ON CONSTRUCTION OF PROBABILISTIC BOOLEAN NETWORKS

HAO JIANG *, XI CHEN †, AND WAI-KI CHING ‡

Abstract. The problem of constructing a sparse Probabilistic Boolean Network (PBN) from a prescribed positive stationary distribution is addressed in this paper. Boolean Networks (BNs) and its extension Probabilistic Boolean Networks (PBNs) stand as two groundbreaking models in the construction of genetic regulatory networks. The network construction problem is ill-posed and challenging. Here we present a novel method to construct a PBN based on a given stationary distribution. A series of sparsest transition probability matrices can be determined once giving the stationary distribution. Fixing the number of nonzero entries in each column of the transition probability matrix (number of transition states in the next step for every gene), the desired sparse transition probability matrix in the sense of maximum entropy can then be uniquely constructed as the linear combinations of the selected sparsest transition probability matrices. The problem of PBN construction for the transition probability matrix therefore can be effectively and efficiently handled utilizing the results obtained for the series of sparsest transition probability matrices.

Key words. Stationary Distribution, Transition Probability Matrix, Probabilistic Boolean Networks, Entropy, Sparsity.

1. Introduction. Rapidly evolving genomic technologies have paved the way for massive amounts of genomic data. A tremendous amount of mathematical and computing approaches have been used to glean the understanding of genetic regulatory networks over the past few decades. Directed graphs could be viewed as the most straightforward way to model a Genetic Regulatory Network (GRN). Bower and Bolouri introduced some classic models of genetic networks [8]. A Bayesian network [7] depicts the gene regulatory process from a probability perspective. In the dynamic-system perspective, differential equations were used to describe the change rate of expression levels. However, a lot more computation time is needed in simulation with much shorter time steps [15]. Discrete Dynamical System (DDS) Model [10], a discrete version of ODEs, assists one to understand the interactions among variables systematically. It has gained a solid foot in quantitative modeling of GRNs. A modified model taking into consideration of time delay effect was proposed as well [6]. From a logical standpoint, the expression of a gene in the network is quantized to be "ON" or "OFF" [9]. This may help in understanding key dynamic properties of a regulatory process. Boolean Networks (BNs) are a class of discrete dynamical systems in that genes interact with each other precisely determined by molecular interactions over a set of Boolean variables. Probabilistic Boolean Network (PBN) model [13, 14] is the extended version of BN that incorporates the stochastic characteristics of GRNs. Each gene is regulated through a set of Boolean functions with corresponding selection probabilities. The model combines deterministic functional aspects and the inherent probabilistic characteristics of complex systems. Given a PBN, the stationary distribution characterizes the network behavior. Efficient method for computing transition probability matrix and resulting stationary distribution was developed in [4, 17].

Network inference from steady-state data is essential in that most microarray data sets are presumed to be obtained from sampling the steady-state. Two algorithms have been proposed [11] to find attractors composing a BN. The problem was also efficiently solved in the sense of maximizing entropy [5]. Here we consider the inverse problem of constructing a PBN based on the prescribed positive stationary probability

*Department of Mathematics, The University of Hong Kong, Hong Kong (haohao@hkusuc.hku.hk)
†Department of Mathematics, The University of Hong Kong, Hong Kong (dlkcissy@hotmail.com)
‡Department of Mathematics, The University of Hong Kong, Hong Kong (wching@hku.hk)

State	$v_1(t)$	$v_2(t)$	$f^{(1)}$	$f^{(2)}$
1	0	0	0	0
2	0	1	1	0
3	1	0	0	1
4	1	1	1	0

Table 2.1: The Truth Table.

distribution. This is an interesting and challenging inverse problem of huge size. It can be formulated into two subproblems: (i) constructing a sparse transition probability matrix from a given positive stationary distribution (Problem 1) and (ii) constructing a PBN from the obtained sparse transition probability matrix (Problem 2). For Problem 2, a favorable result has been obtained through a α-norm addition to the objective function [2].

In this paper we will pay more attention to Problem 1. Given the positive stationary distribution, one can obtain a series of sparsest transition probability matrices. Fixing the number of nonzero entries in the column of transition probability matrix, we can obtain a unique solution among all the linear combinations of the sparsest transition probability matrices in the context of maximum entropy. Details will be elucidated in the following sections with proof. The construction of PBN is then tackled efficiently once knowing the PBN structures for the sparsest transition probability matrices respectively. In this framework, the inverse problem can be efficiently solved.

The remainder of the paper is structured as follows. In Section 2, a brief review on BNs and PBNs will be given. In section 3, we present the mathematical formulation of the inverse problem. Theoretic illustration and proof will be provided. Numerical experiments will be used to demonstrate the effectiveness of our proposed method. Finally, concluding remarks will be given in the last section.

