Results: Two identification algorithms are developed that can be efficiently applied to causal regulation identification of a large-scale network from noisy steady-state experiment data. A distinguished feature of these algorithms is that power law has been explicitly incorporated into estimations, which is one important structural property that most large-scale cellular networks approximately have. Under the condition that parameters of the power law are known and measurement errors are Gaussian, a likelihood maximization approach is adopted. The developed estimation algorithms consist of three major steps. At first, angle minimization between subspaces is utilized to identify chemical elements that have direct influences on a prescribed chemical element, under the condition that the number of direct regulations is known. Second, interference coefficients from prescribed chemical elements are estimated through likelihood maximization with respect to measurement errors. Finally, direct regulation numbers are identified through maximizing a lower bound of an overall likelihood function. These methods have been applied from prescribed chemical elements are estimated through likelihood maximization with respect to measurement errors. Finally, direct regulation numbers are identified through maximizing a lower bound of an overall likelihood function. These methods have been applied to an artificially constructed linear system with 100 elements, a mitogen-activated protein kinase pathway model with 103 chemical elements, some DREAM initiative in silico data and some in vivo data. Compared with the widely adopted total least squares (TLS) method, computation results show that parametric estimation accuracy can be significantly increased and false positive errors can be greatly reduced.

Availability: The Matlab files for the methods are available at http://bioinfo.au.tsinghua.edu.cn/member/ylwang/Matlabfiles_CNI.zip
Contact: tzhou@mail.tsinghua.edu.cn
Supplementary Information: Supplementary data are available at Bioinformatics online.

To whom correspondence should be addressed.
†The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors.

Received on January 19, 2010; revised on May 27, 2010; accepted on June 11, 2010

1 INTRODUCTION

Inferring causal relationships of a cellular network is one of the fundamental problems in understanding cell behaviors. Major reasons behind extensive attentions to this problem appear to lie upon two facts. One is that cellular networks are generally of a large scale and there are a huge number of causal interactions to be estimated. The other is that there are many restrictions on probing signals in biochemical experiments. For example, the perturbations must be non-negative and some chemical elements may not be directly affected, and so on. These make the problem of estimating a cellular network much more difficult than identifying an industrial process (Akutsu et al., 1999; Andrec et al., 2005; Barabasi and Oltvai, 2004; Gardner and Faith, 2005).

With significant development of high-throughout technologies and proteomics analysis methods, several approaches have now been proposed for unraveling the interactions of a cellular network, such as Boolean network methods (Akutsu et al., 1999; Shmulevich et al., 2002; Zheng and Kwoh, 2004), Bayesian network methods (Ferrazzi et al., 2007; Perrin et al., 2003), partial correlation analysis (de la Fuente et al., 2004), differential equation-based time series analysis (Bansal et al., 2006; Sontag, 2008) and so on. However, when applying these methods to large-scale cellular networks, several difficulties may arise (Cantone et al., 2009; Prill et al., 2010). For example, with the number increment of chemical elements in a network, computational burden of most of the methods increases exponentially. To overcome this difficulty, some methods limit the maximum direct regulation number, but this may restrict application ranges of the method itself and give rise to a problem of selecting principal chemical elements. On the other hand, statistical methods such as partial correlation analysis rely on a variety of pairwise correlation measures. This means that an indirect effect may be wrongly recognized as a direct one, and therefore leads to a high false positive rate.

In addition, a ‘top-down’ approach is proposed in Kholodenko et al. (2002) and Andrec et al. (2005) for causal regulation inference from steady-state concentration changes of chemical species. This strategy is based on the total differential formula and total least square (TLS) estimations. Afterwards, the results have been extended to time series data, quasi-steady-state data, etc. (Sontag, 2008), and the ‘non-direct effect’ condition on experiment designs has been significantly weakened (Berman et al., 2007). Our computation experiences, however, show that when data length is of a moderate size, it is very rare that an identified regulation
coefficient is near zero. Recalling that a large-scale system usually has a sparse structure, this means that TLS estimates usually give incorrect information in case that a chemical element has no direct effect on another chemical element. In other words, there may exist significant false positive errors in TLS-based estimates. On the other hand, it is widely recognized that the cellular network identification, distinguishing direct and indirect regulations is of great biological significance (Barabasi and Oltvai, 2004; Cantone et al., 2009; Prill et al., 2010). In actual simulations, however, experimental data alone are rarely sufficient to obtain a statistically sound model. To make things worse, it may even be possible that the identification problem is under-determined if direct regulations exist between every two chemical elements. These imply that in this identification problem, compared with false negative errors, it may be much more difficult to reduce false positive errors. To overcome this difficulty, various interesting methods have been suggested, such as to incorporate qualitative knowledge, to restrict the maximum number of non-zero regulatory inputs, to penalize the sum of direct regulation strengths, etc. (Cantone et al., 2009; Chang et al., 2008; Gardner et al., 2003; Prill et al., 2010).

In this article, we investigate possibilities of utilizing structural information of a large-scale system to reduce both computational complexities and false positive errors. As pointed out in Barabasi and Oltvai (2004), most large-scale cellular networks obey approximately a so-called power law. This observation gives us some important a priori information about the structure of cellular networks. However, to the best of our knowledge, there are few researches so far that have explicitly explored possibilities of utilizing this structural property in cellular network identification. In this article, we incorporate this property into some estimation algorithms, in order to facilitate the inference of direct causal regulations from noisy steady-state experiment data. More specifically, a power law is incorporated into a likelihood function to give an estimate about the direct regulation number. Here, 

\[
\begin{align*}
\sum_{i=1}^{n} x_i &= \sum_{i=1}^{n} y_i = 100
\end{align*}
\]

In this article, causal relations are inferred from steady-state concentration changes of chemical species in a cellular network. This strategy is based on the total differential of a nonlinear function which has also been used in Khodolenko et al. (2002) and Andrec et al. (2005). A brief derivation is given in Appendix A of the Supplementary Material for relations among measurements of species concentrations and direct regulations, which extends the results of Khodolenko et al. (2002) and Andrec et al. (2005) from single perturbation to multiple perturbations. Assume that there are \( n \) chemical elements in a cellular network and in the \( j \)-th experiment, some external perturbations are added on a chemical element which do not directly change the concentration of this species, then, the following relation can be established approximately for a cellular network when the dynamics of species concentrations can be described by a set of ordinary nonlinear differential equations,

\[
\begin{align*}
\text{Ax} &= \text{b}
\end{align*}
\]

