Research Article

Inference of Boolean Networks Using Sensitivity Regularization

Wenbin Liu,[1, 2] Harri Lähdesmäki,[1, 3] Edward R. Dougherty,[4, 5] and Ilya Shmulevich[1]

[1] *Institute for Systems Biology, Seattle, WA 98103, USA*

[2] *College of Computer Science and Engineering, Wenzhou University, Wenzhou 85004, China*

[3] *Institute of Signal Processing, Tampere University of Technology, Tampere 325035, Finland*

[4] *Department of Electrical and Computer Engineering, Texas A&M University, College Station, TX 33101, USA*

[5] *Computational Biology Division, Translational Genomics Research Institute, Phoenix, AZ 77843, USA*

Correspondence should be addressed to Ilya Shmulevich, is@ieee.org

The inference of genetic regulatory networks from global measurements of gene expressions is an important problem in computational biology. Recent studies suggest that such dynamical molecular systems are poised at a critical phase transition between an ordered and a disordered phase, affording the ability to balance stability and adaptability while coordinating complex macroscopic behavior. We investigate whether incorporating this dynamical system-wide property as an assumption in the inference process is beneficial in terms of reducing the inference error of the designed network. Using Boolean networks, for which there are well-defined notions of ordered, critical, and chaotic dynamical regimes as well as well-studied inference procedures, we analyze the expected inference error relative to deviations in the networks' dynamical regimes from the assumption of criticality. We demonstrate that taking criticality into account via a penalty term in the inference procedure improves the accuracy of prediction both in terms of state transitions and network wiring, particularly for small sample sizes.

1. Introduction

The execution of various developmental and physiological processes in cells is carried out by complex biomolecular systems. Such systems are dynamic in that they are able to change states in response to environmental cues and exhibit multiple steady states, which define different cellular functional states or cell types.

The massively parallel dynamics of complex molecular networks furnish the cell with the ability to process information from its environment and mount appropriate responses. To be able to stably execute cellular functions in a variable environment while being responsive to specific changes in the environment, such as the activation of immune cells upon exposure to pathogens or their components, the cell needs to strike a balance between robustness and adaptability.

Theoretical considerations and computational studies suggest that many types of complex dynamical systems can indeed strike such an optimal balance, under a variety of criteria, when they are operating close to a critical phase transition between an ordered and a disordered dynamical regime [1–3]. There is also accumulating evidence that living systems, as manifestations of their underlying networks of molecular interactions, are poised at the critical boundary between an organized and a disorganized state, indicating that cellular information processing is optimal in the critical regime, affording the cell with the ability to exhibit complex coordinated macroscopic behavior [4–8]. Studies of human brain oscillations [9], computer network traffic and the Internet [10, 11], financial markets [12], forest fires [13], neuronal networks supporting our senses [14], and biological macroevolution have also revealed critical dynamics [15].

A key goal in systems biology research is to characterize the molecular mechanisms governing specific cellular behaviors and processes. This typically entails selecting a model class for representing the system structure and state dynamics, followed by the application of computational or statistical inference procedures for revealing the model structure from measurement data [16]. Multiple types of

data can be potentially used for elucidating the structure of molecular networks, such as transcriptional regulatory networks, including genome wide transcriptional profiling with DNA microarrays or other high-throughput technologies, chromatin immunoprecipitation-on-chip (ChIP-on-chip) for identifying DNA sequences occupied by specific DNA binding proteins, computational predictions of transcription factor binding sites based on promoter sequence analysis, and other sources of evidence for molecular interactions [17, 18]. The inference of genetic networks is particularly challenging in the face of small sample sizes, particularly because the number of variables in the system (e.g., genes) typically greatly outnumbers the number of observations. Thus, estimates of the errors of a given model, which themselves are determined from the measurement data, can be highly variable and untrustworthy.

Any prior knowledge about the network structure, architecture, or dynamical rules is likely to improve the accuracy of the inference, especially in a small sample size scenario. If biological networks are indeed critical, a key question is whether this knowledge can be used to improve the inference of network structure and dynamics from measurements. We investigated this question using the class of Boolean networks as models of genetic regulatory networks.

Boolean networks and the more general class of probabilistic Boolean networks are popular approaches for modeling genetic networks, as these model classes capture multivariate nonlinear relationships between the elements of the system and are capable of exhibiting complex dynamics [5, 16, 19–21]. Boolean network models have been constructed for a number of biomolecular systems, including the yeast cell cycle [22, 23], mammalian cell cycle [24], Drosophila segment polarity network [25], regulatory networks of *E. coli* metabolism [26], and Arabidopsis flower morphogenesis [27–29].

At the same time, these model classes have been studied extensively regarding the relationships between their structure and dynamics. Particularly in the case of Boolean networks, dynamical phase transitions from the ordered to the disordered regime and the critical phase transition boundary have been characterized analytically for random ensembles of networks [30–34]. This makes these models attractive for investigating the relationships between structure and dynamics [35, 36].

In particular, the so-called average sensitivity was shown to be an order parameter for Boolean networks [31]. The average sensitivity, which can be computed directly from the Boolean functions specifying the update rules (i.e., state transitions) of the network, measures the average response of the system to a minimal transient perturbation and is equivalent to the Lyapunov exponent [33]. There have been a number of approaches for inferring Boolean and probabilistic Boolean networks from gene expression measurement data [20, 21, 37–44].

We address the relationship between the dynamical regime of a network, as measured by the average sensitivity, and the inference of the network from data. We study whether the assumption of criticality, embedded in the inference objective function as a penalty term, improves the

inference of Boolean network models. We find that for small sample sizes the assumption is beneficial, while for large sample sizes, the performance gain decreases gradually with increasing sample size. This is the kind of behavior that one hopes for when using penalty terms.

This paper is organized as follows. In Section 2, we give a brief definition of Boolean Networks and the concept of sensitivity. Then in Section 3, three measures used in this paper to evaluate the performance of the predicted networks are introduced, and a theoretical analysis of the relationship between the expected error and the sensitivity deviation is presented. Based on this analysis, an objective function is proposed to be used for the inference process in Section 4, while the simulation results are presented in Section 5.

2. Background and Definitions

2.1. Boolean Networks

A Boolean network $G(V, F)$ is defined by a set of nodes $V = \{x_1, \ldots, x_n\}$, $x_i \in \{0, 1\}$ and a set of Boolean functions $F = \{f_1, \ldots, f_n\}$, $f_i : \{0, 1\}^{k_i} \to \{0, 1\}$. Each node x_i represents the expression state of the gene x_i, where $x_i = 0$ means that the gene is OFF, and $x_i = 1$ means it is ON. Each Boolean function $f_i(x_{i_1}, \ldots, x_{i_{k_i}})$ with k_i specific input nodes is assigned to node x_i and is used to update its value. Under the synchronous updating scheme, all genes are updated simultaneously according to their corresponding update functions. The network's state at time t is represented by a vector $x(t) = (x_1(t), \ldots, x_n(t))$ and, in the absence of noise, the system transitions from state to state in a deterministic manner.