2. A Review. In this section, we first give an introduction to BNs and PBNs. We then review some previous works related to PBN construction. For Problem 2, we present a maximum entropy approach in Section 2.1. In Section 2.2, a modified entropy approach is given which can address both Problem 1 and Problem 2. For Problem 2, since the transition matrix is sparse, for the rest of this section, we assume that the transition probability matrix is A with size $2^n \times 2^n$, and each column of A has at most m non-zero entries.

A Boolean Network (BN) $G(V, F)$ consists of a set of vertices $V = \{v_1, v_2, \ldots, v_n\}$ and a list of Boolean functions $F = \{f_1, f_2, \ldots, f_n\}$ where $f_i : \{0,1\}^n \to \{0,1\}$. Define $v_i(t)$ to be the state (0 or 1) of the vertex v_i at time t. The rules of the regulatory interactions among the genes are then represented by $v_i(t+1) = f_i(\mathbf{v}(t)), \quad i = 1, 2, \ldots, n$ where $\mathbf{v}(t) = (v_1(t), v_2(t), \ldots, v_n(t))^T$ is called the Gene Activity Profile (GAP). The GAP can take any possible forms (states) from the set $S = \{(v_1, v_2, \ldots, v_n)^T : v_i \in \{0,1\}\}$ and thus totally there are 2^n possible states in the network. It is known that eventually a BN will enter into a cycle (attractor cycle) and stay there forever. The cycles actually have biological significance such as cell proliferation and apoptosis.

The following is an example of a two-gene BN (taken from [3]) with its truth table given in Table 2.1. From the table we see that if the current network state is 1 then it will make a transition to itself in one step. The next transition step is state 3 if the current state is either 2 or 4. Finally if the current state is 3, the state in the

next step will be 2. The transition probability matrix (Boolean Network matrix) of the 2-gene BN is then given by

$$B = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \tag{2.1}$$

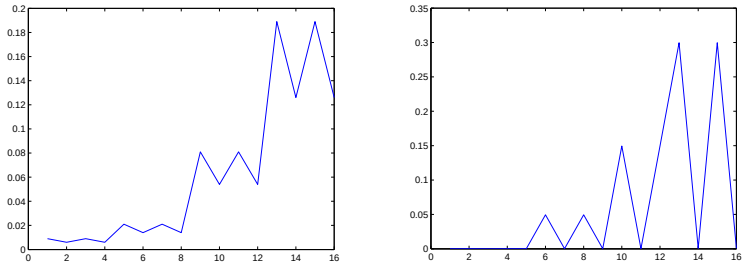
Since the network is deterministic, each column in B has only one non-zero element and the column sum is one. We remark that there is an one-to-one relation between a BN and its corresponding BN matrix.

To overcome the deterministic rigidity of a BN, extension to a probabilistic model is natural. To extend the BN model to a stochastic model, for each vertex v_i in a PBN, instead of having only one Boolean function as in the case of a BN, there are a number of Boolean functions (predictor functions) $f_j^{(i)}(j = 1, 2, \ldots, l(i))$ to be chosen for determining the state of gene v_i. Here $l(i) \leq 2^{2^n}$ and $l(i)$ is the total number of possible Boolean functions of gene i available. Then there are $N = \prod_{i=1}^{n} l(i)$ different possible realizations of BNs. The probability of choosing $f_j^{(i)}$ as the predictor function is $c_j^{(i)}$ where $0 \leq c_j^{(i)} \leq 1$ and $\sum_{j=1}^{l(i)} c_j^{(i)} = 1$ for $i = 1, 2, \ldots, n$. Since there are N possible realizations of BNs and they are characterized by N vector functions $f_1, f_2, \ldots, f_N$ ordered lexicographically. Here $f_1 = (f_1^{(1)}, f_1^{(2)}, \ldots, f_1^{(n)})$ is the first vector function for the first BN and $f_N = (f_{l(1)}^{(1)}, f_{l(2)}^{(2)}, \ldots, f_{l(n)}^{(n)})$ is the last vector function for the Nth BN. Then in an independent PBN (the selection of the Boolean function for each gene is assumed to be independent), the probability of choosing the kth BN having the vector function $(f_{k_1}^{(1)}, f_{k_2}^{(2)}, \ldots, f_{k_n}^{(n)})$ is given by $q_k = \prod_{i=1}^{n} c_{k_i}^{(i)}$, $k = 1, 2, \ldots, N$. We note that the transition process among the states in the set S is a Markov chain process and the transition probability of the Markov chain can also be obtained [4].