The problem discussed in this article is to identify vector \( x \) under the condition that both matrix \( A \) and vector \( b \) are provided. A distinctive characteristic of this problem is that measurement errors exist in both matrix \( A \) and vector \( b \). On the other hand, when \( n \) is large, it is now well known that the distribution of the number of non-zero elements of vector \( x \) obeys approximately the so-called power law. More precisely, let \( n_i \) represent the number of chemical elements that have direct influences on a randomly chosen chemical element in a cellular network, and \( \Pr(\cdot) \) the probability of the occurrence of a random event. Then, there exist a positive number \( \gamma \) and a positive integer \( k_{max} \), such that

\[
\begin{align*}
\Pr(n_i = k) &= \begin{cases} 
\frac{k_{\text{max}}^{-\gamma}}{k_{\text{max}}^\gamma} & 1 \leq k \leq k_{\text{max}} \\
0 & k_{\text{max}} < k \leq n \end{cases}
\end{align*}
\]

It is worthwhile to point out that in structural analysis for a large-scale network, investigations are usually focused on large \( n_i \)'s (Clauset et al., 2009). As there is generally no statistical information about the distribution of small \( n_i \)'s, it is reasonable to assume that all of them have an equal probability to occur. This treatment makes the methods suggested in this article also applicable to identification of small- or moderate-sized networks, as well as consistent with the method proposed in Gardner et al. (2003), which restricts the number of non-zero direct regulations.

2 METHODS

2.1 A general description of the identification procedure

In this article, causal relations are inferred from steady-state concentration changes of chemical species in a cellular network. This strategy is based on the total differential of a nonlinear function which has also been used in Khodolenko et al. (2002) and Andrec et al. (2005). A brief derivation is given in Appendix A of the Supplementary Material for relations among measurements of species concentrations and direct regulations, which extends the results of Khodolenko et al. (2002) and Andrec et al. (2005) from single perturbation to multiple perturbations. Assume that there are \( n \) chemical elements in a cellular network and in the \( j \)-th experiment, some external perturbations are added on a chemical element which do not directly change the concentration of this species, then, the following relation can be established approximately for a cellular network when the dynamics of species concentrations can be described by a set of ordinary nonlinear differential equations,

\[
\begin{align*}
\text{Ax} &= \text{b}
\end{align*}
\]

The problem discussed in this article is to identify vector \( x \) under the condition that both matrix \( A \) and vector \( b \) are provided. A distinctive characteristic of this problem is that measurement errors exist in both matrix \( A \) and vector \( b \). On the other hand, when \( n \) is large, it is now well known that the distribution of the number of non-zero elements of vector \( x \) obeys approximately the so-called power law. More precisely, let \( n_i \) represent the number of chemical elements that have direct influences on a randomly chosen chemical element in a cellular network, and \( \Pr(\cdot) \) the probability of the occurrence of a random event. Then, there exist a positive number \( \gamma \) and a positive integer \( k_{max} \), such that

\[
\begin{align*}
\Pr(n_i = k) &= \begin{cases} 
\frac{k_{\text{max}}^{-\gamma}}{k_{\text{max}}^\gamma} & 1 \leq k \leq k_{\text{max}} \\
0 & k_{\text{max}} < k \leq n \end{cases}
\end{align*}
\]

It is worthwhile to point out that in structural analysis for a large-scale network, investigations are usually focused on large \( n_i \)'s (Clauset et al., 2009). As there is generally no statistical information about the distribution of small \( n_i \)'s, it is reasonable to assume that all of them have an equal probability to occur. This treatment makes the methods suggested in this article also applicable to identification of small- or moderate-sized networks, as well as consistent with the method proposed in Gardner et al. (2003), which restricts the number of non-zero direct regulations.
in which \( r = \left[ l_{n}^{-1} + \sum_{i=0}^{k} s_{i} x_{i}^{-2} \right]^{-1} \). The purposes of this article are to incorporate the above structural information into cellular network identification and to investigate its capabilities in estimation accuracy improvements.

2.2 Identification algorithm

As a first step toward incorporating the so-called power law into cellular network identification, the following two assumptions are adopted:

- Represent measurement errors of matrix \( A \) and vector \( b \), respectively, by \( e_{k} \) and \( e_{k} \). It is assumed that elements of \( \{e_{k}\} \) are independent of each other and have an identical normal distribution \( N(0, \sigma^{2}) \) for a known \( \sigma \).
- Parameters \( k_{m} \), \( y_{m} \), and \( r_{m} \) are known for the power law.

When these pieces of information are available, a natural approach for causal regulation identification from experimental data is likelihood maximization. More specifically, let \( \mathcal{F}_{k} \) represent the number of non-zero elements of a matrix or vector. Then, the likelihood function of measurement errors and the direct regulation number, denote it by \( L(\mathcal{F}_{k}, e_{k}, b) \), can be written as follows.

\[
L(\mathcal{F}_{k}, e_{k}, b) = (2\pi \sigma^{2})^{-m/2} \exp(-\frac{1}{2\sigma^{2}} \sum_{i} e_{i}^{2})
\]

subject to: \( (A - e_{k}) b = e_{k} \) and \( \mathcal{F}_{k} = k \) \( (5) \)

in which \( \text{tr} \) denotes the trace of a square matrix.

Note that \( m, n, k_{m}, y_{m} \), and \( r_{m} \) are assumed to be known. Recalling that \( \ln(\cdot) \) is a monotonically increasing function over \((0, \infty)\), it is obvious that the above maximization problem is equivalent to minimizing a cost function \( h(\mathcal{F}_{k}, e_{k}, b) \) under the conditions of Equation (5). Here, \( h(\mathcal{F}_{k}, e_{k}, b) \)

\[
h(\mathcal{F}_{k}, e_{k}, b) = \text{tr}[(x_{k}^T x_{k})^{2} (e_{k}^T e_{k})] + \sum_{i} \ln(k_{m} + e_{i}^{2})
\]

\( (6) \)

While the above cost function is physically significant, both discrete and continuous variables are included in its optimization. Currently, it appears difficult to derive an analytic expression for the optimal \( k_{m}, e_{k} \) and \( e_{k} \) as well as to develop a globally/locally convergent optimization algorithm. In order to obtain an estimate about a cellular network, the following three major steps are adopted in this article.

