge is also bidirectional, and it will be depicted as merely a line between its vertices. Figure 1 shows a graph representing a pathway from a substrate vertex to a product vertex.

The actual phosphorylation process follows the so-called *Michaelis-Menten* equation[5], which appears in the graph as a label on the edge between the substrate-vertex and the product-vertex. The Michaelis-Menten equation is described as follows. Let the reactant denote the input to the pathway, then the product is calculated by the Michaelis-Menten equation as

$$d/dt\,(\text{product}\,(t)) = k_3 * \text{enzyme} * \text{substrate}\,(t) \,/\, (\text{substrate}\,(t) + k_m)$$

where product (t) is the product concentration at time t, substrate (t) is the substrate concentration, enzyme is the enzyme concentration, and $k_1$, $k_2$, and $k_3$ are the coefficients of the biochemical reaction, whereas $k_m = k_2 / k_1$ and it is assumed that $k_3 \ll k_2$.

The above formalism is used for describing an individual pathway (Cf. Fig. 2). If a pathway can not be divided into any other

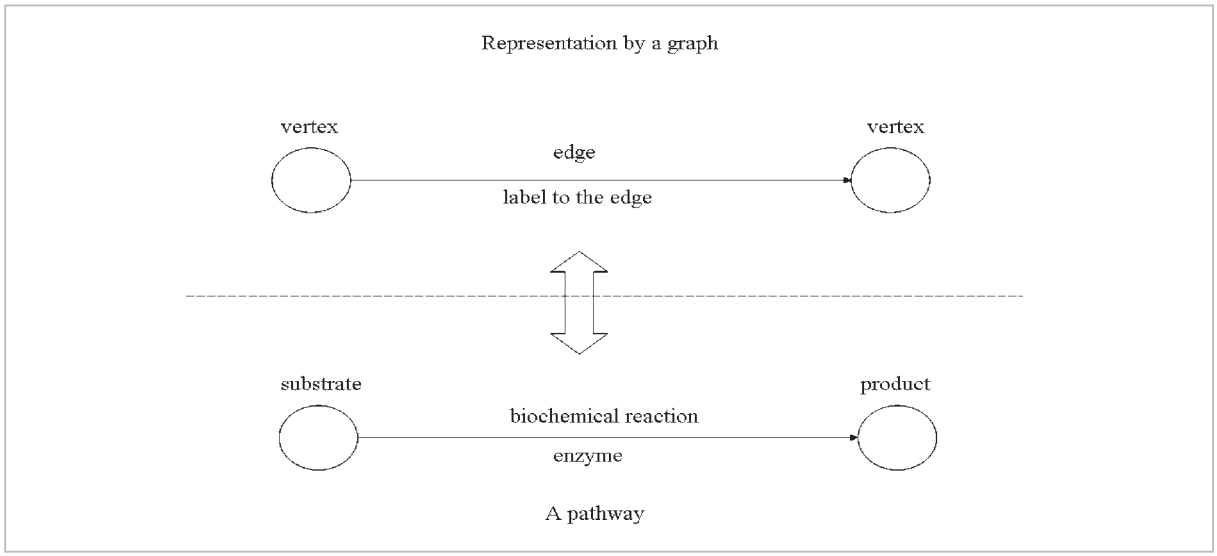


**Fig.1** *A pathway: a graph vs. signals*

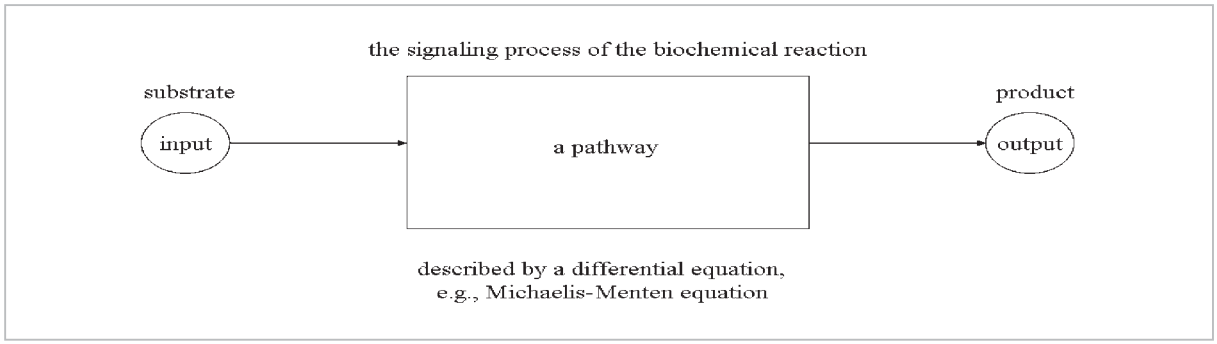


**Fig.2** *The concept of pathway as a system*

pathways, the pathway is called *indivisible* or *atomic*. Atomic pathways are the building blocks from which more complex pathway networks are constructed. Such complex pathways are called *interacted pathways*. The entire pathway network will be constructed by these building blocks. In case of phosphorylation and dephosphorylation states, for example, it is possible to switch between these states via the interacted pathways that are regulated by kinases and phosphotases.

The interaction of different states needs to be investigated considering their influence on biological functions of cells. The MAPK cascade is one of the important pathways with such features.

## 2.3 Example: the MAPK cascade

A MAPK cascade is an important pathway that is at the base of many biological functions in cells, such as in a phosphorylation process. Involved in a MAPK cascade is a number of kinases with the following names:

MAPK: mitogen activating protein kinase,
MAPKK: mitogen activating protein kinase kinase,
MAPKKK: mitogen activating protein kinase kinase kinase,
MAPKKKK: mitogen activating protein kinase kinase kinase kinase,

The resulting pathway is called k-layered, with k being an integer denoting the number of stages in the cascade. The structures of two MAPK cascades are is illustrated in Figs. 3 and 4. From top to down in the order going from upstream to down stream, MAPKKKK phosphorylates MAPKKK, MAPKKK phosphorylates MAPKK, MAPKK phosphorylates

MAPK, which effectuates a phosphorylated protein as output to the process.

When building information models of signal transduction, it is important to structurally analyze the functionality of a signaling pathway network. As an example of a structural model, we show the MAPK cascade of budding yeast *Schizosaccharomyces pombe* in Fig. 4, which involves the kinases Ste 20, Ste 11, Ste 7, Kss 1 / Fus 3 and Far 1 / Ste 12. In this cascade, MAPKKKK is Ste 20, MAPKKK is Ste 11, MAPKK is Ste 7, and there are two MAPK factors, Kss 1 and Fus 3. Furthermore, there are proteins at the bottom of the MAPK cascade, called Far 1 and Ste 12; these proteins, which form the output of the cascade, play an active role in the reproduction of cells. In technical terms, Far 1 is a CDK (Cyclin-