2.2. Sensitivity

The activity of gene x_j in function f_i is defined as

$$\alpha_j^{f_i} = \frac{1}{2^{k_i}} \sum_{x \in \{0,1\}^{k_i}} \frac{\partial f_i(x)}{\partial x_j}, \qquad (1)$$

where $\partial f_i(x)/\partial x_j = f_i(x^{(j,0)}) \oplus f_i(x^{(j,1)})$ is the partial derivative of f_i with respect to x_j, $\oplus$ is addition modulo 2, and $x^{(j,l)} = (x_1, \ldots, x_{j-1}, l, x_{j+1}, \ldots, x_{k_i})$, $l = 0, 1$ [31]. Note that the activity is equivalent to the expectation of the partial derivative with respect to the uniform distribution. Since the partial derivative is itself a Boolean function, its expectation is equal to the probability that a change in the jth input causes a change in the output of the function, and hence the activity is a number between zero and one. The average sensitivity of a function f_i equals the sum of the activities of its input variables:

$$s^{f_i} = \sum_{j=1}^{k_i} \alpha_j^{f_i}. \qquad (2)$$

In the context of random Boolean networks (RBNs), which are frequently used to study dynamics of regulatory network
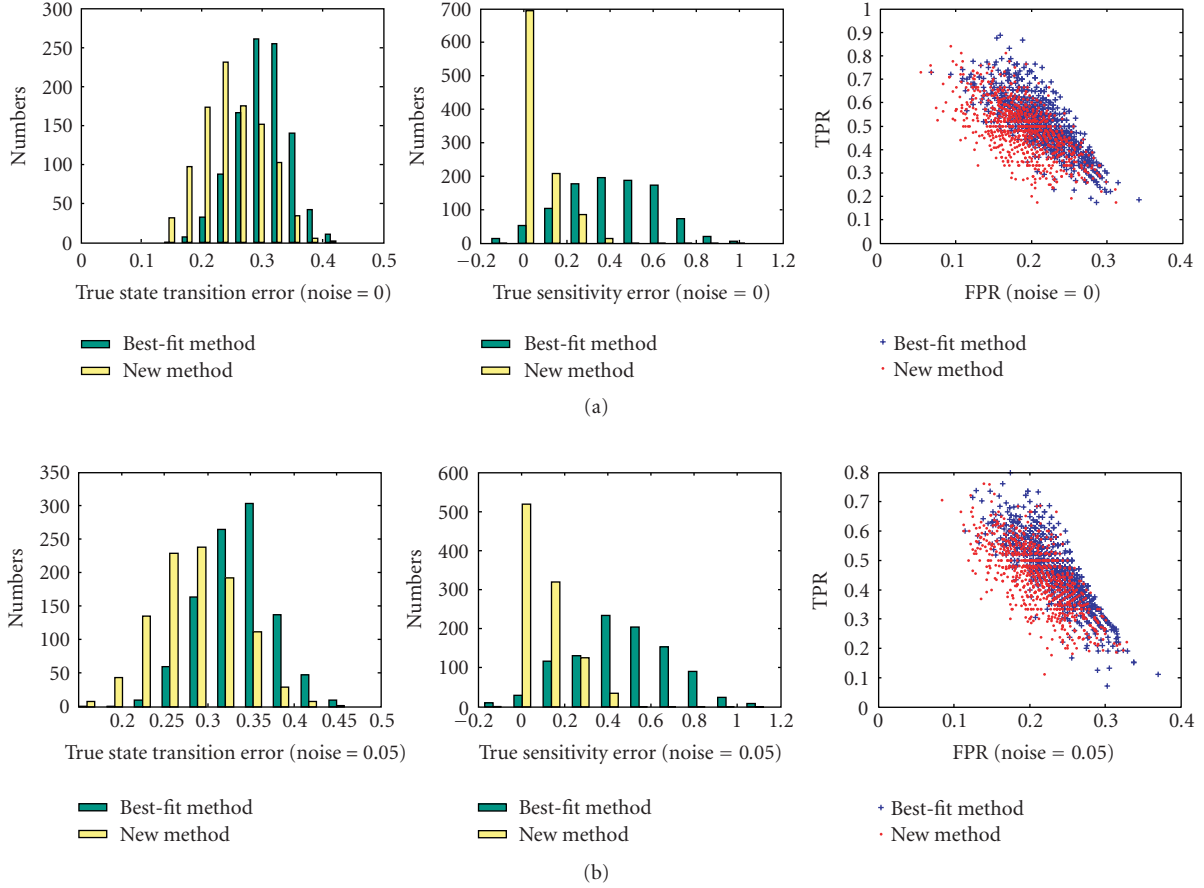
FIGURE 1: Histograms of the true error in state transition and in sensitivity and the ROC distribution for the 1000 random BNs under sample size 10.

models, another important parameter is the bias p of a function f, which is defined to be the probability that the function takes on the value 1. A random Boolean function with bias p can be generated by flipping a p-biased coin 2^k times and filling in the truth table. In other words, the truth table is a realization of 2^k independent Bernoulli (p) random variables. For a function f_i with bias p_i, the expectation of its average sensitivity is

$$E\big[s^{f_i}\big] = \sum_{j=1}^{k_i} E\big[\alpha_j^{f_i}\big] = 2k_i p_i (1 - p_i). \qquad (3)$$

The sensitivity of a Boolean network is then defined as

$$S = \frac{1}{n} \sum_{i=1}^{n} E\big[s^{f_i}\big]. \qquad (4)$$

Sensitivity is in fact a global dynamical parameter that captures how a one-bit perturbation spreads throughout the network and its expectation under the random Boolean network model is equivalent to the well-known phase transition curve [31].

3. Error Analysis

3.1. Performance Measures

There are several ways to measure the performance of an inference method by comparing the state transitions, wiring, or sensitivities with the original network. In this paper, we will use three measures that are described below.

3.1.1. The State Transition Error

This quantity generally shows the fraction of outputs that are incorrectly predicted, and it can be defined as

$$\varepsilon = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{2^n} \sum_{x \in \{0,1\}^n} \big(f_i(x) \oplus f_i'(x)\big) = \frac{1}{n} \sum_{i=1}^{n} \varepsilon_i, \qquad (5)$$

where ε_i denotes the normalized error of the predicted function f'. Additionally, f_i and f' are extended such that they are functions of all the variables (instead of k_i variables) by adding fictitious (i.e., dummy) variables.
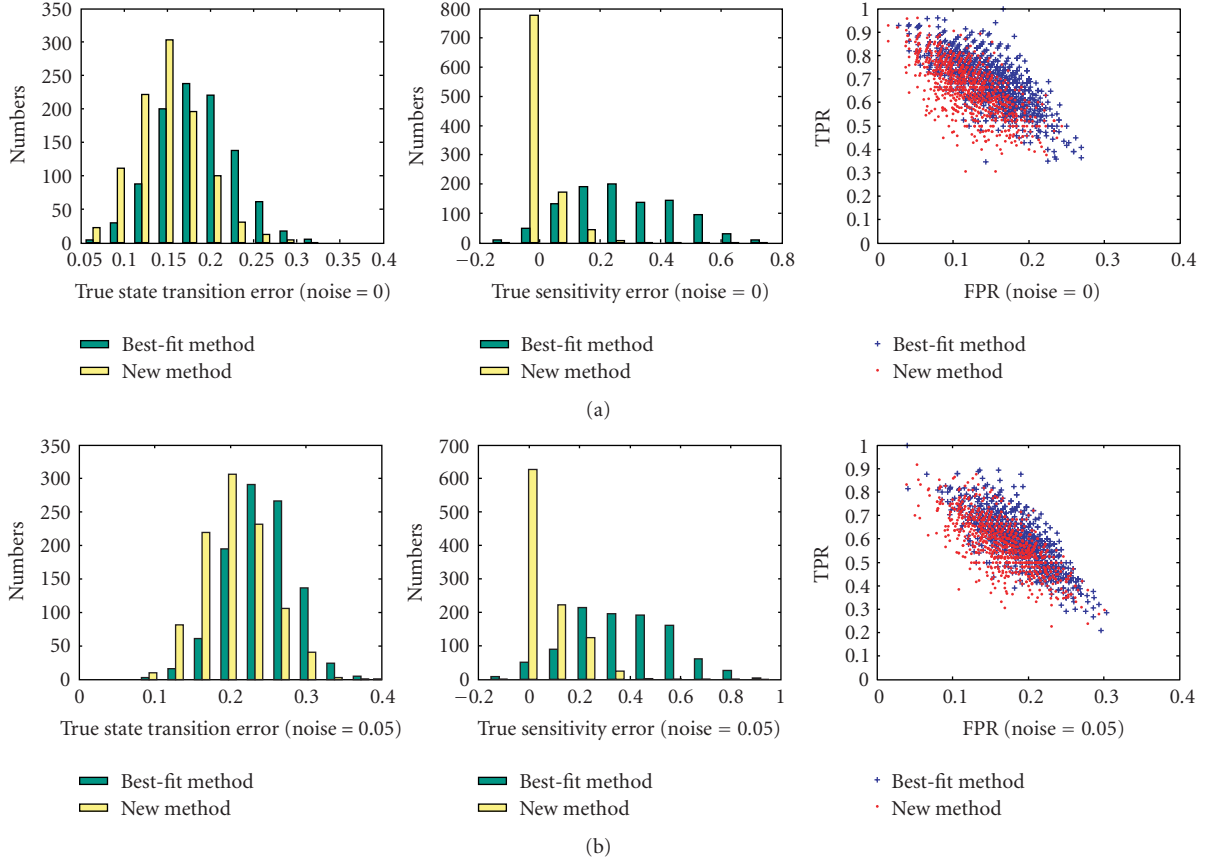
(a)