- For a fixed \( k \), angle minimization between two subspaces is utilized to determine positions of the non-zero elements of vector \( x \).
- Under the condition that positions of non-zero elements are prespecified, \( x \) is estimated through minimizing the first term of the cost function \( h(\mathcal{F}_{k}, e_{k}, b) \) with respect to measurement errors \( e_{k} \).
- The number of non-zero elements of vector \( x \) is obtained through a numerical search which minimizes an upper bound of the cost function \( h(\mathcal{F}_{k}, e_{k}, b) \).

These three steps are investigated, respectively, in the following subsections.

2.2.1 Position determination for direct regulations

When the number of non-zero elements of vector \( x \), denote it by \( k \), is given, there are in principle \( C_{n-1}^{k} \) possibilities to locate these non-zero elements. Here, \( C_{n}^{k} \) denotes the combinatorial number of selecting \( k \) elements from the set \( \{1, 2, ..., n\} \). Note that for a fixed \( k \), \( C_{n-1}^{k} \) increases exponentially with the increment of \( n \). This implies that when a large-scale system is under investigation, in other words, when \( n \) is large, it is usually computationally intractable to consider all these combinations in order to find the optimal locations of the non-zero elements. As a matter of fact, according to our experiences, computation time is currently prohibitive if four direct regulations should be searched for a chemical element within a network having more than 100 species. On the other hand, it is clear from Equation (5) that when there are only \( k \) non-zero elements in vector \( x \), then, only \( k \) columns of matrix \( A \) are used to fit the experimental data of vector \( b \). Denote the \( i \)-th column of matrix \( A \) from the left by \( a_{i} \), the \( i \)-th element of vector \( x \) from the ceiling by \( x_{i} \), \( i = 1, 2, ..., n-1 \). Moreover, assume that the \( j_{p} \)-th element of vector \( x \) has been determined to be non-zero, \( a_{j_{p}} \). Define matrix \( A \) and vector \( \hat{x} \), respectively, as \( A = \{a_{i} \} \) and \( \hat{x} = \{x_{i} \} \).

Then, the first constraint of Equation (5) can be rewritten as

\[
[A - \hat{x} e_{n}] b = e_{n}
\]

in which \( e_{n} = [e_{n}, e_{n}, ..., e_{n}] \) and \( e_{n} \) represents measurement errors of vector \( x_{j} \), \( j = 1, 2, ..., n-1 \).

Note that under this situation, there are no longer restrictions on \( e_{n} \) for \( j \neq j_{p} \). According to the definition of the cost function \( h(\mathcal{F}_{k}, e_{k}, b) \), the corresponding optimal \( e_{n} \) is the zero vector. It can therefore be declared that

\[
j_{p} \in \text{arg} \min_{1 \leq j \leq n} \text{tr}[(\hat{x} a_{j})^{2} (e_{n} e_{n}^{T})]
\]

\( (7) \)

The minimization problem on the right-hand side of the above equation is widely known as the TLS, and has been well settled. In fact, the following results have been established for a long time (Globus and Van Loan, 1989; Van Huffel and Vandewalle, 1991).

\[
j_{p} \in \text{arg} \min_{1 \leq j \leq n} \text{tr}[(\hat{x} a_{j})^{2} (e_{n} e_{n}^{T})] = \epsilon_{2}^{-1} (\hat{\theta} b)
\]

in which \( \epsilon_{2}^{-1} \) stands for the minimal singular value of a matrix. Note that every singular value of a matrix is non-negative. Therefore, the optimal non-zero element position determination problem can be mathematically expressed as

\[
j_{p} \in \text{arg} \min_{1 \leq j \leq n} \epsilon_{2}^{-1} (\hat{\theta} b)
\]

\( (8) \)

While the above results are elegant, they cannot be directly applied to the optimal position determination of the non-zero elements of vector \( x \), as the corresponding minimization problem is still a combinatorial optimization one for which \( b \) is generally hard to find a globally or locally convergent algorithm with polynomial computational complexities. On the other hand, let range(b) and range(\( \hat{x} \)) denote, respectively, the subspaces spanned by vector \( b \) and vectors of matrix \( A \). Then, it can be proved that there is an upper bound of \( \epsilon_{2}^{-1} (\hat{\theta} b) \) that is proportional to the sine of the half of the angle between range(\( \hat{x} \)) and range(b).

More precisely, we have the next theorem, while its proof is deferred to Appendix B of the Supplementary Material.

**Theorem 1.** Let \( \theta_{b} \) represent the angle between range(\( \hat{x} \)) and range(b), \( |\hat{\theta}_{b}| \) the Euclidean norm of vector \( b \). Then,

\[
\epsilon_{2}^{-1} (\hat{\theta} b) \leq 2|\hat{\theta}_{b}| \sin \frac{\theta_{b}}{2}
\]

\( (9) \)

In addition, define matrix \( \hat{A} \) and vector \( \hat{b} \), respectively, as \( \hat{A} = A A^{-1/2} \) and \( \hat{b} = b b^{-1/2} \), and denote the \( j_{p} \)-th column of matrix \( \hat{A} \) from the left by \( \hat{a}_{j} \), \( j = 1, 2, ..., n-1 \). Then, we have the following results about subspace angle minimization, while their proof is given in Appendix C of the Supplementary Material.

**Theorem 2.** Let \( \theta_{b} \) represent the angle between the subspace spanned by vectors \( \hat{a}_{j_{p}} \) and the subspace spanned by \( b \). Moreover, let \( j_{p}^{(\hat{b})} \), \( a_{p}^{(\hat{b})} \), \( a = 1, 2, ..., k \), stand for the positions of the elements of vector \( \hat{A}^{k} \hat{b} \) with the first \( \hat{B} \) biggest magnitudes. Then,

\[
|\hat{b}^{(\hat{b})} a_{p}^{(\hat{b})}| = \text{arg} \min_{1 \leq j \leq n} \theta_{b}
\]

\( (10) \)

From these two theorems, it can be seen that if the minimization of the minimal singular value of \( \hat{A} \hat{b} \) is replaced by the minimization of \( \theta_{b} \), then, the optimal \( j_{p}^{(\hat{b})} \) can be expressed analytically, and therefore the problem of the exponential increment of computation complexities is avoided.
A derivation of this inequality is given in Appendix D of the Supplementary Material. Therefore, through selecting $\mathbf{J}^{w,0}$ that maximizes $\sum_{i=1}^{n} v_i^2$, an upper bound of $\hat{\theta}_i$ and therefore an upper bound of $\mathbf{g}(\hat{\mathbf{A}} \mathbf{b})$, has been minimized.