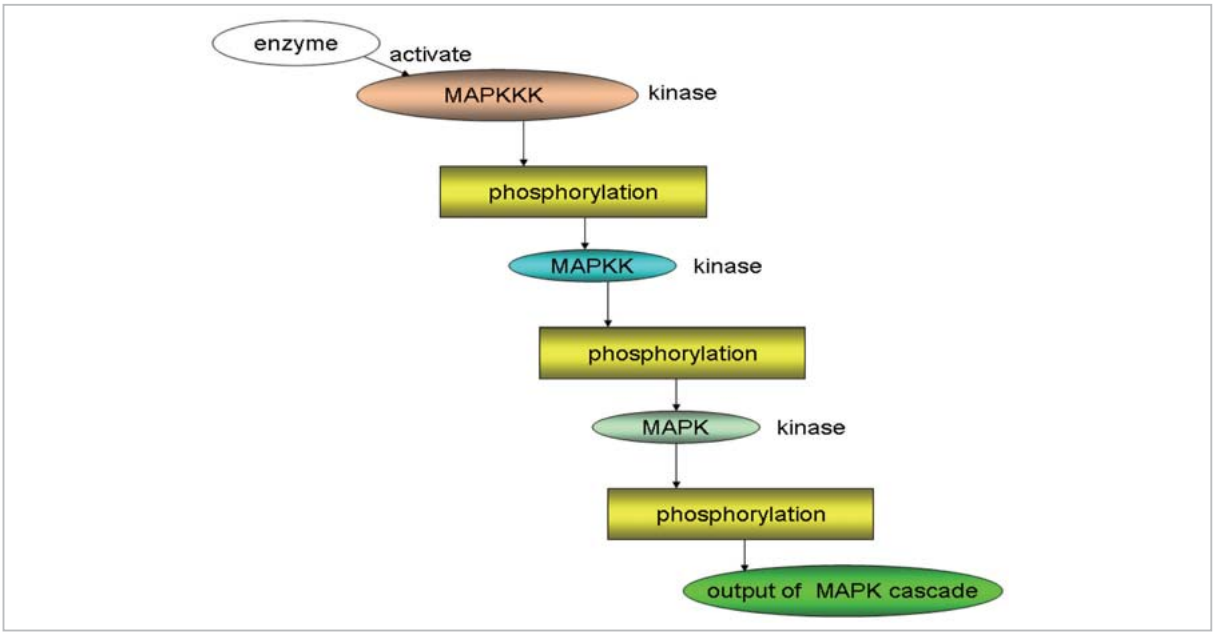


**Fig.3** *The hierarchical structure of MAPK cascade*

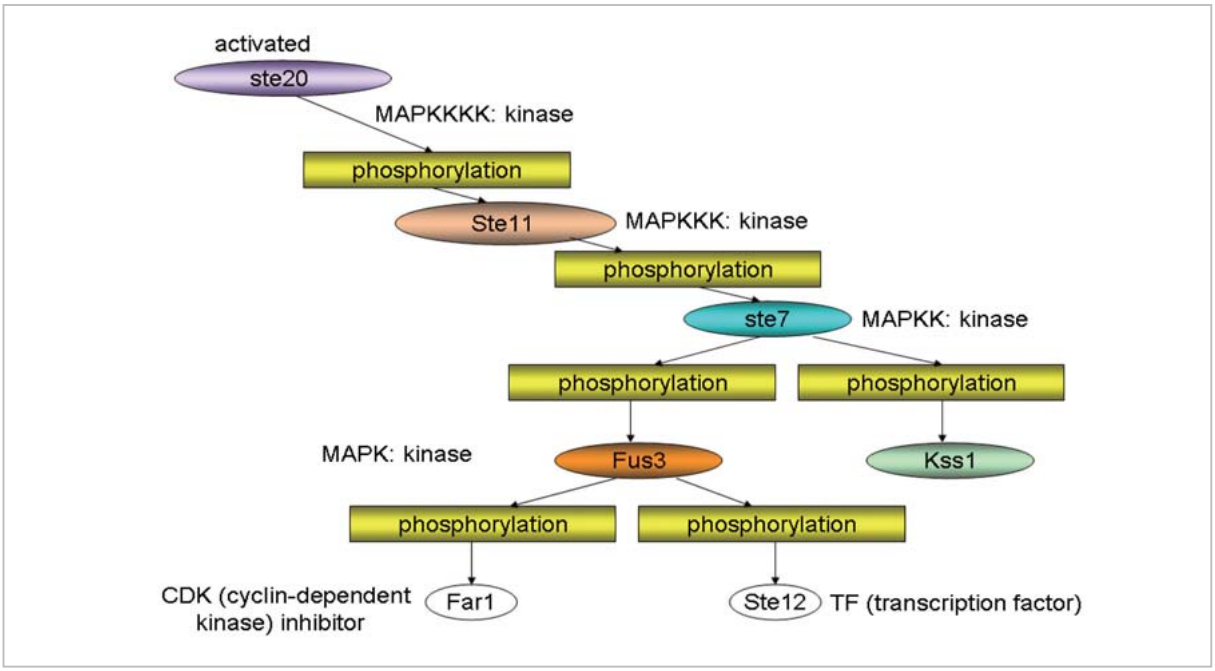


**Fig.4** *An example of a MAPK cascade*

Dependent Kinase) inhibitor, and Ste 2 is a TF (Transcription Factor) Ste 12. Other examples of signal pathways can be found in KEGG[6].

Up to now we have witnessed that the objects — nodes and links of pathway networks — can be represented by the vertexes and edges of graphs. Accordingly, the dynamics features of networks can be investigated based on the computational formulation given above.

## 3 Dynamical analysis of signal transduction networks

### 3.1 Temporal dynamics of signal transduction networks

Based on the graph structure we formulated in previous sections, we will discuss the dynamical features of signal transduction networks in order to systematically understand their information processing mechanism. Basically, the dynamics of signal transduction can be investigated by two major aspects: *spatial dynamics* and *temporal dynamics* of signal transduction. In signal transduction, the spatial dynamics is mainly reflected in diffusion processes. Kholodenko's review paper[7] on spatial dynamics uses the diffusion equation with polar coordinates to formulate the concentration values of kinases in the MAPK cascade constrained by the distance of diffusion.

The current paper focuses mostly on the information processing aspect of signal transduction networks, which clearly have a temporal character, so we limit our discussion to the temporal dynamics of signal transduction. Such dynamics is usually formulated by differential equations, among which the Michaelis-Menten equation mentioned in Section **2** is the most fundamental. The Michaelis-Menten equation describes the biochemical reaction among molecules used for signaling in cells. In the following, we will denote the concentration of any chemical X as [X], describing the number of molecules per unit volume.

A *substrate* X is the input to the pathway under the regulation of an *enzyme* E, and a *product* Y is the output of the pathway. The enzyme plays the role of catalyzer, which triggers the biochemical reaction and transforms the initial substrate into the resulting product.

$$S + E \underset{k_2}{\overset{k_1}{\rightleftarrows}} S \bullet E \xrightarrow{k_3} P + E$$

The expression S•E means that S is bound to E. The parameters $k_1$, $k_2$, and $k_3$ describe the conversion rates and

$$k_m = (k_2 + k_3) / k_1$$

which is called the Michaelis constant.

We can obtain a simple differential equation system for the kinetic dynamics of the product P, where $E_0$ is the total concentration of the enzyme.

$$d/dt\,[P] = k_3\,[E_0] * [X] / (k_m + [X]).$$

The above formulation, however, is only valid under quasi steady-state conditions, i.e., conditions in which the concentration of the substrate-bound enzyme changes at a much slower rate than those of the product and substrate. This allows the enzyme to be treated as a constant, which is $E_0$ in the formula.