(b)

FIGURE 2: Histograms of the true error in state transition and in sensitivity and the ROC distribution for the 1000 random BNs under sample size 15.

3.1.2. The Receiver Operating Characteristic (ROC)

This measurement has been widely used in classification problems. An ROC space is defined by the false positive ratio (FPR) and the true positive ratio (TPR) plotted on the x- and y-axes, respectively, which depicts the relative tradeoffs between true positives and false positives. The FPR and TPR are defined as

$$\text{FPR} = \frac{FP}{N}, \qquad \text{TPR} = \frac{TP}{P}, \tag{6}$$

where TP and FP represent true positive and false positive instances, respectively, while P and N represent the total positive and negative instances, respectively. We will use the ROC distributions to evaluate the accuracy of "wiring" (i.e., the specific input nodes assigned to each node) for the inferred network.

3.1.3. The Sensitivity Error

The sensitivity error measures the deviation in the sensitivity of a predicted network and is defined as

$$\varepsilon_s = \frac{S' - S}{S} = \frac{\Delta S}{S}, \tag{7}$$

where S' is the sensitivity of the predicted network.

3.2. Analysis of Expected Error

All Boolean networks with a fixed number of genes n can be grouped into different families according to the network's sensitivity. Assuming that $G(V, F)$ and $G'(V, F')$ are the original network and another random network, S and S' are their sensitivities, respectively, and $P = \{p_1, \ldots, p_n\}$ and $P' = \{p'_1, \ldots, p'_n\}$ are the biases with which functions $f_i \in F$ and $f'_i \in F'$ are generated. Let $p'_i = p_i + \Delta p_i$. The expectation of the state transition error between them can be written as

$$
\begin{aligned}
E(\varepsilon) &= E\left(\frac{1}{n} \sum_{i=1}^{n} \varepsilon_i\right) \\
&= \frac{1}{n} \sum_{i=1}^{n} E(\varepsilon_i) \\
&= \frac{1}{n} \sum_{i=1}^{n} [p_i(1 - p'_i) + p'_i(1 - p_i)] \\
&= \frac{1}{n} \sum_{i=1}^{n} (p_i + p'_i - 2p_i p'_i) \\
&= \frac{1}{n} \left[\sum_{i=1}^{n} 2p_i(1 - p_i) + \sum_{i=1}^{n} (1 - 2p_i)\Delta p_i\right].
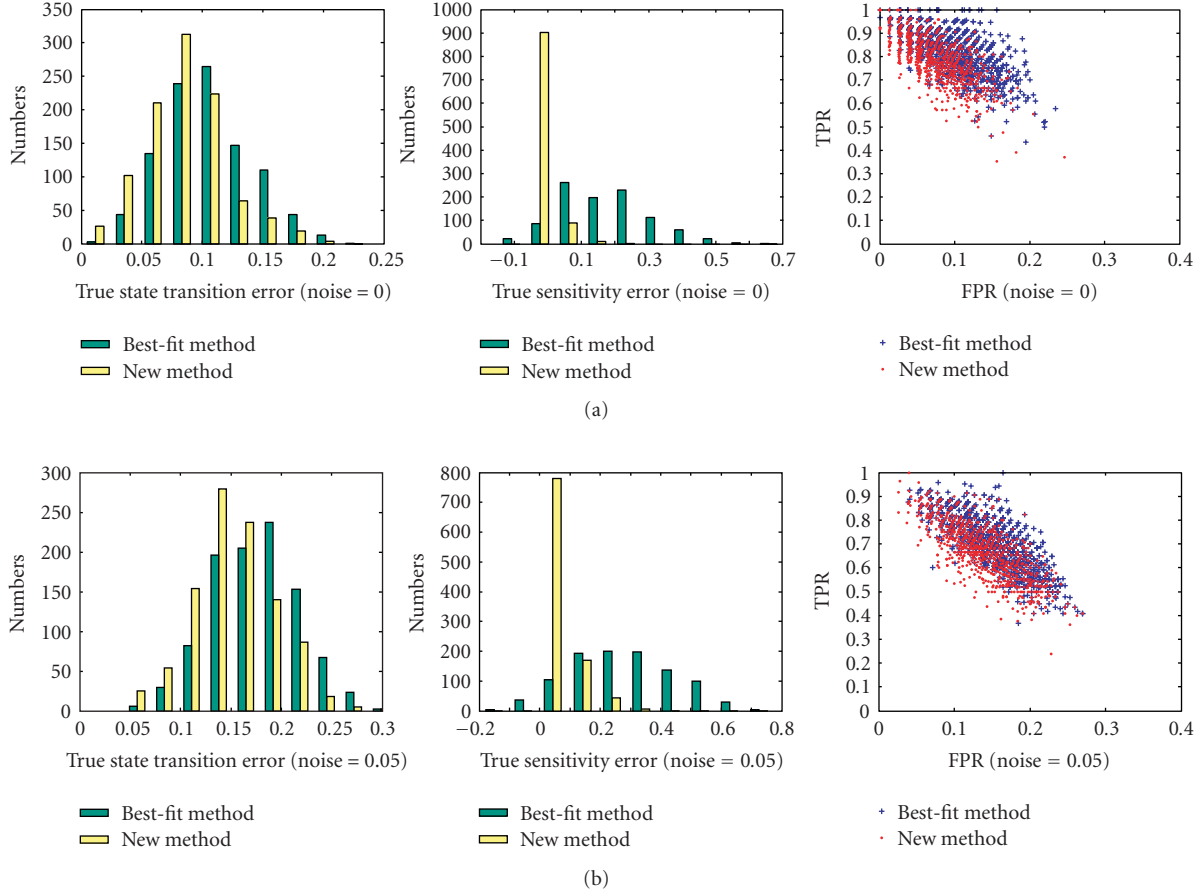\end{aligned}
\tag{8}
$$

FIGURE 3: Histograms of the true error in state transition and in sensitivity and the ROC distribution for the 1000 random BNs under sample size 20.