2.1. A Maximum Entropy Approach. For Problem 2, we have a maximum entropy approach [1]. We remark that the method of maximum entropy has been applied to the determination of unknown parameters from incomplete data such as traffic demand estimation in a transportation network [16]. A detailed discussion on entropy theory can be found in [12]. As we mentioned before, because the entries in $\mathbf{q}$ and A are non-negative, assuming that the $2^n \times 2^n$ transition probability matrix A has at most m non-zero entries, there are at most m^{2^n} BNs constituting this PBN. We label the transition probability matrices of these BNs by $A_1, A_2, \ldots, A_{m^{2^n}}$ systematically. To construct the PBN, the $\mathbf{q}$ has to satisfy the following constraints: (i) $\sum_{i=1}^{m^{2^n}} q_i A_i = A$ (ii) $0 \leq q_i \leq 1$ and $\sum_{i=1}^{m^{2^n}} q_i = 1$. Usually, there are too many feasible solutions. We then adopt entropy as the measurement to narrow down the solution set or even get an unique (optimal) solution. Define

$$F\left(\begin{pmatrix} a_{11} & \cdots & a_{1l} \\ \vdots & \vdots & \vdots \\ a_{l1} & \cdots & a_{ll} \end{pmatrix}\right) = (a_{11}, \ldots, a_{l1}, a_{12}, \ldots, a_{l2}, \ldots, \ldots, a_{1l}, \ldots a_{ll})^T. \tag{2.2}$$

and we let $U = [F(A_1), F(A_2), \ldots, F(A_{m^{2^n}})]$ and $\mathbf{p} = F(A)$. To ensure that $\sum_{i=1}^{m^{2^n}} q_i = 1$, we add a row of $(1, 1, \ldots, 1)$ to the bottom of the matrix U and form a new matrix $\bar{U}$. Meanwhile, we add an entry 1 at the end of the vector $\mathbf{p}$ to get a new vector $\bar{\mathbf{p}}$.



(a) maximum entropy approach (b) modified entropy approach

Fig. 2.1: The probability distribution $\mathbf{q}$

The maximum entropy algorithm can be formulated as follows.

$$\max_{\mathbf{q}} \left\{ \sum_{i=1}^{m^{2^n}} (-q_i \log q_i) \right\} \tag{2.3}$$

subject to $\bar{U}\mathbf{q} = \bar{\mathbf{p}}$ and $0 \leq q_i \ i = 1, 2, \ldots, m^{2^n}$. Newton's method in conjunction with CG method can be applied to solve the above problem. As a demonstration, we consider a PBN with $n = 2$ and $m = 2$. Suppose the observed transition probability matrix is given as follows:

$$A_{2,2} = \begin{pmatrix} 0.1 & 0.3 & 0.5 & 0.6 \\ 0.0 & 0.7 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.5 & 0.0 \\ 0.9 & 0.0 & 0.0 & 0.4 \end{pmatrix}. \tag{2.4}$$

Using our maximum entropy approach, we obtain the solution as shown in Fig. 2.1a. Here we see that this method can be used to identify the major components of the BNs constituting the PBN.

2.2. A Modified Entropy Approach. In this subsection, we introduce another method, namely a modified entropy approach. By adding an α-norm to the objective function (2.3), one can generate a PBN with a more sparse $\mathbf{q}$. The new objective function for Problem 2 is

$$\max_{\mathbf{q}} \left\{ -\sum_{i=1}^{m^{2^n}} q_i \log q_i - \beta \sum_{i=1}^{m^{2^n}} q_i^{\alpha} \right\} \tag{2.5}$$

The first term is the entropy as in (2.3) and the second term is the L_α-norm part which helps in getting a sparse solution $\mathbf{q}$. Here α and β are two parameters. In practice, we adopt grid search method to find optimal values of α and β. For Problem 2, we use the example given in (2.4) to demonstrate this method and compare with the maximum entropy method as well. Using this modified entropy approach, we obtain the solution as shown in Fig.2.1b. The optimal solution is reached when $\alpha = 0.63$ and $\beta = 1.40$. We can see that the modified entropy approach can get a more sparse solution. For Problem 1, the superiority of the modified entropy method can be ensured through reference to [2].

3. Methodology. In this section, we would like to provide the mathematical formulation of the inverse problem. Since it can be divided into two subproblems : (i) constructing a sparse transition probability matrix from the stationary distribution and (ii) approximating the sparse transition probability matrix with PBNs, we will study them one by one.