A question arising naturally from here is how to use the above form to explain the MAPK cascade we already met before. So, we reformulate it as a set of coupled differential equations in which each stage in the cascade corresponds to one equation:

$$d/dt\,[MAPKKK] = k_3\,(MAPKK)\,[MAPKKK\,(0)] * [MAPKKK] / (k_m\,(MAPKKK) + [MAPKKK]),$$

$$d/dt\,[MAPKK] = k_3\,(MAPKK)\,[MAPKKK\,(0)] * [MAPKK] / (k_m\,(MAPKK) + [MAPKK]),$$

$$d/dt\,[MAPK] = k_3\,(MAPK)\,[MAPKK\,(0)] * [MAPK] / (k_m\,(MAPK) + [MAPK]),$$

d/dt [output-of-MAPKcascade] =
$k_3$ (output-of-MAPKcascade) [MAPK (0)] * [output-of-MAPKcascade] / ($k_m$ (output-of-MAPKcascade) + [output-of-MAPKcascade]),

With the appropriate parameters, usually obtained from empirical observations, the system MAPK cascade can be numerically calculated. Basically, the general behavior can be described as an amplification of the initial substrate concentration S(0), resulting in an enhanced signal with higher concentration of the resulting product P (final moment).

One of the well-known functions of the MAPK cascade in cellular signal transduction networks is to act as an amplifier for intracellular signaling processes. An unexpected phenomenon — a fixed-point that occurs at a four-layered MAPK cascade where a feedback is embedded — is observed by simulation[7], which shows that in theory the second messengers' signals can be kept at a constant value during their relay processes within cells.

## 3.2 Fixed point of pathways with feedbacks

Nonlinear dynamical features, such as bifurcation[3], have been reported in signal transduction networks. In order to study the computational and communication capacity of signal transduction networks, it is necessary to make sure how the cellular signals are controlled so that the information flow can be quantitatively measured. This framework forms the basis of the architectural design and performance analysis of engineered ICT systems inspired by signal transduction networks in cells. In this section, we take the fixed-point phenomena as an instance to demonstrate the nonlinear phenomena of signal transduction networks, as reported in[8].

### 3.2.1 The model for the simulation

As shown in Fig. 5 and Fig. 6, the structure of a MAPK cascade is layered, with feedback being embedded into each layer.

The set of coupled differential equations in the case feedback is present then becomes:

d/dt [phosphor-protein in MAPKKK layer] = $-k_3$ [MAPKKK] [MAPKKK (0)] * [MAPKKK] / ($k_m$ (MAPKKK) + [MAPKKK]) $-\int$ [MAPKK] dt,

d/dt [phosphor-protein in MAPKK layer] = $-k_3$ (MAPKK) [MAPKKK (0)] * [MAPKK] / ($k_m$ (MAPKK) + [MAPKK]) $-\int$ [MAPK] dt,

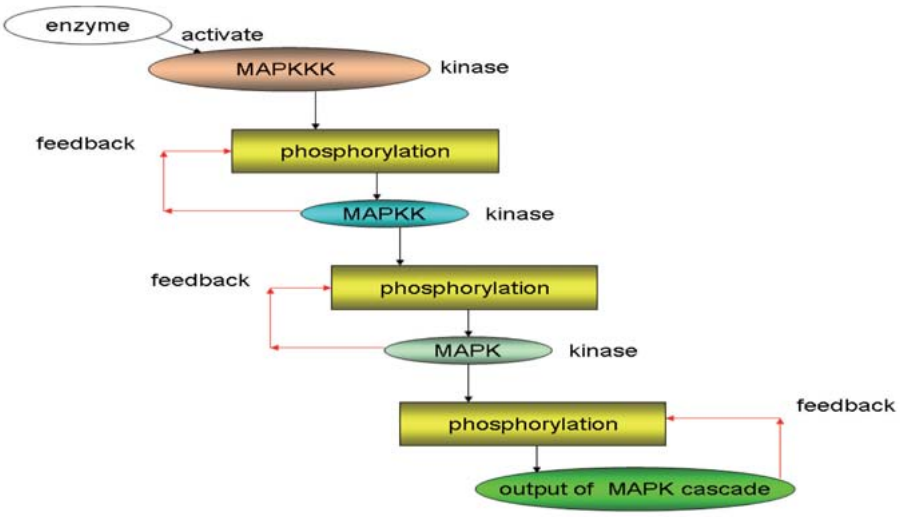


**Fig.5** *Three-layered MAPK cascade with feedback*

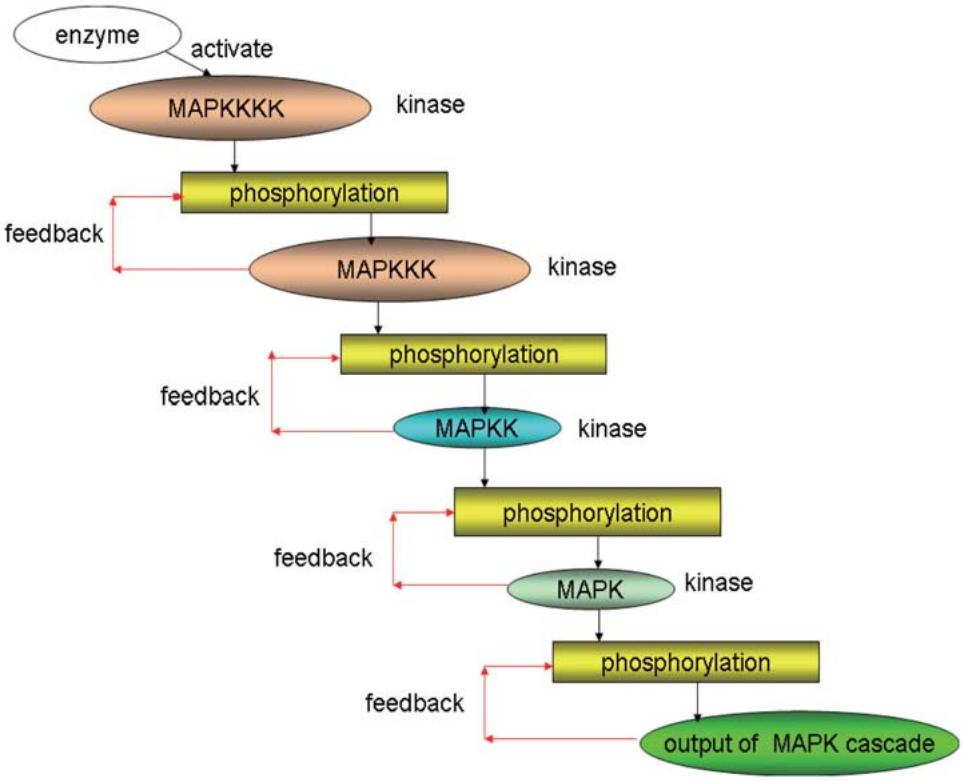


**Fig.6** *Four-layered MAPK cascade with feedback*

d/dt [phosphor-protein in MAPK layer] =
$-k_3$ (MAPK) [MAPKK (0)] * [MAPK] /
($k_m$ (MAPK) + [MAPK])
$-\int$ [output-of-MAPK-cascade] dt,

d/dt [MAPKKK] = $k_3$ (MAPKKK)
[MAPKKKK (0)] * [MAPKKK] /
($k_m$ (MAPKKK) + [MAPKKK]),

d/dt [MAPKK] = $k_3$ (MAPKK)
[MAPKKK (0)] * [MAPKK] /
($k_m$ (MAPKK) + [MAPKK]),

d/dt [MAPK] = $k^3$ (MAPK) [MAPKK (0)] *
[MAPK] / ($k_m$ (MAPK) + [MAPK]),

d/dt [output-of-MAPKcascade] =
$k_3$ (output-of-MAPKcascade) [MAPKK (0)] *
[output-of-MAPKcascade] /
($k_m$ (output-of-MAPKcascade) +
[output-of-MAPKcascade]),

where the integral part is calculated from initial time 0 to the current time t.