It is worthwhile to point out that although relations between $\hat{\theta}_i$ and $\hat{\mathbf{A}}$ are still not very clear, computation experiences show that compared with maximization of $\sum_{i=1}^{n} v_i^2$, minimization of $\hat{\theta}_i$ generally leads to better estimation performances in cellular network identifications. This is illustrated by examples in the next section. However, theoretical reasons for this phenomenon are still under investigations.

2.2.2 Estimation of regulation coefficients

When locations of non-zero elements of vector $\mathbf{x}$ have been determined, their values can be directly estimated, methods are available. On the other hand, it is obvious that the first term of the above cost function is a non-negative and monotonically decreasing function of the above $\mathbf{x}$, this optimum can be obtained through linear searches for which many examples, which include an artificially constructed large-scale linear system, a MAPK pathway model of ordinary nonlinear differential equations, some in silico datasets from the DREAM initiative (Prill et al., 2010), and some in vivo datasets of Gardner et al. (2003) and Cantone et al. (2009). Note that in predicting the behaviors of a network, not only structure of the network, but also regulation directions and strengths among its elements, play important roles. In addition to some widely adopted specifications in network inferences, such as ROC (receiver operating characteristics) curve, PR (precision recall) curve, AUROC (area under a ROC curve), AUPR (area under a PR curve), PPV (positive predictive value), Se (sensitivity), FP (false positive) rate (Cantone et al., 2010), and some in vivo datasets of Gardner et al. (2003) and Cantone et al. (2009). Note that in predicting the behaviors of a network, not only structure of the network, but also regulation directions and strengths among its elements, play important roles. In addition to some widely adopted specifications in network inferences, such as ROC (receiver operating characteristics) curve, PR (precision recall) curve, AUROC (area under a ROC curve), AUPR (area under a PR curve), PPV (positive predictive value), Se (sensitivity), FP (false positive) rate (Cantone et al., 2009; Gardner and Faith, 2005; Prill et al., 2010), etc., the following three specifications are also adopted in this article, which reflect the rate of false sign errors and parameter estimation accuracies, respectively.

To be more specific, consider a cellular network consisting of $n$ chemical elements. Let $x_{ij}, i,j=1,2,\ldots,n$, denote the actual direct effect of the $j$-th chemical element on the $i$-th chemical element, and $\hat{x}_{ij}$ its estimate. Moreover, let $N$ and $P$ represent the total numbers of $x_{ij}=0$ and $x_{ij} \neq 0$, respectively. Furthermore, define $FP$ (false positive number), $FN$ (false negative number), $FS$ (false sign number), $LA$ (low accuracy number) and $AA$ (acceptable accuracy number), respectively, as follows.

- **FP:** the total number of the case that $x_{ij}=0$ but $|\hat{x}_{ij}| > \rho_n$;
- **FN:** the total number of the case that $x_{ij} \neq 0$ but $|\hat{x}_{ij}| \leq \rho_n$;
- **FS:** the total number of the case that $x_{ij} \hat{x}_{ij} < 0$;
- **LA:** the total number of the case that $x_{ij} \neq 0$ and $\hat{x}_{ij} \neq 0$, but $|\hat{x}_{ij} - x_{ij}|/|x_{ij}| > \rho_s$;
- **AA:** the total number of the case that $x_{ij}=0$ and $|\hat{x}_{ij}| \leq \rho_n$, as well as the case that $x_{ij} \neq 0$, $\hat{x}_{ij} \neq 0$ and $|\hat{x}_{ij} - x_{ij}|/|x_{ij}| > \rho_s$.

Algorithm I for Cellular Network Inference:

1. Initialize vector $\mathbf{s}$ and $\mathbf{J}(0)$, respectively, as $\mathbf{s} = 0$ and a large positive number, for example, $\mathbf{J}(0) = 10^{10}$.
2. Compute matrix $\mathbf{A}$ and vector $\hat{\mathbf{A}}\mathbf{b}$. Denote the $i$-th row element of $\hat{\mathbf{A}}\mathbf{b}$ by $\gamma_i$, and assume that $|\gamma_i| \geq |\gamma_{i+1}| \geq \cdots \geq |\gamma_{m-1}|$.
3. Construct matrix $\hat{\mathbf{A}}$, let $\hat{\mathbf{A}} = [\mathbf{a}_1 \mathbf{a}_2 \cdots \mathbf{a}_n]$. Compute the value of $J(\hat{k})$ which is defined as
   \[
   J(\hat{k}) = \frac{\mathbf{s}^T(\hat{\mathbf{A}}\mathbf{b})}{2\rho_n} - \left\{ \begin{array}{ll}
   \ln(\rho_n), & \hat{k} \leq k_{\text{est}} \\
   \ln(\rho_n) - \ln(k_{\text{est}} - k), & k_{\text{est}} < k \leq n
   \end{array} \right.
   \]
   (14)
   (4) If $J(\hat{k}) \leq J(\hat{k}-1)$, replace $\hat{k}$ by $\hat{k}+1$ and repeat the above Step 3. If $J(\hat{k}) \geq J(\hat{k}-1)$, go to the next step.
5. Perform singular value decomposition of matrix $[\hat{\mathbf{A}}_v \mathbf{b}]$. Denote its right singular vector associated with its minimum singular value $\mathbf{g}([\hat{\mathbf{A}}_v \mathbf{b}])$ by $v$. Let $v_i$ represent the $i$-th row element of vector $v$. Replace the $j$-th row element of vector $\mathbf{x}$ by $-v_i$, $j = 1, 2, \ldots, k-1$.

Algorithm II for Cellular Network Inference: this algorithm is completely the same as that of Algorithm I, except that in the second step, matrix $\mathbf{A}$ is replaced by matrix $\hat{\mathbf{A}}$.
Here, $\rho_u$ and $\rho_r$ represent acceptable magnitude bounds, respectively, for absolute and relative parametric estimation errors. Their values may vary with applications.

With the above numbers, the following three specifications are defined, which are used in this article to evaluate the effectiveness of a network identification algorithm.

- FS rate (rate of false sign errors): $\text{FS}/P$;
- LA rate (rate of estimates with low accuracy): $\text{LA}/P$;
- AA rate (rate of estimates with acceptable accuracy): $\text{AA}/(P+N)$.

Moreover, note that the sum of the numbers of false negative estimates and true positive estimates equals to $P$. We therefore have $\text{PPV} = \frac{P - \text{FN}}{P}$ and $\text{FP rate} = \frac{\text{FP}}{N}$.