3.1. Construction of Sparse Transition Probability Matrix. Assume the number of genes is n, $\boldsymbol{\pi} = (\pi_1, \pi_2, \ldots, \pi_N)$ is a positive stationary probability distribution and $N = 2^n$. Once knowing the stationary distribution $\boldsymbol{\pi}$, we can obtain $N-1$ sparsest transition probability matrices ($t_i = (\frac{1}{\pi_i})/(\sum_{i=1}^{N} \frac{1}{\pi_i}), i = 1, 2, \ldots, N$):

$$T_1 = \begin{bmatrix} 1-t_1 & 0 & \cdots & 0 & 0 & t_N \\ t_1 & 1-t_2 & \cdots & 0 & 0 & 0 \\ 0 & t_2 & \ddots & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & t_{N-2} & 1-t_{N-1} & 0 \\ 0 & 0 & \cdots & 0 & t_{N-1} & 1-t_N \end{bmatrix}$$

$$\vdots$$

$$T_{N-1} = \begin{bmatrix} 1-t_1 & t_2 & \cdots & 0 & 0 & 0 \\ 0 & 1-t_2 & t_3 & 0 & 0 & 0 \\ 0 & 0 & 1-t_3 & t_4 & 0 & 0 \\ 0 & 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 & 1-t_{N-1} & t_N \\ t_1 & 0 & \cdots & 0 & 0 & 1-t_N \end{bmatrix}$$

Next, we consider the linear combinations of the $N-1$ sparsest transition probability matrices. Suppose we fix the number of nonzero entries in each column of the transition probability matrix k, we are interested in getting desired matrix in the sense of maximizing entropy. Mathematically speaking, we are interested in the solution to the following problem:

$$\max_{p_{ij}} \left\{ -\sum_{j=1}^{N} \pi_j \left(\sum_{i=1}^{N} p_{ij} \log p_{ij} \right) \right\}$$

subject to

$$\begin{cases} \sum_{i=1}^{N} p_{ij} = 1, \ j = 1, 2, \ldots, N \\ P\boldsymbol{\pi} = \boldsymbol{\pi}, P = [p_{ij}], p_{ij} \geq 0, \ i, j = 1, 2, \ldots, N. \\ P = \sum_{j=1}^{k-1} q_{i_j} T_{i_j} \\ q_{i_j} > 0, \sum_{j=1}^{k-1} q_{i_j} = 1, \ \{i_1, i_2, \ldots, i_{k-1}\} \subset \{1, 2, \ldots, N\} \end{cases} \tag{3.1}$$

THEOREM 3.1. *Fixing the number of nonzero entries in each column of the transition probability matrix k, we can have C_{N-1}^{k-1} optimal solutions.*

Proof. There are $N-1$ sparsest transition probability matrices. We also note that for all the $N-1$ matrices, the number of nonzero entries in each column is 2. Together with the particular characteristics embedded in these matrices we need exactly $k-1$ matrices to make a transition probability matrix with k nonzero entries in each column. We prove that the uniformly(probability $\frac{1}{k-1}$) linear combination of any $k-1$ matrices satisfy the above optimization problem.

Assume $P = \sum_{j=1}^{k-1} q_{i_j} T_{i_j}$, then the entropy can be expressed as follows:

$$\begin{aligned}&-\sum_{j=1}^{N}\pi_j\sum_{i=1}^{N}p_{ij}\log p_{ij} = -\sum_{j=1}^{N}\pi_j[(1-t_j)\log(1-t_j)+\sum_{l=1}^{k-1}q_{i_l}t_j\log(q_{i_l}t_j)]\\ &= -\sum_{j=1}^{N}\pi_j[(1-t_j)\log(1-t_j)+t_j\log(t_j)] - \sum_{j=1}^{N}\pi_j t_j\sum_{l=1}^{k-1}[q_{i_l}\log(q_{i_l})].\end{aligned}$$

Since $\pi_i, t_i, i = 1, 2, \ldots, N$ are already known, they can be viewed as constant; apart from that, because of the convexity of the function $x\log(x)$, we can easily achieve the optimal solution when

$$q_{i_c} = \frac{1}{k-1} \quad \text{for} \quad c \in \{1, 2, \ldots, k-1\}.$$

This clearly states the irrelevance in selection of the $(k-1)$ matrices. Therefore, one can have C_{N-1}^{k-1} kinds of combinations as long as all the combination coefficients are $\frac{1}{k-1}$. This completes the proof. □

After proof for the above theorem, we know that for a fixed number k, there are C_{N-1}^{k-1} optimal solutions among all the $k-1$ linear combinations of the sparsest transition probability matrices with the same maximum entropy. If we regard the optimal entropy as a function of k, we have the following theorem.