### 3.2.2 Simulation

Owing to the fact that the essential process of intracellular communication exhibits non-linear dynamical behaviors in a biochemical framework in the model above, it is possible to figure out what kind of parameters of Michaelis-Menten kinetics behind the fixed-point phenomena is used by the molecular mechanism of signal transduction networks through the MAPK cascade.

The conditions are set as follows:

- The initial concentration of the enzyme = 0.45.
- In the MAPKKKK, MAPKKK, MAPKK, MAPK layer, the product acts as the kinase for the succeeding MAPKKK, MAPKK, MAPK layers and output of the entire MAPK cascade at its bottom.
- In each layer of MAPK cascade $k_m = 0.1$ and $k_3 = 0.01$,

- The initial concentration of substrates (phosphor-proteins) in all the pathway-like units is set as 0.45.
- The step/sample number is 10
- The initial concentration of product is set as 0.001.

Now follows a simulation of a 4-layered MAPK cascade

Let y = f (x) denote the signaling process from MAPK denoted as x to protein of the entire MAPK cascade y, we observed that

$$f(0.0001030) = 0.0001030.$$

where x = y = 0.0001030.

This is a fixed-point-like phenomenon. The value 0.0001030 determines the crucial point of the phase transition between the monotonic decreasing mode and the fixed-point-like mode.

### 3.3 Robustness

At the system level of a pathway network, dynamical features of pathway networks are the key to understand the cellular signaling mechanism. One of the most important dynamical features of pathway networks is robustness. The robustness of a pathway network can be investigated through different means, for example, stability analysis is an efficient one in the case of the Mos-p MAPK cascade pathway, where Mos-p is a kinase/protein that is in the phosphorylation state, whereas Mos is the same protein in the dephosphorylation state.

The cell has high robustness against external disturbances. As Kitano[9] points out, cancer is an example of robustness in a cell. In contrast with the robustness of pathway networks in the cell, conventional communication networks are very fragile, for example, the Achelles' heel phenomenon in internet networks occurs when failures occur. This contrast motivates us to quantitatively describe the biological robustness[10] of pathway networks in order to investigate the possibility of applying the knowledge of the biological robustness in pathway networks to the design of information networks in the future.

In the previous sections, we discussed the graph structure that is usual in computer science. It is obvious that the robustness of cellular signaling processes described in terms of nonlinear dynamics is tightly connected with the "dynamical" graph structure of pathway networks. In general, the parameters of the differential equations that describe the biochemical reactions of pathway networks can be defined as labels that correspond to the graph, where the vertices and edges of the graph are defined for the pathway network in the previous section. These topics belong to the field called *Dynamical Networks*, which refers to the integration of nonlinear dynamics and graph theory. In this section we define nonlinear dynamical systems in a matrix formulation that corresponds to the graph of the pathway network.

Based on these schemes, we can formulate a robustness mechanism of cellular signaling pathway networks.

#### 3.3.1 Basic concepts

The concept of robustness is defined as follows.

*Definition of Robustness:*

Robustness refers to a mechanism that can guarantee and realize the state transition of a (usually dynamic) system from vulnerable and unstable states to sustainable (stable) states when the system suffers from disturbance that is outside the environment and that is unexpected in most cases[11].

Let us define a nonlinear system W to describe the cellular signaling mechanism in cells:

W (X, S, U,Y, Q)

where

X — input to the system
Y — output of the system
S — the states of the system
U — feedback

Q — parameter vector

The system description is given as follows:

$$d/dt\,(X) = A\,X + B\,S + C\,U$$
$$Y = D\,S$$

where A, B, C, D are constant matrixes, described by Q.

All the above-mentioned variables are vectors.

We focus on the definition of robustness in terms of the system state. Then we define the robustness by the function G (.) satisfying the condition that

$$S = G\,(S, E)$$

when the system suffered disturbance E

The robustness feature of a pathway is expressed in several different aspects of pathway networks. Stability is among them. The above-mentioned formulation makes it possible to use the states of the system for describing the robustness where the robustness mechanism is interpreted as the mechanism for providing the steady states. The Mos-p MAPK pathway[12] is an example explaining/analyzing the pathway stability for the robustness of the corresponding pathway network. In this pathway, the term "stable steady states" denoted as SS is used to describe the state transition of the pathway network within a certain domain[12]. The Mos-p MAPK pathway includes a MAPK cascade. As Huang et al[13] reported, the nonlinear dynamical features of different phosphorylation processes in a MAPK cascade vary, i.e., the phosphorylation concentration versus time curve is different for each layer of a MAPK cascade. The significance of this phenomenon is obvious if we reveal the fixed-point of the MAPK cascade presented in the previous section.

### 3.3.2 Stability analysis: From an example of the Mos-p MAPK pathway for explanation of stability

The Mos-p MAPK pathway that demonstrates the transition between the stable/unstable steady states in cells is reported in[12]. In order to make the formulation of the corresponding model, we need the *Hill coefficient* and the *Hill equation*.

The Hill equation is given as

$$\delta = ([L]^n) / (K_d + [L]^n)$$
$$= [L]^n / (K_A{}^n + [L]^n),$$

where

$\delta$ is the concentration of the phosphorylated protein,
[L] is the ligand concentration,
$K_d$ is the equilibrium dissociation constant,
$K_A$ is the ligand concentration occupying half of the binding sites,
and n is the Hill coefficient denoting the cooperativity of binding. It describes the nonlinear degree of the product's response to the ligand.

The cooperativity indicates the degree of the biochemical reactions for binding the substrate and enzyme. Here ligand means an enzyme protein that can bind with another molecule. In a Mos-p MAPK pathway, Mos-p is the ligand.

The coefficient n mentioned above is normally denoted as $n_{Hill}$.

$n_{Hill} = 1$ refers to the case of Michaelis-Menten kinetics, which corresponds to the case of completely independent binding, regardless of how many additional ligands are already bound.

$n_{Hill} > 1$ shows the case of positive cooperativity,

$n_{Hill} < 1$ shows the case of negative cooperativity.

By the above, the SS state can be quantitatively described. When the SS is achieved, the activation degree (concentration) of Mos-p can be formulated as a fixed-point of the following form:

$$\text{Mos-p} = f\,(\text{Mos-p}),$$

where f (.) refers to the Mos-p MAPK cascade pathway.

This shows the steady state of Mos-p, which is consistent with the phenomena reported by Ferrell et al.[12].

Based on the ultrasencitivity quantization, we have a closer look at the Mos-p MAPK pathway from two viewpoints:

#### 3.3.2.1 A brief look at the system

We consider the input of the system as the proteins Mos-P and malE-Mos, and the output as the protein MAPK. Here two signals concerning MAPK are involved — activated MAPK and phosphorylated MAPK (phos • MAPK for short).