### 3.1 Example 1: an artificial linear network

In this subsection, the suggested algorithms are tested on an artificially constructed network in which all the adopted assumptions are completely satisfied. More precisely, a network with 100 elements is utilized which obeys the power law with $k_{\text{init}} = 1$ and $\gamma = 2.5$. The number of direct regulations, as well as regulation coefficients, to a network element, are randomly generated. Moreover, 500 trails have been performed to investigate the statistical properties of the estimates.

The numerical simulation settings are as follows. Denote the matrix consisting of the actual interactions among network elements by $X_0$. Every column of $X_0$ is independently generated according to the next three steps.

- At first, number of non-zero elements of a column is randomly generated according to the power law with $k_{\text{init}} = 1$ and $\gamma = 2.5$. Assume it to be $k$.
- Locations of non-zero elements are determined by the Matlab m-file randperm.m for random permutations. That is, elements of the set $\{1, 2, \ldots, 100\}$ are at first randomly permuted, and then the first $k$ elements are adopted as the numbers of the rows in this column with a non-zero entry. Denote them by $j_{\alpha}^{\beta}$.
- The entry of the $j_{\alpha}^{\beta}$-th row of this column is generated independently according to a uniform distribution over $[-2^{-\beta}, 2^{-\beta}]$, $\alpha = 1, \ldots, k$. All the other entries are assigned to be zero.

On the other hand, a $10^3 \times 10^2$ dimensional real matrix $A_0$ is randomly generated whose entries are independent and uniform samples from $[1, 5]$. Moreover, two $10^3 \times 10^2$ dimensional matrices $e_A$ and $e_B$ are generated with independent samples of a normal distribution $N(0, 4)$. Finally, matrices $A$ and $B$ are generated as $A = A_0 + e_A$ and $B = A_0X_0 + e_B$.

After the production of matrices $A$ and $B$, every column of $X_0$ is estimated independently on the basis of the data of matrix $A$ and that of the corresponding column of matrix $B$. Estimation performances are evaluated using the aforementioned specifications with $\rho_u = 10^{-3}$ and $\rho_r = 5\%$.

In Table 1, both the empirical expectations and the empirical standard deviations (SDs) are reported for the aforementioned specifications, on the basis of 500 independent simulation trails. In Figure 1, the empirical expectations of the corresponding ROC and PR curves are shown, together with their empirical SDs.

#### Table 1. Estimation Performances for a linear network with 100 species

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>FP rate</td>
<td>0.0184</td>
<td>0.0027</td>
</tr>
<tr>
<td>FS rate</td>
<td>0.0231</td>
<td>0.0160</td>
</tr>
<tr>
<td>LA rate</td>
<td>0.5459</td>
<td>0.0639</td>
</tr>
<tr>
<td>AA rate</td>
<td>0.9666</td>
<td>0.0044</td>
</tr>
<tr>
<td>PPV</td>
<td>0.4101</td>
<td>0.0393</td>
</tr>
<tr>
<td>Se</td>
<td>0.6992</td>
<td>0.0802</td>
</tr>
<tr>
<td>AUROC</td>
<td>0.8457</td>
<td>0.0408</td>
</tr>
<tr>
<td>AUROC</td>
<td>0.6038</td>
<td>0.0954</td>
</tr>
</tbody>
</table>

*Significance level: $\alpha = 2$, elements of $X_0$ are from $[1, 5]$.

#### Figure 1

Averaged ROC and PR curves of Example 1 with a fixed noise SD $\sigma = 2.0000$. (a) Averaged ROC curve. (b) Averaged PR curve. Filled circle, Algorithm I; filled triangle, Algorithm II, filled square, TLS.

To investigate influences of noise strengths on estimation performances, the above simulations have been repeated with different noise level $\sigma$. In actual computations, the value of $\sigma$ is taken from the set of $\{0.01, 0.10, 0.22, 0.50, 0.71, 1.00, 1.25, 1.50, 1.55, 1.60, 1.73, 1.90, 2.00, 3.00, 3.50, 4.00, 5.00, 6.00\}$.

Figure 2 shows variations of the empirical expectations of PPV and FP rate of the estimates, together with their SDs. Variations of other specifications, such as AUROC, AUPR, etc., are provided in Appendix E of the Supplementary Material.
Algorithm II with the specifications Sensitivity, AUROC and AUPR. Not very surprising, as in the extreme situation, that is, when $\sigma$ are adopted in performance assessment, then, when measurement errors are not very significant, the above simulations show that the TLS method performs better than both Algorithms I and II. This is not very surprising, as in the extreme situation, that is, when $\sigma = 0$, Equation (2) is replaced by $Ax=b$ for a linear network, but there is no guarantee that minimization of $\theta_2$ or $\theta_2$ always leads to actual positions of non-zero elements of vector $x$.

3.2 A nonlinear MAPK pathway model

In actual applications, it can rarely be expected that assumptions adopted in algorithm developments are perfectly satisfied. In this subsection, the suggested estimation algorithms are tested on a MAPK pathway model, which consists of 103 chemical elements and is described by a set of first-order ordinary nonlinear differential equations. These equations take completely the same form as that of Equation (S1) of the Supplementary Material. This model is originally built in Schoeberl et al. (2002) and capable of explaining many biological observations. The structure of the model is given in Figure S4 of Appendix F of the Supplementary Material.

Causal interactions among chemical elements of this network are identified among one of its steady states or stable equilibria. To apply the suggested algorithms, parameters for the power law are required. To obtain this information and evaluate estimation accuracies, the Jacobian matrix of the nonlinear function vector $U(Ax_k, \delta_k, \mu_k[257])_{103}$ is first computed at the selected stable equilibrium $x_0$, which is further used to calculate the actual interactions among chemical elements using Equation (S3) in the Supplementary Material. Based on these information, parameters of the power law are estimated through counting the number of non-zero $r_{ij}$ with a fixed $i$, $i,j=1,2,\ldots,103$, and fitting the logarithm of the corresponding empirical probabilities. Using this method, $\gamma=1.4718$ and $k_{\min} = 1$ are obtained. On the other hand, using the likelihood maximization-based method of Clauset et al. (2009), $\gamma=0.8000$ and $k_{\min} = 1$ are obtained. Figure S5 of Appendix F of the Supplementary Material shows the calculated empirical probabilities and the estimated power laws.