Using the relationship between sensitivity and bias in Section 2.2, we have

$$E(\Delta s_i)/2k_i = \frac{E(s^{f_i'} - s^{f_i})}{2k_i}$$
$$= p_i' - p_i - ((p_i')^2 - p_i^2) \qquad (9)$$
$$= (1 - 2p_i)\Delta p_i - (\Delta p_i)^2.$$

Then,

$$E(\varepsilon) = \frac{1}{n}\left[\sum_{i=1}^{n}E(s^{f_i}/k_i) + \sum_{i=1}^{n}E(\Delta s^{f_i}/(2k_i)) + \sum_{i=1}^{n}(\Delta p_i)^2\right]. \qquad (10)$$

If we further assume that both networks' connectivity is constant, $K = k_i$ ($i = 1,\ldots,n$), then

$$E(\varepsilon) = \frac{1}{nK}\sum_{i=1}^{n}E(s^{f_i}) + \frac{1}{2nK}\sum_{i=1}^{n}E(\Delta s^{f_i}) + \frac{1}{n}\sum_{i=1}^{n}\Delta p_i^2 \qquad (11)$$
$$= \frac{S}{K} + \frac{\Delta S}{2K} + \frac{1}{n}\sum_{i=1}^{n}\Delta p_i^2.$$

This means that the expectation of the state transition error $E(\varepsilon)$ generally depends on the original network's connectivity

K, its sensitivity S, the sensitivity deviation ΔS, and the mean quadratic terms of the bias deviation $(1/n)\sum_{i=1}^{n}\Delta p_i^2$.

(1) If $\Delta p_i = 0$, then ΔS will be 0. In this case, each function f_i' keeps the same bias with that of the original network, and then

$$E(\varepsilon) = \frac{S}{K}. \qquad (12)$$

(2) If $\Delta p_i \neq 0$ and ΔS is 0, the predicted network still stays in the same sensitivity class, and

$$E(\varepsilon) = \frac{S}{K} + \frac{1}{n}\sum_{i=1}^{n}\Delta p_i^2. \qquad (13)$$

(3) If $(1/n)\sum_{i=1}^{n}\Delta p_i^2$ is relatively small compared with $\Delta S/2K$, we can treat it as a constant c. Then,

$$E(\varepsilon) = \frac{S}{K} + \frac{\Delta S}{2K} + c. \qquad (14)$$

In this case, $E(\varepsilon)$ will have a linear relationship with ΔS. This indicates that $\Delta S > 0$ will surely introduce additional error $\Delta S/2K$. Our simulations indicate that the inference method we use (best-fit, see below) yields a network with $\Delta S > 0$ in most cases.
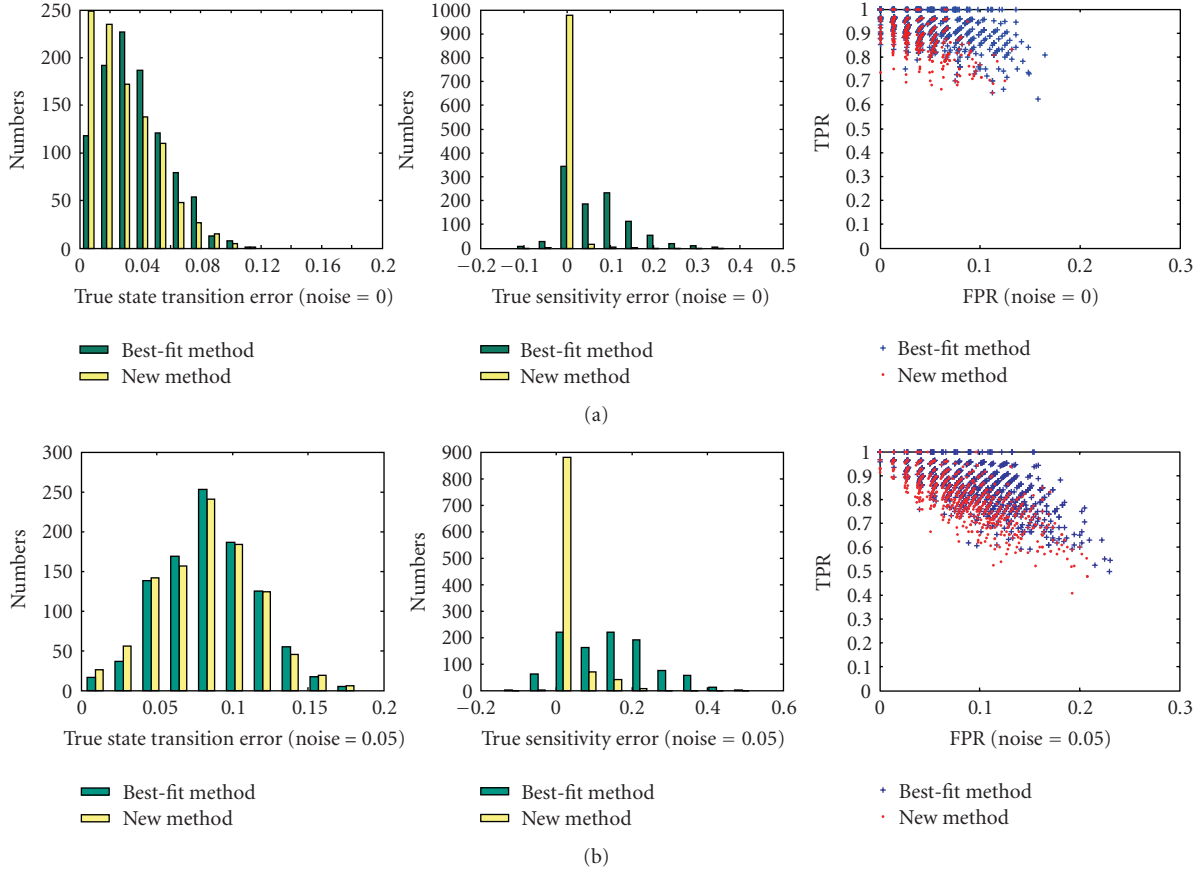
(a)

(b)

FIGURE 4: Histograms of the true error in state transition and in sensitivity and the ROC distribution for the 1000 random BNs under sample size 30.

4. Inference Method

To infer a Boolean network, for each target node we need to apply some optimization criterion to each set of input variables and Boolean function on those input variables and then choose the variable set and corresponding Boolean function that minimizes the objective function. The first step in our proposed procedure is to find variable sets and Boolean functions that provide good target prediction. Based upon time-series observations, given a target node $X_i(t+1)$ and an input-node vector $\widetilde{X}_i(t) = (X_{i_1}(t) \dots , X_{i_{k_i}}(t))$, the best predictor, f_i, minimizes the error, $\varepsilon_i(f) = \Pr[f(\widetilde{X}_i(t)) \neq X_i(t+1)]$, among all possible predictors f. Finding the best predictor for a given node means finding the minimal error among all Boolean functions over all input-variable combinations. We consider three variable combinations. Since we will optimize via an objective function containing a sensitivity penalty term, we will select a collection of input-variable sets and select the minimal-error Boolean function over each input-variable set. This is accomplished by using the plugin (resubstitution) estimate of the error $\varepsilon_i(f)$, which is given by the number of times $f(\widetilde{X}_i(t)) \neq X_i(t+1)$ in the data divided by the number of times, the pair $((\widetilde{X}_i(t), X_i(t+1))$ is observed in the data. This procedure is equivalent to the best-fit extension method [45]. We make use of an

efficient algorithm for solving the best-fit extension problem and finding all functions having error smaller than a given threshold [37]. We then select the four best variable sets and corresponding Boolean functions as candidates. The limitation of four candidates is based on computational considerations; in principle, there is no such limitation.