Based on the results from the experiment reported in[12], we can use the Michaelis-Menten equation to obtain that the response of Mek is monotonic to Mos-P, to malE-Mos, or to Mos-P and malE-Mos. Additionally, the response of activate • MAPK or phos • MAPK is monotonic to the systems' inputs Mos-P or to malE-Mos or to Mos-P and male-Mos, under the condition that there is no feedback in the pathway networks, which are branched at the routes from Mos-P and malE-Mos to Mek. The phosphorylation effect can be witnessed at Mek, MAPK and Mos-P.

A graph description for this pathway network is given in Fig. 6

Considering the robustness again, we realize that we need the feedback (Cf. Fig. 7) and related nonlinear cellular signaling mechanism to help us understand the robustness within this pathway network.

#### 3.3.2.2 Modeling the temporal dynamics of the system

Let us use the Hill-coefficient-based formula to describe the temporal behavior of cellular signaling. Assuming the time series of

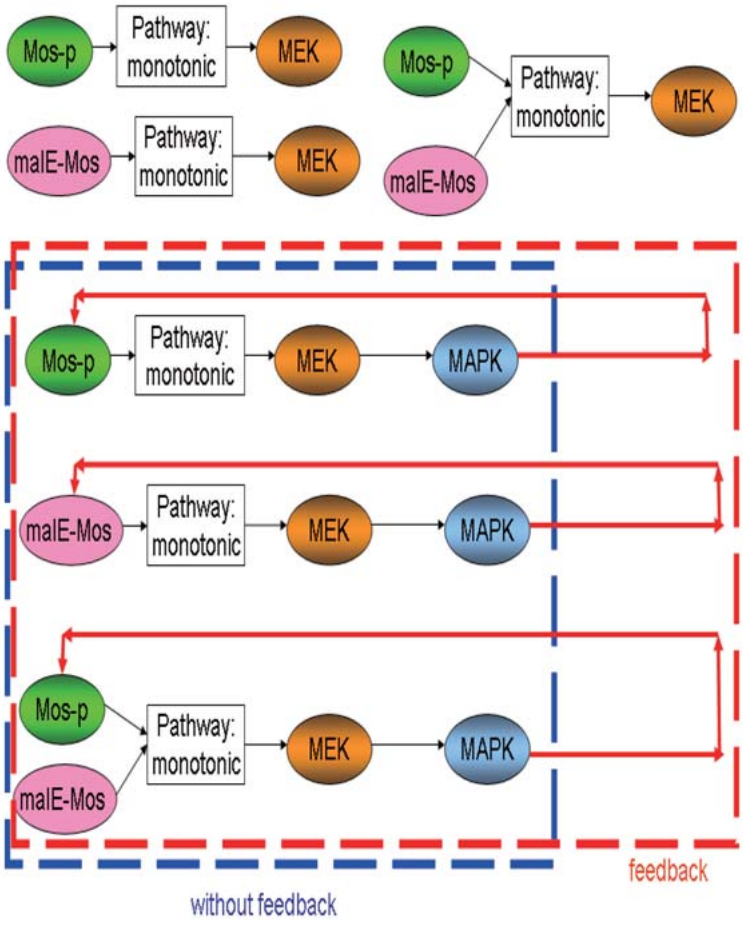


**Fig.7** Cascade without feedback and with feedback

malE-Mos as

$$d/dt\ [moleE\text{-}Mos] = f1\ ([molE\text{-}Mos], coefficients)$$

where f1 () takes the form of a monotic function with a decreasing order, we obtain

$$[moleE\text{-}Mos]\ (n+1) - [molE\text{-}Mos]\ (n) = coefficient\ [molE\text{-}Mos]\ (t)$$

This is the time difference equation used for numerical calculation.

Let m denote the step, m is an integer and the [mole-Mos] is set as 0. Then we obtain

$$[active\ MAPK\ tot]\ (m+1) = ([Mos\text{-}P]\ (m) + 1000)^H / (EC50^H + ([Mos\text{-}P]\ (m) + 1000)^H)$$

where

H refers to the Hill coefficient nH = 5
EC50 = 20 (nM)
[molE-Mos] (t = 0) is set as 1000

Then

$$[Mos\text{-}P]\ (m+2) = 0.5 * [active\ MAPK]\ (m+1) = 0.5 * [MAPK\ (tot)]\ ([Mos\text{-}P]\ (m) + 1000)^H / (EC50^H + ([Mos\text{-}P]\ (m) + 1000)^H)$$

Now assume that [Mos-P] (m+3) = [Mos-P] (m+2) + [Mos-P] (m). Let the moment n+1 correspond to m+3, n correspond to m when we set the time reference of Mos-P according to the initial time of going through the pathway and the final time. The moments of m+1 and m+2 refer to the internal states of the pathway network as a system. Consequently we have

$$[Mos\text{-}P]\ (n+1) = 0.5 * ([Mos\text{-}P]\ (n) + 1000)^H / (EC50^H + ([Mos\text{-}P]\ (n) + 1000)^H) + [Mos\text{-}P]\ (n)$$

The Mos-P MAPK pathway derived computing process then becomes

$$[x\ (n+1)] - [x\ (n)] = 0.5 * ([x\ (n)] + 1000)^5 / (20^5 + ([x\ (n)] + 1000)^5) - [x\ (n)]$$

where

x is an integer > = 0,
[x (0)] is set as 5.

Let $f((x\ (t)) = 0.5 * (x\ (t) + 1000)^5 / (20^5 + (x\ (t) + 1000)^5) - x\ (t)$

$$d/dt\ (x\ (t)) = x\ (t+1) - x\ (t) = f((x\ (t))$$

We define a function in general:

$$d/dt\ (x\ (t)) = f((x\ (t))$$

when [molE-Mos] is treated as a constant.

It is obvious that the above system is still nonlinear even though the control input [molE-Mos] is a constant.

Considering the dynamical feature of control input [molE-Mos], we have that

$$d/dt\ (x\ (t)) = 0.5 * (x\ (t) + u\ (t))^5 / (20^5 + (x\ (t) + u\ (t))^5) - x\ (t)$$

where x (t) = Mos-p refers to the state of the system and u (t) = [molE-Mos] refers to the control input.

Since this is a coupled system, it is necessary to decouple the different signals in order to efficiently control the system.

The description of the pathway network as a system is normally established by a differential equation. Denote the system by the following equations

$$d\ [X\ (t)] / dt = f\ (t, X\ (t), U\ (t))$$
$$Y\ (t) = g\ (t, X\ (t))$$

where t is time, U (t) is the input, X (t) is the state, and Y (t) is the output. This is a general

form that includes the case of nonlinear systems.

Based on the instance of Mos-p MAPK pathway we discussed before, the cellular pathway network is modeled as a controller-centered system where feedback is embedded. Different constraints can be used to formulate specific objects in applications, e.g., a matrix-based representation could be

$$dX(t)/dt = A(t)X(t) + B(t)U(t)$$
$$Y(t) = C(t)X(t)$$

where A (t), B (t) and C (t) are matrixes.

In this instance, the variable is one dimensional and a nonlinear relation exists between state transitions. So we have that

$$d/dt\,(x(t)) = f(x(t), u(t))$$

where

x (t) = [Mos-p], this is the state;

u (t) = [molE-Mos], this is the control input; this equation is established under the steady states of the Mos-p MAPK pathway.