It is worthwhile to mention that although there are 103 chemical elements in the MAPK pathway model, some of the species do not have direct influences on their own concentration change speeds, that is, for these species, $\frac{\partial x_i}{\partial x_i}=0$. In this case, the interference parameter $r_{ij}$ of Equation (S3) in the Supplementary Material is not well defined, and therefore the suggested methods, as well as the TLS method, cannot be applied. On the other hand, numerical simulations show that for some species with self-concentration influences, the first-order Taylor series approximation of Equation (S6) in the Supplementary Material is not valid, which happens when $0(||d(x_0)||^2 ||b||_2^2)$ is of the same order of or greater than the magnitude of $2(||d(x_0)|| ||b||_2 ||\delta||_2)$. In fact, using the definition of $r_{ij}$, Equation (S9) in the Supplementary Material can be rewritten as

$$\sum_{k=1, k\neq i}^{m} r_{ik} R_{U_k} = R_{U_i} \left[ 1 + \kappa_i(x^{(i)}|d^{(i)}, \delta_i) \right], \quad i=1,2,\ldots,m$$

in which $\kappa_i(x^{(i)}|d^{(i)}, \delta_i) = O(||d^{(i)}||^2_2 ||\delta||_2 ||\delta_i||^2_2)$.

Simulations show that at the selected stable equilibrium, for some species, $\kappa_i(x^{(i)}|d^{(i)}, \delta_i)$ has a magnitude $>0.5$, and in the worst case, this magnitude is very close to $10^5$. Under these situations, the approximation of Equation (S9) in the Supplementary Material is no longer valid.
AUROC and AUPR are considered, the TLS method has a better AUROC 0.7765 0.7807 0.7012 0.6982 0.7745 AA rate 0.7806 0.7995 0.3692 0.4033 0.0207 Table 2. Estimation performances for a MAPK pathway model

<table>
<thead>
<tr>
<th></th>
<th>Algorithm I</th>
<th>Algorithm II</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \gamma = 0.8000 )</td>
<td>( \gamma = 1.4718 )</td>
<td>( \gamma = 0.8000 )</td>
</tr>
<tr>
<td>FP rate</td>
<td>0.1949</td>
<td>0.1748</td>
</tr>
<tr>
<td>FS rate</td>
<td>0.1005</td>
<td>0.1096</td>
</tr>
<tr>
<td>LA rate</td>
<td>0.2785</td>
<td>0.2785</td>
</tr>
<tr>
<td>AA rate</td>
<td>0.7806</td>
<td>0.7995</td>
</tr>
<tr>
<td>PPV</td>
<td>0.1706</td>
<td>0.1865</td>
</tr>
<tr>
<td>Se</td>
<td>0.6575</td>
<td>0.6575</td>
</tr>
<tr>
<td>Se</td>
<td>0.4804</td>
<td>0.4806</td>
</tr>
</tbody>
</table>

\( N = 3592, P = 219 \)

Based on these considerations, 37 species are finally chosen whose steady-state concentration change can be appropriately approximated by the first-order Taylor series and whose concentration change speed can be directly influenced by its own concentration.

In data generations, kinetic parameters \( p_j \), and initial values of \( x_i \) are changed in a way similar to that of Andre et al. (2005) and Khodolenko et al. (2002). That is, when direct influences on the \( i \)-th species are to be estimated, only the values of these \( p_j \), \( j \in \{ 1, 2, \ldots, 247 \} \), are permitted to be changed which do not explicitly alter the value of the nonlinear function \( f_i(x, p) \). More specifically, an appropriate \( p_j \), is selected together with \( 8 \sim 12 \) \( x_i \) that are, respectively, changed to 0.9999 or \( 0.1 \) at the initial time. Here, both \( \alpha_j \) and \( \beta_j \) are independent and uniform random samples from \([0.9, 1]\). Steady-state concentration of every species in the network is calculated before and after a perturbation using the toolbox Simulink of the commercial software MATLAB. To every calculated relative concentration change at the steady states, that is, \( \delta_i \), a random number is added which is independently generated according to the normal distribution with mean 0 and SD \( 10^{-5} \). Estimation performances with parameters or initial values of \( x_i \) are shown in Table 2 when experiment data length takes an integer between 142 and 146, while Figure 3 shows the corresponding ROC and PR curves. The actual data length varies slightly with the number of direct regulations from external perturbations on the chemical element under investigation. From this table, it is clear that the TLS method is superior to the methods of this article in false negative errors. On the other hand, in both false positive errors and parameter estimation accuracy, the methods of this article have prominent advantages. Moreover, in either false sign errors or estimates with large relative errors, both Algorithms I and II are also better than the TLS method. Furthermore, except Sensitivity, Algorithm I outperforms Algorithm II with respect to all the other adopted specifications. These results are consistent with those obtained in the previous linear network simulation example. However, when AUROC and AUPR are considered, the TLS method has a better performance than Algorithm II.

It is also clear from these estimation results that although curve fitting and likelihood maximization lead to very different estimates for parameters of the power law, these differences do not lead to significant changes in network estimation performances, no matter whether Algorithm I or II is used. This may imply that these two estimation procedures are robust against estimation errors of these parameters. In fact, several other artificial \( k_{\text{min}} \) and \( \gamma \) have also been utilized, and consistent observations have been obtained.

In addition, a detailed analysis of the false negative errors of Algorithm I with the curve fitting-based parametric estimates of the power law, shows that there are totally 75 direct regulation coefficients that have been incorrectly estimated as zero. With respect to these regulations, the results of the corresponding TLS estimates are as follows.

- Forty-one of them have a sign different from its actual value.
- In the remaining thirty-four estimates, only one of them has a relative error belonging \([-5, 5]\). As a matter of fact, most of them have a relative error with a magnitude > 100%.

Consistent phenomena have also been observed for the false negative estimates of Algorithm I with the likelihood maximization based estimates of the power law parameters, as well as those of Algorithm II. These results are summarized in Table S2 of the Supplementary Material.

To investigate convergence properties of the suggested algorithms, the species numbered 11 in Schoeberl et al. (2002), that is, (EGF-EGFR1), is selected for which its direct regulation coefficients from other species are estimated with the data length \( m \) varying from \( m = 20 \) to \( m = 10^4 \). In these investigations,
This implies that results for the TLS estimate in these figures are valid only \( \hat{\theta} \text{triangle}, \text{Algorithm II}; \text{filled square}, \text{TLS}. \text{(continuous line: network estimations, errors are also computed for the estimated of Algorithm II are poor.} \)

But performances of both Algorithms I and II are still very poor compared to the reported participants. It appears safe to declare that Algorithm I works well with the DREAM3 Size 100 subchallenges. In addition, our estimation procedures have significantly lower computational complexities. To be more specific, for the best performer of these subchallenges, it is reported that approximately 78 h have been consumed to obtain an estimate with a high-end cluster. But for both Algorithms I and II, <2 s were required even if a personal computer is utilized, which is equipped with 2.00 GB RAMs and two 2.40 GHz Intel(R) Core(TM) Quad CPUs.