Because we have a small data sample, if we were to use the resubstitution error estimates employed for variable selection as error estimates for the best Boolean functions, we would expect optimistic estimates. Hence, for each selected variable set, we estimate the error of the corresponding Boolean function via the .632 bootstrap [46, 47]. A bootstrap sample consists of N equally likely draws with replacement from the original sample consisting of N data pairs $(\widetilde{X}_i(t), X_i(t+1))$. For the zero-bootstrap estimator, $\widehat{\varepsilon}_N^b$, the function is designed on the bootstrap sample and tested on the points left out, this is done repeatedly, and the bootstrap estimate is the average error made on the left-out points. $\widehat{\varepsilon}_N^b$ tends to be a high-biased estimator of the true error, since the number of points available for design is on average only 0.632 N. The .632 bootstrap estimator attempts to correct this bias via a weighted average,

$$\widehat{\varepsilon}_N^{b632} = 0.368\varepsilon_N^{\text{res}} + 0.632\widehat{\varepsilon}_N^b, \qquad (15)$$

where $\varepsilon_N^{\text{res}}$ is the original resubstitution estimate.
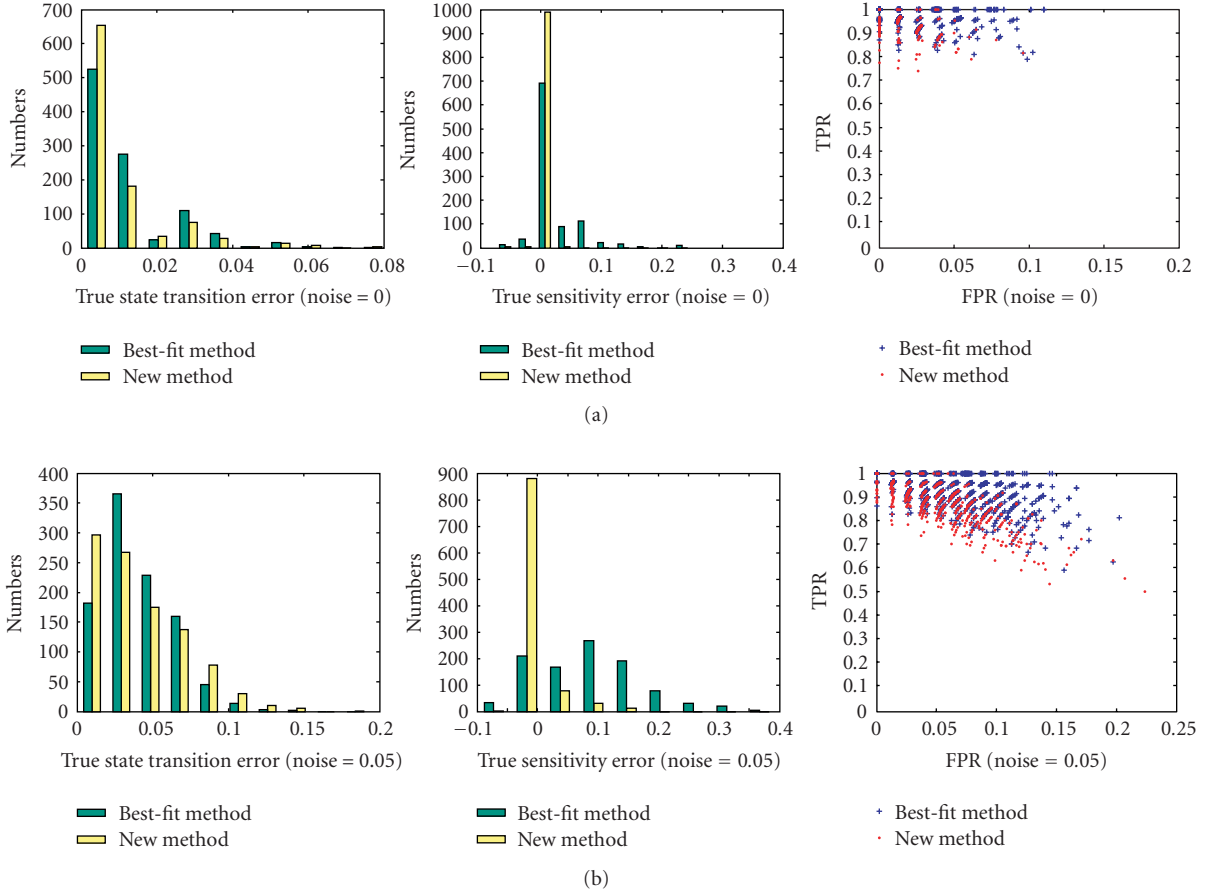
(a)

(b)

FIGURE 5: Histograms of the true error in state transition and in sensitivity and the ROC distribution for the 1000 random BNs under sample size 40.

We summarize the procedure as follows.

(1) For each three-variable set V, do the following.

 (i) Compute the resubstitution errors of all Boolean functions using the full sample data set.

 (ii) Choose the Boolean function, f_V, possessing the lowest resubstitution error as the corresponding function for V.

 (iii) Bootstrap the sample and compute the zero-bootstrap error estimate for f_V.

 (iv) Compute the .632 bootstrap error estimate for f_V using the computed resubstitution and zero-bootstrap estimates.

(2) Select the four input-variable sets whose corresponding functions possess the lowest .632 bootstrap estimates.

This procedure is the same as the one used in [48] to evaluate the impact of different error estimators on feature-set ranking. It was demonstrated there that the bootstrap tends to outperform cross-validation methods in choosing good feature sets. While it was also observed that bolstering tends to be slightly better than bootstrap, bolstering cannot be applied in discrete settings, so it is not a viable option in our context.

Motivated by the analysis in Section 3, we refine the inference process by incorporating the sensitivity. We construct an objective function

$$F_{\text{obj}} = \varepsilon' + |\varepsilon_s|, \qquad (16)$$

where ε' represents the bootstrap-estimated error of the previously selected Boolean function and ε_s is the sensitivity error. The first item represents the prediction error, while the second represents the "structural error" associated with general network dynamics. Our hypothesis is that a better inference should have a small error in both state transition and sensitivity, and consequently, the value of its objective function F_{obj} should be minimal. Of the four input-variable sets selected via prediction error for a target node, we use the one with minimal objective function F_{obj} for network construction.

5. Simulation Results

All simulations are performed for random Boolean networks with $n = 10$ and $K = 3$. For a given BN, we randomly generate m pairs of input and output states. We also consider the effect of noise, with 5% noise added to the output states of each gene by flipping its value with probability 0.05.
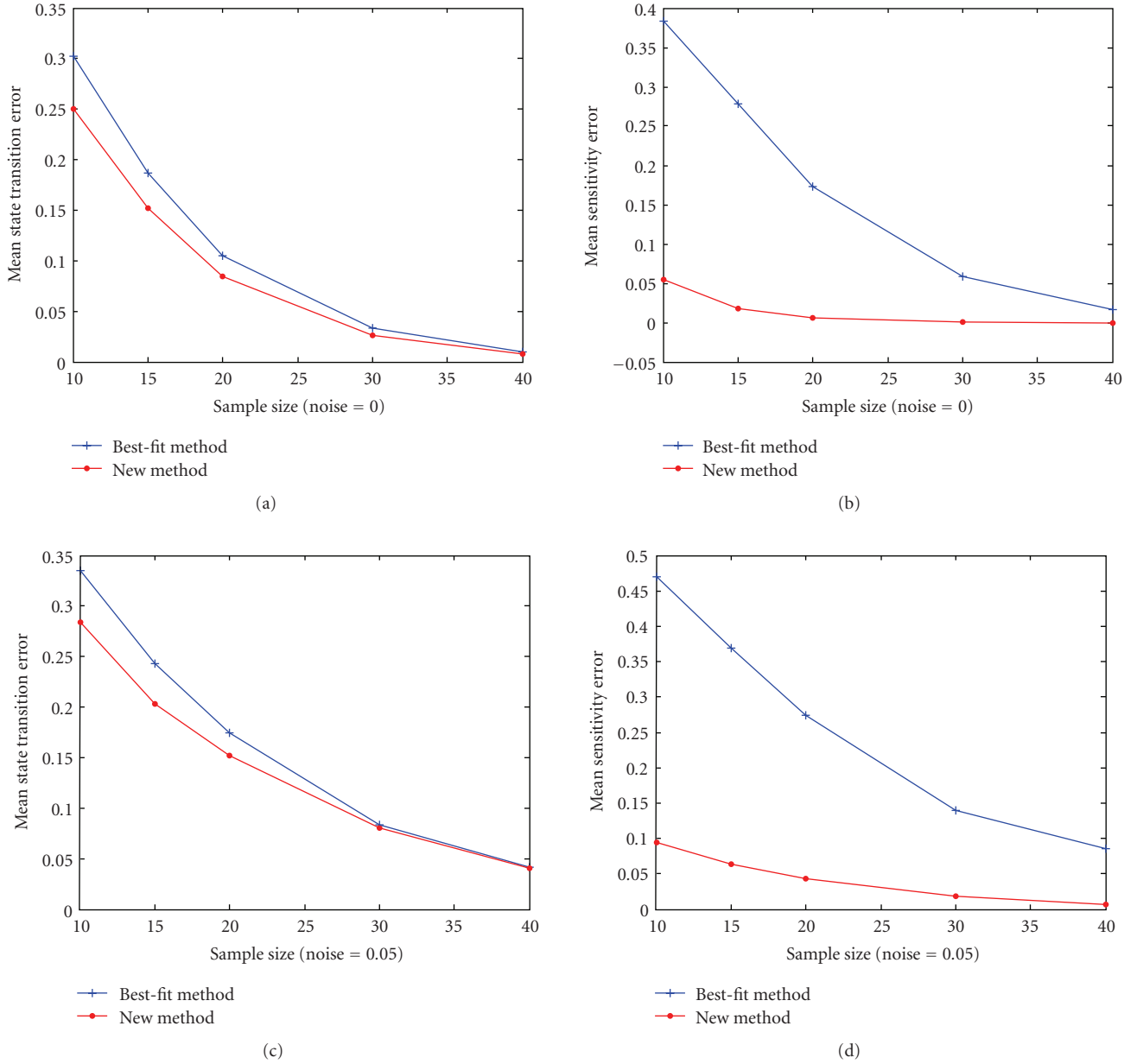
(a)