The output for detecting the signals of phosphorylation is given as follow:

$$y(t) = phos \bullet [MAPK] = g((x(t), u(t))$$

where

$$g((x(t), u(t)) = 0.5 * (x(t) + u(t))^3 / (20^5 + (x(t) + u(t))^3)$$

under the condition that the Mos-p MAPK pathway is in the steady state.

From the above discussion of pathway systems, we have presented a dynamics-based representation for formulating the robustness of a dynamical system, which is motivated by the biological robustness in cellular pathway networks, but which can be modified and extended to any abstract dynamical system where feedback is embedded.

## 4 Error correcting codes for cellular signaling pathways

How can we carry out reliable information processing by pathway networks with dynamical features? An important element underlying cellular signaling is the robustness of molecular pathways. Mechanisms such as those resembling error correcting codes may play an essential role in this framework.

### 4.1 Molecular coding for molecular communication

A reversible molecular switch is the basis of information representation in cellular informatics as shown in Fig. 7. Two kinds of reversible molecular switches exist in cells.

*The molecular switch of phosphorylation and dephosphorylation:* The phosphorylation state of a signaling protein is defined as 1, whereas the dephosphorylation state of a signaling protein is defined as 0. The phosphorylation process is regulated by kinase, the dephosphorylation process is regulated by phosphatase.

*The switch of GTP-bound and GDP-bound*: As shown in Fig. 8, the GTP-bound state of GTPase set by the so-called GEF and GEF pathway is defined as 1 and the GDP-bound state of GTPase set by the so-called GAP and GAP pathway is defined as 0.

The MAPK cascade consists of several phosphorylation processes. So multiple binary codes can be generated, e.g. four bits generated by a four-layered MAPK cascade (see Fig. 9).

The above molecular switches allow us to formulate a mathematical model for information processing based on an abstraction of the data structure corresponding to the signaling pathways.

The above framework only involves switches, and it does not include the redundancy typically associated with error correcting codes. If we intend to include such codes, then the resulting mechanisms should be compatible with the mechanisms encountered in cell communication.

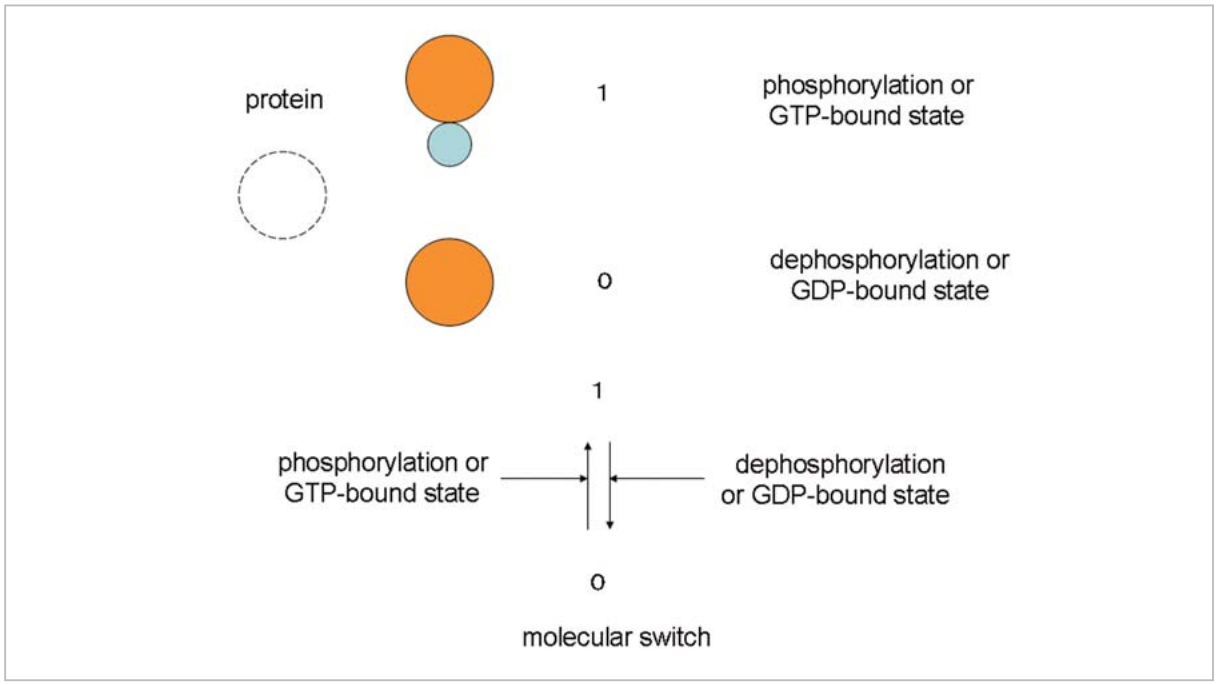


**Fig.8** *Molecular Switch for binary codes (bits)*

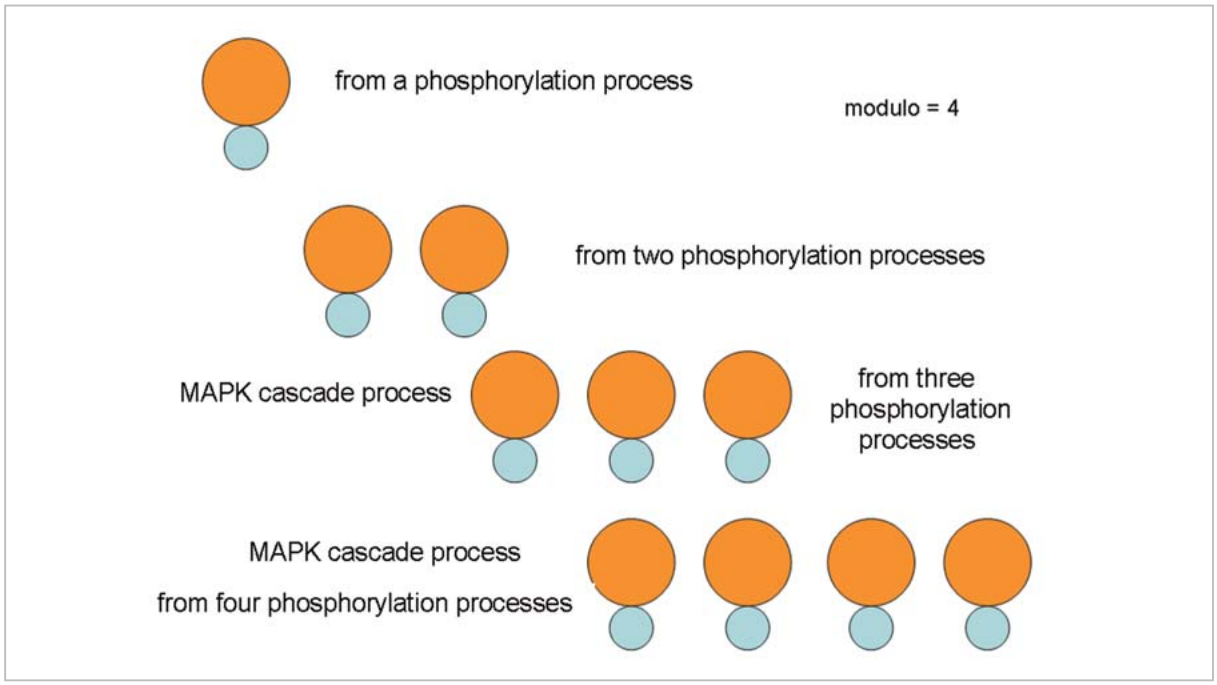


**Fig.9** *Four bits generated from a MAPK cascade*

### 4.2 LDPC coding for pathways

Error correcting codes are mathematical constructs, and their design involves a different philosophy from the way in which biological mechanisms have evolved, which includes an element of randomness. Fortunately, there is an error correcting code, of which the design---though mathematically well-founded---involves a random element as well. Called *Low Density Parity Code (LDPC)*, this code is, surprisingly, very efficient, in the sense that it allows transmission of information at rates very close to the theoretical maximum. These codes were developed in 1960 by Gallager, and they had been long forgotten due to the perception of them being impracti-

cal, only to be rediscovered by MacKay in 1996. Ironically, the randomness in LDPC codes is very attractive in the framework of biological systems, and this is the reason why we describe them in more detail.