THEOREM 3.2. *Optimal entropy for the transition probability matrix as a function of k (number of nonzero entries in each column) is a increasing function.*

Proof. From Theorem 3.1, we know the optimal entropy function $E(k)$ as follows:

$$\begin{aligned}&E(k) = -\sum_{j=1}^{N}\pi_j[(1-t_j)\log(1-t_j)+t_j\log(t_j)] - \sum_{j=1}^{N}\pi_j t_j\sum_{l=1}^{k-1}[\tfrac{1}{k-1}\log(\tfrac{1}{k-1})]\\ &= -\sum_{j=1}^{N}\pi_j[(1-t_j)\log(1-t_j)+t_j\log(t_j)] + [\sum_{j=1}^{N}\pi_j t_j]\log(k-1)\end{aligned}$$

and $E(k+1) - E(k) = [\sum_{j=1}^{N}\pi_j t_j]\log(1+\frac{1}{k-1}) > 0$.
This completes the proof. □

Since for a fixed number k in transition probability matrix construction, we can have C_{N-1}^{k-1} optimal solutions within all the possible $(k-1)$ linear combinations of the sparsest transition probability matrices. Among the C_{N-1}^{k-1} possibilities, we would like to choose the most preferable one, the one which maximizes the entropy rate of the Markov chain. The following theorem helps answering the question in that it has found the transition probability matrix with the largest entropy among all the transition probability matrices.

THEOREM 3.3. *The solution to the following optimization problem:*

$$\max_{p_{ij}}\left\{-\sum_{j=1}^{N}\pi_j\left(\sum_{i=1}^{N}p_{ij}\log p_{ij}\right)\right\}$$

subject to

$$\begin{cases}\sum_{i=1}^{N}p_{ij} = 1,\ j = 1, 2, \ldots, N\\ P\boldsymbol{\pi} = \boldsymbol{\pi},\ P = [p_{ij}]\\ p_{ij} \geq 0,\ i, j = 1, 2, \ldots, N.\end{cases} \tag{3.2}$$

is achieved when $p_{ij} = \pi_i,\ i, j = 1, 2, \ldots, N$.

Proof. The above problem is equivalent to the following minimization problem:

$$\min_{p_{ij}}\left\{\sum_{j=1}^{N}\pi_j\left(\sum_{i=1}^{N}p_{ij}\log p_{ij}\right)\right\}$$

with the same constraints. We can use Lagrange multiplier method, if we rewrite

$$\begin{cases} f = \sum_{j=1}^{N} \pi_j (\sum_{i=1}^{N} p_{ij} \log p_{ij}); \\ k_i = \sum_{j=1}^{N} p_{ij}\pi_j - \pi_i, \ i = 1, 2, \ldots, N; \\ h_j = \sum_{i=1}^{N} p_{ij} - 1, \ j = 1, 2, \ldots, N; \\ g_{ij} = -p_{ij}, \ i, j = 1, 2, \ldots, N. \end{cases}$$

Then we have for $\boldsymbol{\lambda} \in R^N, \mathbf{r} \in R^N, \boldsymbol{\mu} \in R^{N^2}$, such that

$$\begin{cases} \nabla f + \nabla \mathbf{h}\boldsymbol{\lambda} + \nabla \mathbf{k}\mathbf{r} + \nabla \mathbf{g}\boldsymbol{\mu} = \mathbf{0}; \\ \mathbf{g}\boldsymbol{\mu} = \mathbf{0}; \\ \boldsymbol{\mu} \geq \mathbf{0}; \end{cases}$$

where $\mathbf{g} = [g_{11}, g_{12}, \ldots, g_{1N}, g_{21}, g_{22}, \ldots, g_{2N}, \ldots, g_{NN}]$, $\mathbf{h} = [h_1, h_2, \ldots, h_N]$ and $\mathbf{k} = [k_1, k_2, \ldots, k_N]$. We note that $p_{ij} > 0$, so the non-negativity constraints are indeed inactive. Then we can have the following equations:

$$\begin{cases} \pi_j(1 + \log(p_{ij})) + \lambda_j + r_i\pi_j = 0 & i, j = 1, 2, , \ldots, N \\ \sum_{i=1}^{N} p_{ij} = 1 & j = 1, 2, \ldots, N \\ \sum_{j=1}^{N} p_{ij}\pi_j = \pi_i & i = 1, 2, \ldots, N. \end{cases}$$

We have $p_{ij} = e^{-1-\frac{\lambda_j}{\pi_j}-r_i}, \ i, j = 1, 2, \ldots, N$. Using the condition $\sum_{i=1}^{N} p_{ij} = 1, \ j = 1, 2, \ldots, N$, we get

$$e^{-1-\frac{\lambda_j}{\pi_j}} \sum_{i=1}^{N} e^{-r_i} = 1, \quad j = 1, 2, \ldots, N. \tag{3.3}$$