However, performances of both Algorithms I and II are still very poor for the DREAM4 Size 100 subchallenges. This may possibly be due to that the adopted assumptions on noises \( \varepsilon \) have been seriously deteriorated.

On the other hand, except Sensitivity, both Algorithms I and II outperform the TLS method in all these subchallenges. Moreover, except in some specifications such as FS rate and FP rate, etc., Algorithm II is slightly superior to Algorithm I. Moreover, except in some specifications such as FS rate and FP rate, etc., Algorithm II is slightly superior to Algorithm I. However, it should be emphasized that these results are not directly comparable because the reported participants were completely blind to the structures and dynamics of the networks.

\[ \text{Expectation of squared estimation errors} \]

\[ \text{Squares of estimation bias} \]

\[ \text{Data length} \]

\[ \text{Number of direct regulations} \]

In addition, our estimation procedures have significantly lower computational complexities. To be more specific, for the best performer of these subchallenges, it is reported that approximately 78 h have been consumed to obtain an estimate with a high-end cluster. But for both Algorithms I and II, <2 s were required even if a personal computer is utilized, which is equipped with 2.00 GB RAMs and two 2.40 GHz Intel(R) Core(TM) Quad CPUs.

However, performances of both Algorithms I and II are still very poor for the DREAM4 Size 100 subchallenges. This may possibly be due to that the adopted assumptions on noises \( \varepsilon \) have been seriously deteriorated.

On the other hand, except Sensitivity, both Algorithms I and II outperform the TLS method in all these subchallenges. Moreover, except in some specifications such as FS rate and FP rate, etc., Algorithm II is slightly superior to Algorithm I. However, it should be emphasized that these results are not directly comparable because the reported participants were completely blind to the structures and dynamics of the networks.

\[ \text{Expectation of squared estimation errors} \]

\[ \text{Squares of estimation bias} \]

\[ \text{Data length} \]

\[ \text{Number of direct regulations} \]

In addition, our estimation procedures have significantly lower computational complexities. To be more specific, for the best performer of these subchallenges, it is reported that approximately 78 h have been consumed to obtain an estimate with a high-end cluster. But for both Algorithms I and II, <2 s were required even if a personal computer is utilized, which is equipped with 2.00 GB RAMs and two 2.40 GHz Intel(R) Core(TM) Quad CPUs.

However, performances of both Algorithms I and II are still very poor for the DREAM4 Size 100 subchallenges. This may possibly be due to that the adopted assumptions on noises \( \varepsilon \) have been seriously deteriorated.

On the other hand, except Sensitivity, both Algorithms I and II outperform the TLS method in all these subchallenges. Moreover, except in some specifications such as FS rate and FP rate, etc., Algorithm II is slightly superior to Algorithm I. However, it should be emphasized that these results are not directly comparable because the reported participants were completely blind to the structures and dynamics of the networks.

\[ \text{Expectation of squared estimation errors} \]

\[ \text{Squares of estimation bias} \]

\[ \text{Data length} \]

\[ \text{Number of direct regulations} \]

In addition, our estimation procedures have significantly lower computational complexities. To be more specific, for the best performer of these subchallenges, it is reported that approximately 78 h have been consumed to obtain an estimate with a high-end cluster. But for both Algorithms I and II, <2 s were required even if a personal computer is utilized, which is equipped with 2.00 GB RAMs and two 2.40 GHz Intel(R) Core(TM) Quad CPUs.

However, performances of both Algorithms I and II are still very poor for the DREAM4 Size 100 subchallenges. This may possibly be due to that the adopted assumptions on noises \( \varepsilon \) have been seriously deteriorated.

On the other hand, except Sensitivity, both Algorithms I and II outperform the TLS method in all these subchallenges. Moreover, except in some specifications such as FS rate and FP rate, etc., Algorithm II is slightly superior to Algorithm I. However, it should be emphasized that these results are not directly comparable because the reported participants were completely blind to the structures and dynamics of the networks.

\[ \text{Expectation of squared estimation errors} \]

\[ \text{Squares of estimation bias} \]

\[ \text{Data length} \]

\[ \text{Number of direct regulations} \]

In addition, our estimation procedures have significantly lower computational complexities. To be more specific, for the best performer of these subchallenges, it is reported that approximately 78 h have been consumed to obtain an estimate with a high-end cluster. But for both Algorithms I and II, <2 s were required even if a personal computer is utilized, which is equipped with 2.00 GB RAMs and two 2.40 GHz Intel(R) Core(TM) Quad CPUs.

However, performances of both Algorithms I and II are still very poor for the DREAM4 Size 100 subchallenges. This may possibly be due to that the adopted assumptions on noises \( \varepsilon \) have been seriously deteriorated.

On the other hand, except Sensitivity, both Algorithms I and II outperform the TLS method in all these subchallenges. Moreover, except in some specifications such as FS rate and FP rate, etc., Algorithm II is slightly superior to Algorithm I. However, it should be emphasized that these results are not directly comparable because the reported participants were completely blind to the structures and dynamics of the networks.

\[ \text{Expectation of squared estimation errors} \]

\[ \text{Squares of estimation bias} \]

\[ \text{Data length} \]

\[ \text{Number of direct regulations} \]

In addition, our estimation procedures have significantly lower computational complexities. To be more specific, for the best performer of these subchallenges, it is reported that approximately 78 h have been consumed to obtain an estimate with a high-end cluster. But for both Algorithms I and II, <2 s were required even if a personal computer is utilized, which is equipped with 2.00 GB RAMs and two 2.40 GHz Intel(R) Core(TM) Quad CPUs.

However, performances of both Algorithms I and II are still very poor for the DREAM4 Size 100 subchallenges. This may possibly be due to that the adopted assumptions on noises \( \varepsilon \) have been seriously deteriorated.