The LDPC code is defined by a partite graph BG = <V, E>, where the vertex set $V = V_1 \cup V_2$. $V_1$ is an ordered set of nodes, each of which is labeled by a binary number. The set $V_1$ thus denotes a binary code word. $V_2$ is a set of which each node is labeled by a binary number equaling the parity value of the sum of the labels of the nodes in $V_1$ that are connected to the node in $V_2$ by an edge in E.

As shown in Fig. 10, at first let $V_{12}$ be 1 and $V_{21}$ be 0 (parity summation), then $V_{11}$ should accordingly be assigned the value 1. This is the encoding scheme. The decoding is carried out in a reverse way, as shown in Fig. 11.

Assume that $V_{11}$ is lost during information transmission, then in order to restore the value of $V_{11}$ we need to use the information of $V_{12}$ and $V_{21}$. Because $V_{12}$ is 1 and the parity summation is 0, we can infer that $V_{11}$ should be 1.

Figure 12 gives the encoding process of 5 bits in $V_1$ (labeled as $L_1$), where the known bits are put into the condition part of an "IF THEN" rule and the unknown bits are put into the conclusion part of this rule.

The nodes in $V_1$ are called $L_1$-units and the ones in $V_2$ are called $L_2$-units, owing to the fact that they are described by two separated vertex sets in a bipartite graph.

Within the signal transduction model followed in this paper, the two bipartite parts associated with the generation of an LDPC can be designed as shown in Fig. 13, where the information processing units corresponding to the phosphorylation/dephosphorylation pathway and GEF/GAP pathway are denoted as $L_1$-unit and $L_2$-unit, respectively.