On the other hand, with $\sum_{j=1}^{N} p_{ij}\pi_j = \pi_i, \ i = 1, 2, \ldots, N$ we get

$$e^{-r_i} \sum_{j=1}^{N} \pi_j e^{-1-\frac{\lambda_j}{\pi_j}} = \pi_i, \ i = 1, 2, \ldots, N. \tag{3.4}$$

Using Equation (3.3), we know that $e^{-1-\frac{\lambda_j}{\pi_j}}$ is constant. If we define $e^{-1-\frac{\lambda_j}{\pi_j}} = C$, then from Equation (3.4) we have $e^{-r_i}C = \pi_i, \ i = 1, 2, \ldots, N$. This implies that $p_{ij} = e^{-r_i-1-\frac{\lambda_j}{\pi_j}} = e^{-r_i}C = \pi_i, \ i, j = 1, 2, \ldots, N$. We note that

$$\nabla^2 f + \sum_{i=1}^{N} \lambda_i \nabla^2 h_i + \sum_{j=1}^{N} r_j \nabla^2 k_j + \sum_{l=1}^{N} \sum_{m=1}^{N} \mu_{lm} \nabla^2 g_{lm} = Diag(\frac{1}{p})$$

where $p = [p_{11}, \ldots, p_{1N}, p_{21}, \ldots, p_{2N}, \ldots, p_{NN}]^T$ and $\frac{1}{p} = [\frac{1}{p_{11}}, \ldots, \frac{1}{p_{1N}}, \frac{1}{p_{21}}, \ldots, \frac{1}{p_{NN}}]^T$. It's clear that the Hessian matrix is positive definite, hence $p_{ij} = \pi_i, \ i, j = 1, 2, \ldots, N$ is the global optimal minimum point of f subject to the constraints. □

As $T_{std} = [p_{ij}], \ p_{ij} = \pi_i, \ i, j = 1, 2, \ldots, N$ is the transition probability matrix with maximum entropy, we can differentiate the C_{N-1}^{k-1} transition probability matrices via the Euclidean Distance of the matrix with T_{std}. Fixing the number k, for matrix $TestM$ in C_{N-1}^{k-1} transition probability matrices, the smallest distance between $TestM$ and T_{std} indicates that the corresponding matrix contains the most abundant information. We therefore choose that matrix as our preferable transition probability matrix.

3.2. Construction of PBNs from the Selected Transition Probability Matrices. In the process of constructing PBNs from the transition probability matrix, a favorable result has been obtained through the technique of adding α-Norm to the objective function. We thus can utilize the algorithm to get the desired PBNs.

For $T_i, i = 1, 2, \ldots, N$, as there are only 2 nonzero entries in each column in these matrices, it is very fast to get the desired PBNs for all the $(N-1)$ matrices. Without loss of generality, we hypothesize for some Boolean matrices B_l^j (with coefficients $coef_l^j$), we have

$$T_j = \sum_{l=1}^{M} coef_l^j \cdot B_l^j, \quad j = 1, 2, \ldots, N.$$

For a fixed number k, if we assume the optimal transition probability matrix is in the following expression:

$$OP_k = \frac{1}{k-1} \sum_{j=1}^{k-1} T_{i_j}, \quad \{i_j, j = 1, 2, \ldots, k-1\} \subset \{1, 2, \ldots, N\}.$$

Then we can directly get the desired PBNs for OP_k without further computation:

$$OP_k = \frac{1}{k-1} \sum_{j=1}^{k-1} \sum_{l=1}^{M} coef_l^{i_j} \cdot B_l^{i_j}.$$

This would save much time for the construction of PBNs, especially when k is large, as the major computational complexity was efficiently reduced from $O(k^3 2^{3n})$ to $O(k 2^{3n})$ [1], thereby offering a new perspective in PBN construction.

4. Numerical Experiments. In this section, we conduct a numerical experiment to illustrate the effectiveness of our proposed method. We suppose the number of genes is $n = 2$, then the number of states $N = 2^n = 4$. We further assume that the stationary distribution is $\boldsymbol{\pi} = (0.1\ 0.2\ 0.3\ 0.4)$. Once we know the information, it's straightforward to construct the set of sparsest probability transition matrices T_1, T_2, T_3 and matrix with maximum entropy T_{std}.