(b)

(c)

(d)

FIGURE 6: Mean state transition and sensitivity error for sample sizes ranging from 10 to 40, computed with zero noise and 5% noise.

From the perspective of network inference, performance is best characterized via a distance function between the ground-truth network and the inferred network, more specifically, by the expected distance between the ground-truth and inferred network as estimated by applying the inference to a random sample of ground-truth networks [49]. In our case, we have chosen the normalized state-transition error as the distance between the networks.

First, we investigate the performance of the new method on networks with different sensitivities, $S = 0.8, 0.9, 1.0, 1.2, 1.4$, on sample sizes ranging from 10 to 40. There are total of 200 networks for each value of the sensitivity. The left columns of Figures 1–5 are the histograms of the distribution of the true state-transition

error (Section 3.1.1) for both the traditional best-fit method (combined with .632 bootstrap) and the new proposed method. They show that the proposed method reduces this error dramatically in small sample situations. As sample size increases, the performance of both methods becomes closer. The middle columns of Figures 1–5 are the histograms of the distribution of the sensitivity error (Section 3.1.3). As can be seen, the best-fit method usually ends up with a network with larger sensitivity in small sample cases, while the proposed method can find a network operating in the same or nearby dynamic regime. The right columns of Figures 1–5 are the ROC distributions of both methods (Section 3.1.2). The proposed method has approximately the same *TPR* as the best-fit method but with a lower *FPR*. This means that the
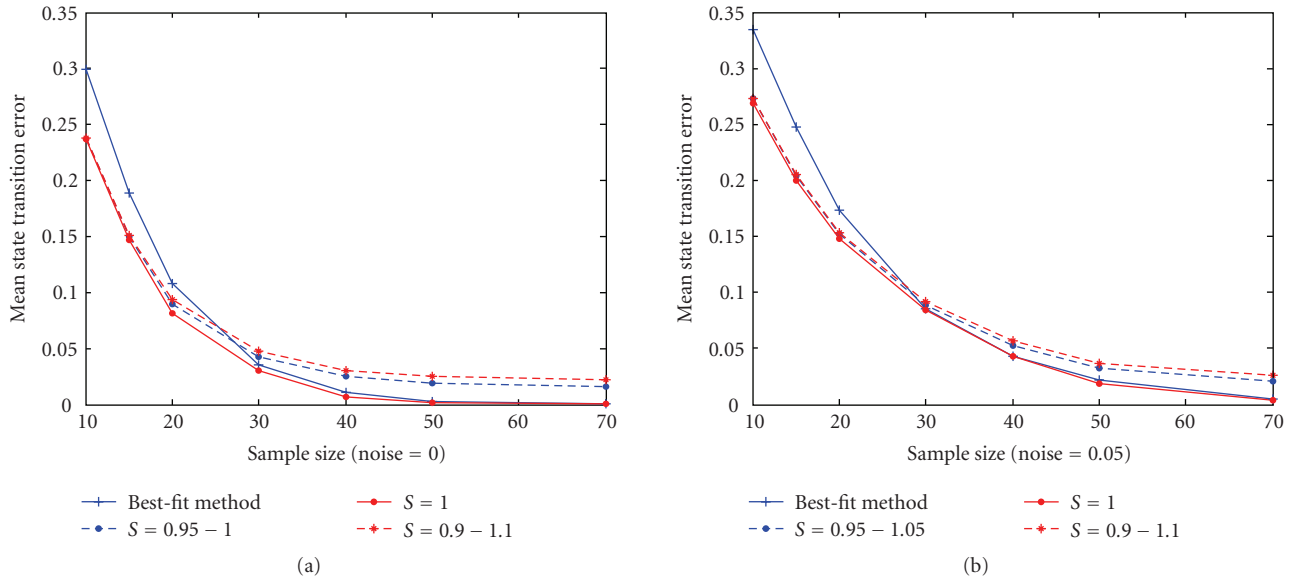
FIGURE 7: Mean state transition error of different sensitivity deviation for sample sizes ranging from 10 to 70, computed with zero noise and 5% noise.
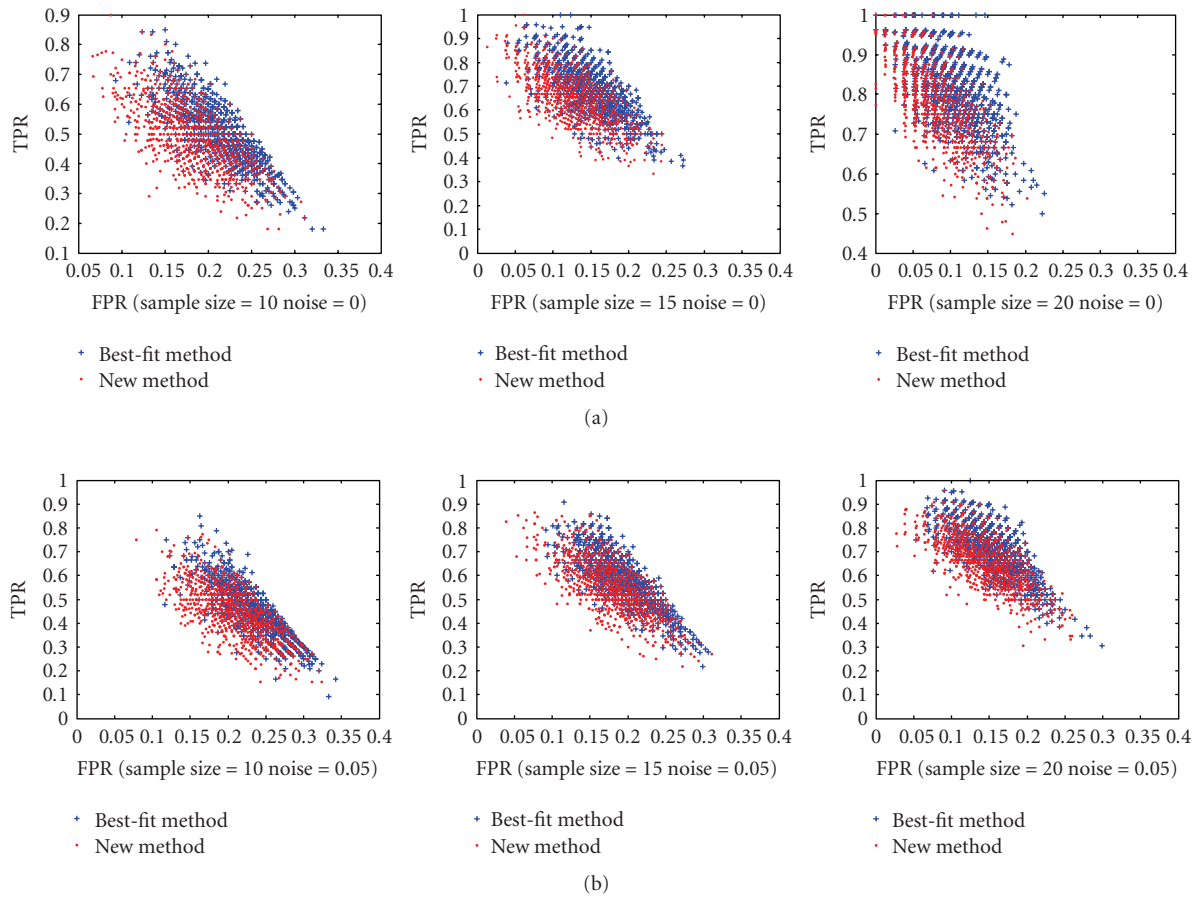