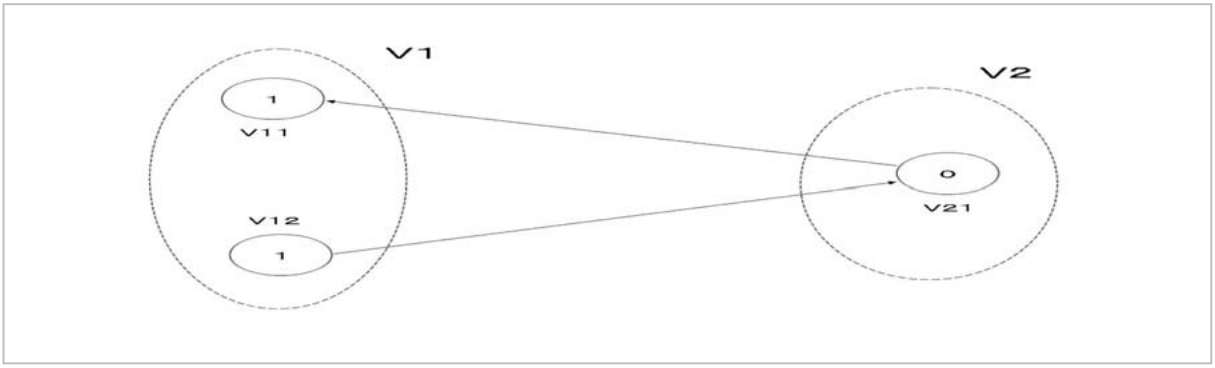


**Fig. 10** *A simple example of an LDPC encoding follows below*

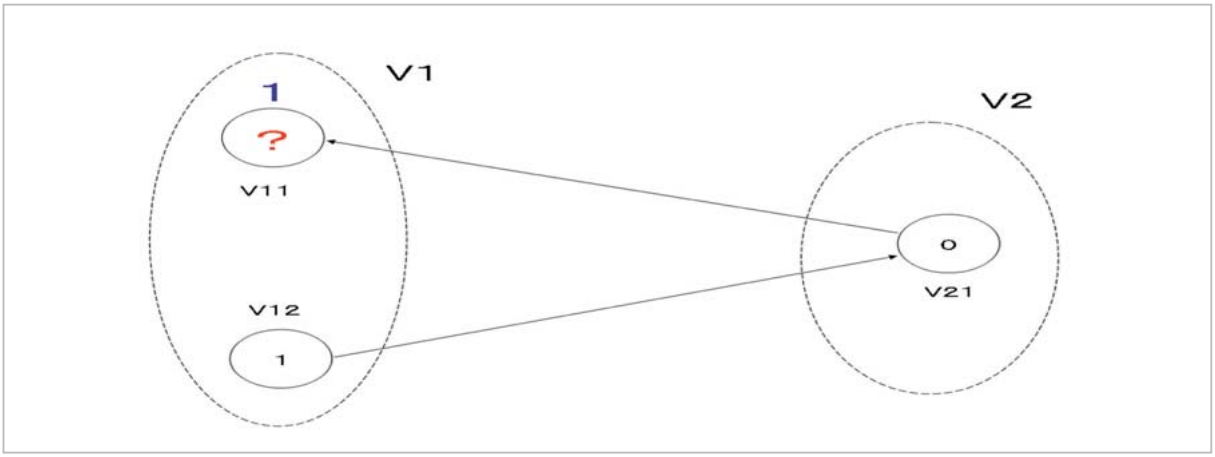


**Fig. 11** *A simple example of an error correcting LDPC code*

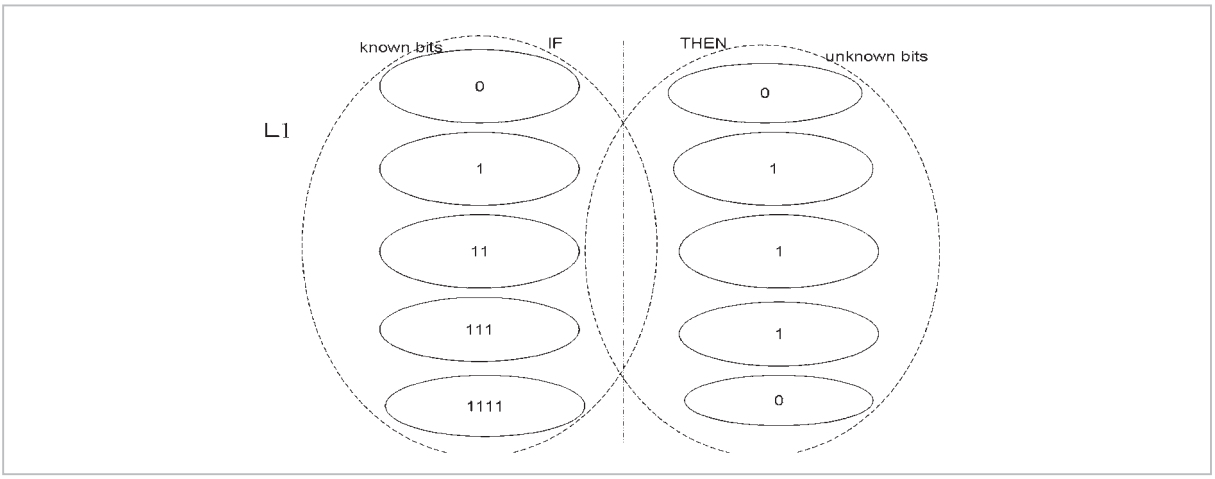


**Fig.12** *The rules for generating words*

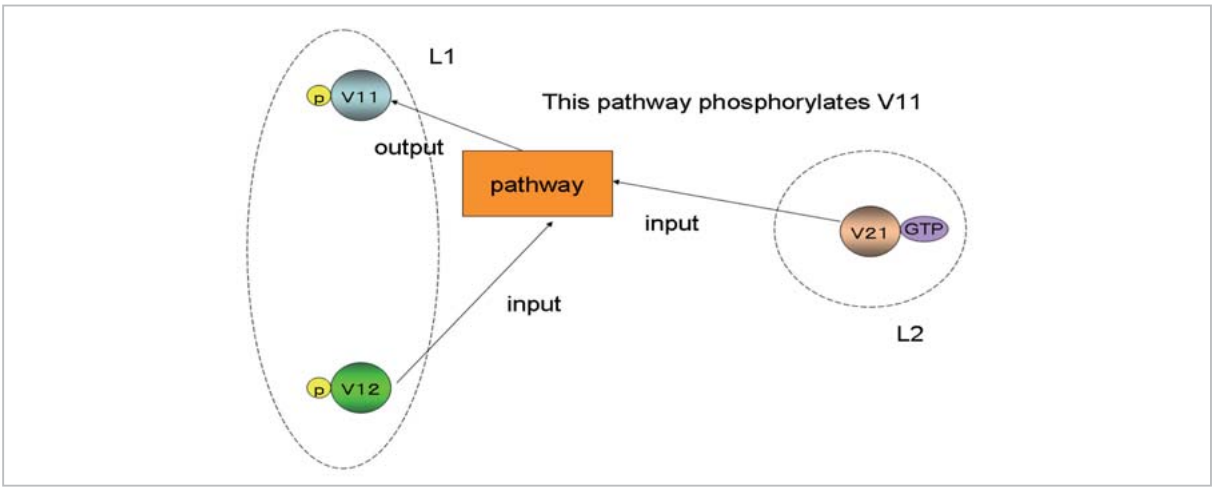


**Fig.13** *An example of a pathway that is mapped to the model in Fig. 11*

As shown in the above figure, the bipartite graph to encode an LDPC code is mapped from a pathway whose input is $V_{12}$ when in the phosphorylation state and $V_{21}$ when in the GTP-bound state and whose output is $V_{11}$ when in the phosphorylation state. This pathway corresponds to the graph in Figs. 10 and 11 for encoding and decoding an LDPC code.

The above model allows us to encode phosphorylation pathways and the dephosphorylation pathways in terms of an LDPC code, which is capable of approaching the Shannon limit. The derived encoding/decoding model is bidirectional, symmetrical and "implicitly-binary" (i.e., its binary form can be formulated in terms of n bits of the $L_1$-units of the model), which differs from the previous model given in[14] that is one-directional, asymmetrical (from phosphorylation/dephosphorylation to GTP-bound/GDP-bound states) and "explicitly-binary" (i.e., the phosphorylation/dephosphorylation mechanism directly encodes the code z).

## 5 Conclusions

In this paper we have described pathways in biological organisms that behave de-facto like communication systems on cellular scales. The robustness of biological communication systems offers an important lesson for the design of man-made communication systems: key concepts in this context are parallelism, adaptability, and structural stability. We have also sketched LDPC codes, which have much in common---due to their randomness---with the unstructured character of bio-

logical systems. This may suggest that LDPC codes form an important avenue of research in the realization of communication systems that include the above key concepts. Cellular signal transduction networks form a rich source of inspiration in the design of next-generation communication systems, which will have greater robustness, greater capacity, and greater adaptability, but which will also be less visible to its users, forming a transparent ever-present environment, such as the *overlay-network* in[15].

## *References*


1 A. L. Barabasi and E. Bonabeau, "Sacle-free Networks", Scientific American, Vol.288, No.5, pp.60-69, 2003.

2 R. Milo, et al., "Network Motifs: Simple Building Blocks of Complex Networks", Science, Vol.296, pp.910-913, 2002.

3 Hiroaki Kitano and Takehito Azuma, "Systems Biology and Control", ISCIE Journal 'Systems, Control and Information', Vol.48, No.3, pp.104-111, 2004 (in Japanese).

4 B. Alberts, et al., "The Molecular Biologyo of the Cell", Fourth Edition, New York: Garland Science, 2002.

5 Uri Alon, "An Introduction to Systems Biology: Design Principles of Biological Circuits", Boca Raton: Chapman Hill/CRC, 2006.

6 KEGG: Kyoto Encyclopedia of Genes and Genomes, http://www.kegg.jp/

7 B. N. Kholodenko, "Cell-signaling Dynamics in Time and Space", Nature Review Molecular Cell Biology, Vol.7, pp.165-176, 2006.

8 J.-Q. Liu, "On Quantitative Aspect of an Information Processing Model Inspired by Signaling Pathways in Cells: An Empirical Study", IPSJ SIG Technical Report, 2007-MPS-43, pp.21-24, 2007.

9 H. Kitano, "Cancer as a Robust System: Implications for Anticancer Therapy", Nature Reviews Cancer, Vol.4, pp.227-235, 2004.

10 H. Kitano, "Biological Robustness", Nature Reviews Genetics, Vol.5, pp.826-837, 2004.

11 J.-Q. Liu and K. Leibnitz, "Modeling the Dynamics of Cellular Signaling for Communication Networks", In: Bio-inspired Computing and Communication Networks, Y. Xiao, F. Hu (Eds.), Boca Raton: Auerbach Publications / CRC Press, in press, 2008.

12 J. E. Ferrell Jr. and E. M. Machleder, "The Biochemical Basis of an All-or-none Cell Fate Switch in Xnopus Oocytes", Science, Vol.280, 895-898, 1998.

13 C.-Y. F. Huang and J. E. Ferrell Jr., "Ultrasensitivity in the Mitogen-activated Protein Kinase Cascade", Proceedings of the National Academy of Sciences, Vol.93, pp.10078-10083, 1996.

14 J.-Q Liu and K. Shimohara, "Biomolecular Computation for Bionanotechnology", BostonlLondon: Artech House, 2007.

15 Z. Li and P. Mohapatra, "QRON: QoS-aware Routing in Overlay Networks", IEEE Journal on Selected Areas in Communications, Vol.22, pp.29-40, 2004.



***Jian-Qin Liu**, Ph.D. / Dr.Informatics*

*Expert Researcher, Kobe Advanced ICT Research Center, Kobe Research Laboratories*

*Bioinformatics, Network Informatics, Biomolecular Computation*



***Ferdinand Peper**, Ph.D.*

*Senior Researcher, Kobe Advanced ICT Research Center, Kobe Research Laboratories*

*Nanocomputing, Cellular Automata, Noise*