$$T_1 = \begin{pmatrix} 0.52 & 0.00 & 0.00 & 0.12 \\ 0.48 & 0.76 & 0.00 & 0.00 \\ 0.00 & 0.24 & 0.84 & 0.00 \\ 0.00 & 0.00 & 0.16 & 0.88 \end{pmatrix}, \quad T_2 = \begin{pmatrix} 0.52 & 0.00 & 0.16 & 0.00 \\ 0.00 & 0.76 & 0.00 & 0.12 \\ 0.48 & 0.00 & 0.84 & 0.00 \\ 0.00 & 0.24 & 0.00 & 0.88 \end{pmatrix}$$

$$T_3 = \begin{pmatrix} 0.52 & 0.24 & 0.00 & 0.00 \\ 0.00 & 0.76 & 0.16 & 0.00 \\ 0.00 & 0.00 & 0.84 & 0.12 \\ 0.48 & 0.00 & 0.00 & 0.88 \end{pmatrix}, \quad T_{std} = \begin{pmatrix} 0.1 & 0.1 & 0.1 & 0.1 \\ 0.2 & 0.2 & 0.2 & 0.2 \\ 0.3 & 0.3 & 0.3 & 0.3 \\ 0.4 & 0.4 & 0.4 & 0.4 \end{pmatrix}.$$

Applying the α-Norm algorithm proposed in [2] to construct the PBNs, we get the major components of BNs and corresponding coefficients:

$$T_1 \approx 0.12\times \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} + 0.16\times \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix} + 0.24\times \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} + 0.48\times \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Table 4.1: 2-norm of the C_{N-1}^{k-1} matrices with T_{std} for different k

$k=2$	T_1	T_2	T_3
2-norm with T_{std}	0.9763	1.0554	0.9208
$k=3$	$\frac{1}{2}(T_1+T_2)$	$\frac{1}{2}(T_1+T_3)$	$\frac{1}{2}(T_2+T_3)$
2-norm with T_{std}	0.8834	0.8704	0.8482
$k=4$	$\frac{1}{3}(T_1+T_2+T_3)$		
2-norm with T_{std}	0.8374		

Table 4.2: PBNs of optimal transition probability matrices for different k

	PBNs
$k=2$	$0.12\times T_{13}+0.16\times T_{23}+0.24\times T_{33}+0.48\times T_{43}$
$k=3$	$\frac{1}{2}[0.12\times T_{12}+0.16\times T_{22}+0.24\times T_{32}+0.48\times T_{42}+0.12\times T_{13}+0.16\times T_{23}+0.24\times T_{33}+0.48\times T_{43}]$
$k=4$	$\frac{1}{3}[0.12\times T_{11}+0.16\times T_{21}+0.24\times T_{31}+0.48\times T_{41}+0.12\times T_{12}+0.16\times T_{22}+0.24\times T_{32}+0.48\times T_{42}+0.12\times T_{13}+0.16\times T_{23}+0.24\times T_{33}+0.48\times T_{43}]$

$$T_2 \approx 0.12\times\begin{pmatrix}1&0&0&0\\0&1&0&1\\0&0&1&0\\0&0&0&0\end{pmatrix}+0.16\times\begin{pmatrix}1&0&1&0\\0&1&0&0\\0&0&0&0\\0&0&0&1\end{pmatrix}+0.24\times\begin{pmatrix}1&0&0&0\\0&0&0&0\\0&0&1&0\\0&1&0&1\end{pmatrix}+0.48\times\begin{pmatrix}0&0&0&0\\0&1&0&0\\1&0&1&0\\0&0&0&1\end{pmatrix}.$$

$$T_3 \approx 0.12\times\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&1&1\\0&0&0&0\end{pmatrix}+0.16\times\begin{pmatrix}1&0&0&0\\0&1&1&0\\0&0&0&0\\0&0&0&1\end{pmatrix}+0.24\times\begin{pmatrix}1&1&0&0\\0&0&0&0\\0&0&1&0\\0&0&0&1\end{pmatrix}+0.48\times\begin{pmatrix}0&0&0&0\\0&1&0&0\\0&0&1&0\\1&0&0&1\end{pmatrix}.$$

Now we are ready to construct PBNs for certain given conditions. Let $k=2$, with the support of theorem 3.3, we can find the desired optimal transition probability matrix: T_3 having smallest distance with T_{std}. Let $k=3$, we can find the desired optimal transition probability matrix: $\frac{1}{2}(T_2+T_3)$. When $k=4$, in a similar manner, the desired optimal transition probability matrix is $\frac{1}{3}(T_1+T_2+T_3)$. For better illustration, see Table 4.1.

Since we have already got the decomposition for transition probability matrices T_1, T_2, T_3, the corresponding PBNs for the optimal transition probability matrices with different k can then be efficiently constructed.