FIGURE 8: The distribution of ROC with sensitivity deviation from 0.9 to 1.1 for sample sizes 10, 15, and 20, computed with zero noise and 5% noise.

recovered connections will have higher reliability. Figure 6 shows the mean error in state transition and sensitivity under different samples sizes, for zero noise and 5% noise.

In practice, we do not know the network sensitivity, so that the assumed value in the inference procedure may not agree with the actual value. Hence, the inference procedure must be robust to this difference. Under the assumption of criticality for living systems, it is natural to set $S = 1$ in the inference procedure. Moreover, assuming that a living system remains near the border between order and disorder, the true sensitivity of gene regulatory networks will remain close to 1 under the Boolean formalism. Thus, to investigate robustness, we generated 1000 networks with sensitivities $S = 0.9, 0.95, 1.0, 1.05, 1.1$, and then inferred them using $S = 1$ with the proposed method. The mean state transition errors of both methods are shown in Figure 7.

When the actual sensitivity is 1, the method helps for small samples and the performances become close for large samples, analogous to Figure 6. When the true network deviates from the modelling assumption, $S = 1$, the proposed method helps for small samples and results in some loss of performance for large samples. This kind of behavior is what one would expect with an objective function that augments the error. In effect, the sensitivity is a penalty term in the objective function that is there to impose constraint on the optimization. In our case, when the true sensitivity is not equal to 1, the sensitivity constraint $S = 1$ yields smaller sensitivity error than the best-fit method in small sample situations, while the sensitivity error of the best-fit method is smaller for large samples. In sum, the constraint is beneficial for small samples.

Finally, the performance of the new method with regard to wiring for small sensitivity deviation is presented in Figure 8. It shows that the new method can achieve the same *TPR* with a lower *FPR* under a small sensitivity deviation in small sample situations.

6. Conclusions

Sensitivity is a global structural parameter of a network which captures the network's operating dynamic behavior: ordered, critical, or chaotic. Recent evidence suggests that living systems operate at the critical phase transition between ordered and chaotic regimes. In this paper, we have proposed a method to use this dynamic information to improve the inference of Boolean networks from observations of input-output relationships. First, we have analyzed the relationship between the expectation of the error and the deviation of sensitivity, showing that these quantities are strongly correlated with each other. Based on this observation, an objective function is proposed to refine the inference approach based on the best-fit method. The simulation results demonstrate that the proposed method can improve the predicted results both in terms of state transitions, sensitivity, and network wiring. The improvement is particularly evident in small sample size settings. As the sample size increases, the performance of both methods becomes similar. In practice, where one does not know the sensitivity of the true network, we have assumed it to be 1, the critical value, and investigated

inference performance relative to its robustness to the true sensitivity deviating from 1. For small samples, the kind we are interested in when using such a penalty approach, the proposed method continues to outperform the best-fit method.

For practical applications, one can apply an optimization strategy, such as genetic algorithms, to attain suboptimal solutions instead of the brute force searching strategy used in this paper. As the final chosen function for each gene generally lies within the top three candidates in our simulations, one can just select from a few top candidate functions for each gene instead of using all of the possible $\binom{n}{k}$ candidates. Finally, it should be noted that the ideas presented here could also be incorporated into other inference methods, such as the ones in [40, 41].

References

[1] C. G. Langton, "Computation at the edge of chaos: phase transitions and emergent computation," *Physica D*, vol. 42, no. 1–3, pp. 12–37, 1990.

[2] P. Krawitz and I. Shmulevich, "Basin entropy in Boolean network ensembles," *Physical Review Letters*, vol. 98, no. 15, Article ID 158701, 4 pages, 2007.

[3] N. H. Packard, "Adaptation towards the edge of chaos," in *Dynamic Patterns in Complex Systems*, J. A. S. Kelso, A. J. Mandell, and M. F. Shlesinger, Eds., pp. 293–301, World Scientific, Singapore, 1988.

[4] P. Rämö, J. Kesseli, and O. Yli-Harja, "Perturbation avalanches and criticality in gene regulatory networks," *Journal of Theoretical Biology*, vol. 242, no. 1, pp. 164–170, 2006.

[5] S. A. Kauffman, *The Origins of Order: Self-Organization and Selection in Evolution*, Oxford University Press, New York, NY, USA, 1993.

[6] I. Shmulevich, S. A. Kauffman, and M. Aldana, "Eukaryotic cells are dynamically ordered or critical but not chaotic," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, no. 38, pp. 13439–13444, 2005.

[7] R. Serra, M. Villani, and A. Semeria, "Genetic network models and statistical properties of gene expression data in knock-out experiments," *Journal of Theoretical Biology*, vol. 227, no. 1, pp. 149–157, 2004.

[8] M. Nykter, N. D. Price, M. Aldana, et al., "Gene expression dynamics in the macrophage exhibit criticality," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 105, no. 6, pp. 1897–1900, 2008.

[9] K. Linkenkaer-Hansen, V. V. Nikouline, J. M. Palva, and R. J. Ilmoniemi, "Long-range temporal correlations and scaling behavior in human brain oscillations," *Journal of Neuroscience*, vol. 21, no. 4, pp. 1370–1377, 2001.

[10] S. Valverde and R. V. Solé, "Self-organized critical traffic in parallel computer networks," *Physica A*, vol. 312, no. 3-4, pp. 636–648, 2002.

[11] K. Fukuda, H. Takayasu, and M. Takayasu, "Origin of critical behavior in Ethernet traffic," *Physica A*, vol. 287, no. 1-2, pp. 289–301, 2000.

[12] T. Lux and M. Marchesi, "Scaling and criticality in a stochastic multi-agent model of a financial market," *Nature*, vol. 397, no. 6719, pp. 498–500, 1999.

[13] B. D. Malamud, G. Morein, and D. L. Turcotte, "Forest fires: an example of self-organized critical behavior," *Science*, vol. 281, no. 5384, pp. 1840–1842, 1998.

[14] O. Kinouchi and M. Copelli, "Optimal dynamical range of excitable networks at criticality," *Nature Physics*, vol. 2, no. 5, pp. 348–352, 2006.

[15] K. Sneppen, P. Bak, H. Flyvbjerg, and M. H. Jensen, "Evolution as a self-organized critical phenomenon," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 92, no. 11, pp. 5209–5213, 1995.