On the other hand, except Sensitivity, both Algorithms I and II outperform the TLS method in all these subchallenges. Moreover, except in some specifications such as FS rate and FP rate, etc., Algorithm II is slightly superior to Algorithm I. However, it should be emphasized that these results are not directly comparable because the reported participants were completely blind to the structures and dynamics of the networks.

\[ \text{Expectation of squared estimation errors} \]

\[ \text{Squares of estimation bias} \]

\[ \text{Data length} \]

\[ \text{Number of direct regulations} \]

In addition, our estimation procedures have significantly lower computational complexities. To be more specific, for the best performer of these subchallenges, it is reported that approximately 78 h have been consumed to obtain an estimate with a high-end cluster. But for both Algorithms I and II, <2 s were required even if a personal computer is utilized, which is equipped with 2.00 GB RAMs and two 2.40 GHz Intel(R) Core(TM) Quad CPUs.

However, performances of both Algorithms I and II are still very poor for the DREAM4 Size 100 subchallenges. This may possibly be due to that the adopted assumptions on noises \( \varepsilon \) have been seriously deteriorated.

On the other hand, except Sensitivity, both Algorithms I and II outperform the TLS method in all these subchallenges. Moreover, except in some specifications such as FS rate and FP rate, etc., Algorithm II is slightly superior to Algorithm I. However, it should be emphasized that these results are not directly comparable because the reported participants were completely blind to the structures and dynamics of the networks.

\[ \text{Expectation of squared estimation errors} \]

\[ \text{Squares of estimation bias} \]

\[ \text{Data length} \]

\[ \text{Number of direct regulations} \]

In addition, our estimation procedures have significantly lower computational complexities. To be more specific, for the best performer of these subchallenges, it is reported that approximately 78 h have been consumed to obtain an estimate with a high-end cluster. But for both Algorithms I and II, <2 s were required even if a personal computer is utilized, which is equipped with 2.00 GB RAMs and two 2.40 GHz Intel(R) Core(TM) Quad CPUs.

However, performances of both Algorithms I and II are still very poor for the DREAM4 Size 100 subchallenges. This may possibly be due to that the adopted assumptions on noises \( \varepsilon \) have been seriously deteriorated.

On the other hand, except Sensitivity, both Algorithms I and II outperform the TLS method in all these subchallenges. Moreover, except in some specifications such as FS rate and FP rate, etc., Algorithm II is slightly superior to Algorithm I. However, it should be emphasized that these results are not directly comparable because the reported participants were completely blind to the structures and dynamics of the networks.

\[ \text{Expectation of squared estimation errors} \]

\[ \text{Squares of estimation bias} \]

\[ \text{Data length} \]

\[ \text{Number of direct regulations} \]

In addition, our estimation procedures have significantly lower computational complexities. To be more specific, for the best performer of these subchallenges, it is reported that approximately 78 h have been consumed to obtain an estimate with a high-end cluster. But for both Algorithms I and II, <2 s were required even if a personal computer is utilized, which is equipped with 2.00 GB RAMs and two 2.40 GHz Intel(R) Core(TM) Quad CPUs.

However, performances of both Algorithms I and II are still very poor for the DREAM4 Size 100 subchallenges. This may possibly be due to that the adopted assumptions on noises \( \varepsilon \) have been seriously deteriorated.

On the other hand, except Sensitivity, both Algorithms I and II outperform the TLS method in all these subchallenges. Moreover, except in some specifications such as FS rate and FP rate, etc., Algorithm II is slightly superior to Algorithm I. However, it should be emphasized that these results are not directly comparable because the reported participants were completely blind to the structures and dynamics of the networks.

\[ \text{Expectation of squared estimation errors} \]

\[ \text{Squares of estimation bias} \]

\[ \text{Data length} \]

\[ \text{Number of direct regulations} \]
approaches) network (Cantone et al., 2009). The SOS test network and the IRMA network have, respectively, only 9 and 5 genes. Generally, the power law cannot be applied to networks with a size of this order. But when the distribution of Equation (3) is utilized, we found that it is still efficient in improving performances of TLS estimates. The details of the estimation results are provided in Appendix H of the Supplementary Material. In these cases, as the size of the involved network is not very large, it is feasible to consider all the possible positions of non-zero elements of vector \( \mathbf{x} \) when the number of direct regulations is prescribed. To see the conservativeness of the upper bound of Equation (14), the corresponding estimation results have also been included in the Supplementary Material.

The results of Appendix H of the Supplementary Material show that performances of the suggested methods are close but slightly better than the reported estimates. On the other hand, combinatorial optimization-based estimate does not provide significant estimation performance improvements, although its computation time is almost 10 times longer than that of both Algorithms I and II.

In the above estimations, only curve fitting-based estimates for the power law parameters are utilized. But it has also been found that final estimation performances do not change appreciably even if the likelihood maximization-based estimates for these parameters are utilized.

4 CONCLUSION

In this article, possibilities have been investigated for incorporating the so-called power law into identification of direct causal regulations in a large-scale cellular network. Under the assumption that measurement errors are independently and normally distributed, a likelihood maximization approach is suggested. On the basis of relations between subspace angles and matrix singular values, two procedures are suggested to estimate these interactions. These methods are applied to an artificially constructed linear large-scale network, a MAPK pathway model, some recently developed performance assessment networks and the \textit{in vivo} data set of a nine-gene network and a five-gene network. Comparisons with the widely adopted TLS method show that the suggested methods have distinguished advantages on both reduction of false positive errors and improvement of parametric estimation accuracy. Moreover, one of the supposed algorithms has a much faster convergence speed when either estimation error or estimation bias is considered.

While most of the reported results are encouraging, the developed estimation procedures are still far from satisfaction of practical application requirements. This has been made very clear by the unsatisfactory performances with the sub-challenges of DREAM4. As further researches, it is interesting to investigate parameter estimation from experiment data for the power law, simultaneous identification of direct causal regulations for multiple species, appropriate statistical descriptions of noise and measurement errors, as well as extending the procedures to estimations with other kinds of experiment data, such as quasi-steady-state data, time series data with non-uniform sampling, etc.

ACKNOWLEDGEMENTS

The authors would like to express their gratitude to Profs Q.C. Zhao and S. Li, Tsinghua University, China, for various invaluable discussions.

Funding: National Natural Science Foundation of China (grant numbers 60625305, 60721003, in parts); 973 Program of China (grant number 2009CB320602, in parts).

Conflict of Interest: none declared.

REFERENCES


Fernandez, F. et al. (2007) Bayesian approaches to reverse engineer cellular systems: a simulation study on nonlinear Gaussian networks. BMC Bioinformatics, 8 (Suppl. 1), S2.