If we rewrite the PBNs for all the probability transition matrices T_1, T_2, T_3 in the following way:

$$\begin{cases} T_1 \approx 0.12\times T_{11}+0.16\times T_{21}+0.24\times T_{31}+0.48\times T_{41};\\ T_2 \approx 0.12\times T_{12}+0.16\times T_{22}+0.24\times T_{32}+0.48\times T_{42};\\ T_3 \approx 0.12\times T_{13}+0.16\times T_{23}+0.24\times T_{33}+0.48\times T_{43}. \end{cases}$$

Then, for different values of k, the optimal transition probability matrices are given in Table 4.2.

5. Conclusions. In this paper, we have proposed a novel perspective to tackle the problem of PBNs construction from a prescribed positive stationary distribution. Compelling support in theory and efficiency in realization constitute a powerful demonstration for our developed model. The promising result obtained may open up a new era towards unraveling mysterious inherent mechanisms in genetic regulatory networks.

Acknowledgments. Research support in part by HKRGC Grant 7017/07P, HKU strategic theme grant on computational sciences.

REFERENCES

[1] X. Chen, W. Ching, X.S. Chen, Y. Cong and N. Tsing, *Construction of Probabilistic Boolean Networks from a Prescribed Transition Probability Matrix: A Maximum Entropy Rate Approach*, to appear in East Asian Journal of Applied Mathematics, (2011).

[2] X. Chen, L. Li, W. Ching, N. Tsing, *A Modified Entropy Approach for Construction of Probabilistic Boolean Networks*, ISB2010, Suzhou, Computational Systems Biology Proceedings, (2010), pp. 243–250.

[3] W. Ching, X. Chen and N. Tsing, *Generating Probabilistic Boolean Networks from a Prescribed Transition Probability Matrix*, IET on Systems Biology, 6 (2009), pp. 453–464.

[4] W. Ching, S. Zhang, M. Ng and T. Akutsu, *An Approximation Method for Solving the Steady-state Probability Distribution of Probabilistic Boolean Networks*, Bioinformatics, 23 (2007), pp. 1511–1518.

[5] W. Ching and Y.Cong, *A New Optimization Model for the Construction of Markov Chains*, CSO2009, Hainan, IEEE Computer Society Proceedings, (2009), pp. 551–555.

[6] H. Jiang, W. Ching, K. Aoki-Kinoshita and D. Guo, *Delay Discrete Dynamical Models for Genetic Regulatory Networks*, ISB2010, SuZhou, Computational Systems Biology Proceedings, (2010), pp. 93–100.

[7] N. Friedman, M. Linial, I. Nachman and D. Pe'er, *Using Bayesian Networks to Analyze Expression Data*, J. Comput. Biol., 7 (2000), pp. 601–620.

[8] M. Gibson and E. Mjolsness, *Modeling the Activity of Single Gene*, J. Bower and H. Bolouri Ed., Computational Modeling of Genetic and Biochemical Networks. eds. Cambridge MA: MIT Press, (2001), chapter 1.

[9] S. Kauffman, *Metabolic Stability and Epigenesis in Randomly Constructed Gene Nets*, J. Theoret. Biol., 22 (1969), pp. 437–467.

[10] M. Song, Z. OuYang and Z. Liu, *Discrete Dynamical System Modeling for Gene Regulatory Networks of 5-hxymethylfurfural Tolerance for Ethanologenic Yeast*, IET Syst. Biol., 3 (2009), pp. 203–218.

[11] R. Pal, I. Ivanov, A. Datta, M. Bittner and E.Dougherty, *Generating Boolean Networks with a Prescribed Attractor Structure*, Bioinformatics, 21 (2005), pp. 4021–4025.

[12] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, McGraw-Hill International Editions, New York, (1989).

[13] I. Shmulevich, E. Dougherty, S. Kim and W. Zhang, *Probabilistic Boolean Networks: A Rule-based Uncertainty Model for Gene Regulatory Networks*, Bioinformatics, 18 (2002), pp. 261–274.

[14] I.Shmulevich and E.R. Dougherty, *Probabilistic Boolean Networks The Modeling and Control of Gene Regulatory Networks*, SIAM, Society for Industrial and Applied Mathematics, Philadelphia, PA, (2010), pp. 1–55.

[15] P. Smolen, D. Baxter and J. Byrne, *Mathematical Modeling of Gene Networks*, Neuron, 26 (2000), pp. 567–580.

[16] A. Wilson, *Entropy in Urban and Regional Modelling*, Pion, London, (1970).

[17] S. Zhang, W. Ching, M. Ng and T. Akutsu, *Simulation Study in Probabilistic Boolean Network Models for Genetic Regulatory Networks*, Journal of Data Mining and Bioinformatics, 1 (2007), pp. 217–240.