[16] I. Shmulevich and E. R. Dougherty, *Genomic Signal Processing*, Princeton Series in Applied Mathematics, Princeton University Press, Princeton, NJ, USA, 2007.

[17] S. A. Ramsey, S. L. Klemm, D. E. Zak, et al., "Uncovering a macrophage transcriptional program by integrating evidence from motif scanning and expression dynamics," *PLoS Computational Biology*, vol. 4, no. 3, e1000021, pp. 1–25, 2008.

[18] H. Lähdesmäki, A. G. Rust, and I. Shmulevich, "Probabilistic inference of transcription factor binding from multiple data sources," *PLoS ONE*, vol. 3, no. 3, p. e1820, 2008.

[19] S. Kauffman, "Homeostasis and differentiation in random genetic control networks," *Nature*, vol. 224, no. 5215, pp. 177–178, 1969.

[20] I. Shmulevich, E. R. Dougherty, S. Kim, and W. Zhang, "Probabilistic Boolean networks: a rule-based uncertainty model for gene regulatory networks," *Bioinformatics*, vol. 18, no. 2, pp. 261–274, 2002.

[21] I. Shmulevich, E. R. Dougherty, and W. Zhang, "From Boolean to probabilistic Boolean networks as models of genetic regulatory networks," *Proceedings of the IEEE*, vol. 90, no. 11, pp. 1778–1792, 2002.

[22] F. Li, T. Long, Y. Lu, Q. Ouyang, and C. Tang, "The yeast cell-cycle network is robustly designed," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 101, no. 14, pp. 4781–4786, 2004.

[23] M. I. Davidich and S. Bornholdt, "Boolean network model predicts cell cycle sequence of fission yeast," *PLoS ONE*, vol. 3, no. 2, p. e1672, 2008.

[24] A. Fauré, A. Naldi, C. Chaouiya, and D. Thieffry, "Dynamical analysis of a generic Boolean model for the control of the mammalian cell cycle," *Bioinformatics*, vol. 22, no. 14, pp. e124–e131, 2006.

[25] R. Albert and H. G. Othmer, "The topology of the regulatory interactions predicts the expression pattern of the segment polarity genes in *Drosophila melanogaster*," *Journal of Theoretical Biology*, vol. 223, no. 1, pp. 1–18, 2003.

[26] A. Samal and S. Jain, "The regulatory network of *E. coli* metabolism as a Boolean dynamical system exhibits both homeostasis and flexibility of response," *BMC Systems Biology*, vol. 2, article 21, 2008.

[27] C. Espinosa-Soto, P. Padilla-Longoria, and E. R. Alvarez-Buylla, "A gene regulatory network model for cell-fate determination during *Arabidopsis thaliana* flower development that is robust and recovers experimental gene expression profiles," *The Plant Cell*, vol. 16, no. 11, pp. 2923–2939, 2004.

[28] L. Mendoza and E. R. Alvarez-Buylla, "Dynamics of the genetic regulatory network for *Arabidopsis thaliana* flower morphogenesis," *Journal of Theoretical Biology*, vol. 193, no. 2, pp. 307–319, 1998.

[29] L. Mendoza, D. Thieffry, and E. R. Alvarez-Buylla, "Genetic control of flower morphogenesis in *Arabidopsis thaliana*: a logical analysis," *Bioinformatics*, vol. 15, no. 7-8, pp. 593–606, 1999.

[30] B. Derrida and Y. Pomeau, "Random networks of automata: a simple annealed approximation," *Europhysics Letters*, vol. 1, no. 2, pp. 45–49, 1986.

[31] I. Shmulevich and S. A. Kauffman, "Activities and sensitivities in Boolean network models," *Physical Review Letters*, vol. 93, no. 4, Article ID 048701, 4 pages, 2004.

[32] M. Aldana, S. Coppersmith, and L. P. Kadanoff, "Boolean dynamics with random couplings," in *Perspectives and Problems in Nonlinear Science*, E. Kaplan, J. E. Marsden, and K. R. Sreenivasan, Eds., pp. 23–89, Springer, New York, NY, USA, 2002.

[33] B. Luque and R. V. Solé, "Lyapunov exponents in random Boolean networks," *Physica A*, vol. 284, no. 1–4, pp. 33–45, 2000.

[34] B. Luque and R. V. Solé, "Phase transitions in random networks: simple analytic determination of critical points," *Physical Review E*, vol. 55, no. 1, pp. 257–260, 1997.

[35] Y. D. Nochomovitz and H. Li, "Highly designable phenotypes and mutational buffers emerge from a systematic mapping between network topology and dynamic output," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, no. 11, pp. 4180–4185, 2006.

[36] M. Nykter, N. D. Price, A. Larjo, et al., "Critical networks exhibit maximal information diversity in structure-dynamics relationships," *Physical Review Letters*, vol. 100, no. 5, Article ID 058702, 4 pages, 2008.

[37] H. Lähdesmäki, I. Shmulevich, and O. Yli-Harja, "On learning gene regulatory networks under the Boolean network model," *Machine Learning*, vol. 52, no. 1-2, pp. 147–167, 2003.

[38] X. Zhou, X. Wang, and E. R. Dougherty, "Gene prediction using multinomial probit regression with Bayesian gene selection," *EURASIP Journal on Applied Signal Processing*, vol. 2004, no. 1, pp. 115–124, 2004.

[39] W. Zhao, E. Serpedin, and E. R. Dougherty, "Inferring gene regulatory networks from time series data using the minimum description length principle," *Bioinformatics*, vol. 22, no. 17, pp. 2129–2135, 2006.

[40] S. Marshall, L. Yu, Y. Xiao, and E. R. Dougherty, "Inference of a probabilistic Boolean network from a single observed temporal sequence," *EURASIP Journal on Bioinformatics and Systems Biology*, vol. 2007, Article ID 32454, 15 pages, 2007.

[41] R. Pal, I. Ivanov, A. Datta, M. L. Bittner, and E. R. Dougherty, "Generating Boolean networks with a prescribed attractor structure," *Bioinformatics*, vol. 21, no. 21, pp. 4021–4025, 2005.

[42] I. Tabus and J. Astola, "On the use of MDL principle in gene expression prediction," *EURASIP Journal on Applied Signal Processing*, vol. 2001, no. 4, pp. 297–303, 2001.

[43] T. Akutsu, S. Miyano, and S. Kuhara, "Inferring qualitative relations in genetic networks and metabolic pathways," *Bioinformatics*, vol. 16, no. 8, pp. 727–734, 2000.

[44] T. Akutsu, S. Miyano, and S. Kuhara, "Algorithms for identifying Boolean networks and related biological networks based on matrix multiplication and fingerprint function," *Journal of Computational Biology*, vol. 7, no. 3-4, pp. 331–343, 2000.

[45] E. Boros, T. Ibaraki, and K. Makino, "Error-free and best-fit extensions of partially defined Boolean functions," *Information and Computation*, vol. 140, no. 2, pp. 254–283, 1998.

[46] B. Efron, "Estimating the error rate of a prediction rule: improvement on cross-validation," *Journal of the American Statistical Association*, vol. 78, no. 382, pp. 316–331, 1983.

[47] B. Efron and R. J. Tibshirani, *An Introduction to the Bootstrap*, vol. 57 of *Monographs on Statistics and Applied Probability*, Chapman & Hall, New York, NY, USA, 1993.

[48] C. Sima, S. Attoor, U. Brag-Neto, J. Lowey, E. Suh, and E. R. Dougherty, "Impact of error estimation on feature selection," *Pattern Recognition*, vol. 38, no. 12, pp. 2472–2482, 2005.

[49] E. R. Dougherty, "Validation of inference procedures for gene regulatory networks," *Current Genomics*, vol. 8, no. 6, pp. 351–359, 2007.