JOURNAL OF COMPUTATIONAL BIOLOGY
Volume 22, Number 11, 2015
Mary Ann Liebert, Inc.
Pp. 1025–1033
DOI: 10.1089/cmb.2015.0072

Bayesian Inference for Duplication–Mutation with Complementarity Network Models

AJAY JASRA,[1] ADAM PERSING,[2] ALEXANDROS BESKOS,[2]
KARI HEINE,[2] and MARIA DE IORIO[2]

ABSTRACT

We observe an undirected graph G without multiple edges and self-loops, which is to represent a protein–protein interaction (PPI) network. We assume that G evolved under the duplication–mutation with complementarity (DMC) model from a seed graph, G_0, and we also observe the binary forest Γ that represents the duplication history of G. A posterior density for the DMC model parameters is established, and we outline a sampling strategy by which one can perform Bayesian inference; that sampling strategy employs a particle marginal Metropolis–Hastings (PMMH) algorithm. We test our methodology on numerical examples to demonstrate a high accuracy and precision in the inference of the DMC model's mutation and homodimerization parameters.

Key words: duplication–mutation with complementarity (DMC) model, particle marginal Metropolis–Hastings (PMMH), protein–protein interaction (PPI) network, sequential Monte Carlo (SMC).

1. INTRODUCTION

As a result of breakthroughs in biotechnology and high-throughput experiments thousands of regulatory and protein–protein interactions have been revealed, and genome-wide protein–protein interaction (PPI) data are now available. Protein–protein interactions are one of the most important components of biological networks, as they are fundamental to the functioning of cells. To gain a better understanding of why these interactions take place, it is necessary to view them from an evolutionary perspective. The evolutionary history of PPI networks can help answer many questions about how present-day networks have evolved and provide valuable insight into molecular mechanisms of network growth (Kreimer et al., 2008; Pereira-Leal et al., 2007). However, inferring network evolution history is a statistical and computational challenging problem as PPI networks of extant organisms provide only snapshots in time of the network evolution. There has been recent work on reconstructing ancestral interactions (e.g., Dutkowski and Tiuryn, 2007; Gibson and Goldberg, 2009; Patro et al., 2012). The main growth mechanism of PPI network is gene duplication and divergence (mutations) (Wagner, 2001); all proteins in a family evolve from a common ancestor through gene duplications and mutations, and the protein network reflects the entire history of the

[1]Department of Statistics & Applied Probability, National University of Singapore, Singapore, Singapore.
[2]Department of Statistical Science, University College London, London, United Kingdom.

genome evolution (Vazquez et al., 2003). In this article we follow Li et al. (2013), and we develop computational methods to infer the growth history and the parameters under the given model incorporating not only the topology of observed networks, but also the duplication history of the proteins contained in the networks. In their article, Li et al. (2013) propose a maximum likelihood approach. The authors establish a neat representation of the likelihood function, and it is this representation that is used in this article. The duplication history of the proteins can be inferred independently by phylogenetic analysis (Patro et al., 2012; Pinney et al., 2007).

The approach we adopt here is first to obtain a numerically stable estimate of the likelihood function, under fixed parameters; this is achieved via the sequential Monte Carlo (SMC) method (see Doucet et al., 2000, and Gordon et al., 1993). This approach can then be used to infer the parameters of the model, from a Bayesian perspective, as well as the growth history, via a Markov chain Monte Carlo (MCMC) method. To the best of our knowledge, this has not been considered in the literature, although related ideas have appeared for simpler models in Wang et al. (2014). Our computational strategy not only improves on likelihood estimation in comparison to Li et al. (2013), but also provides a natural setup to perform posterior inference on the parameters of interest.

This article is structured as follows. In section 2, we detail the model and associated computational method for statistical inference. In section 3, our numerical results are presented. In section 4, the article is concluded with some discussion of future work.

2. MODEL AND METHODS

We follow similar notation and exposition as in Li et al. (2013) to introduce the protein–protein interaction network, its duplication history, and the duplication–mutation with complementarity (DMC) model (Vazquez et al., 2003). In particular, the notions of adjacency and duplication are made concrete there. We also introduce the associated Bayesian inference problem with which this work is primarily concerned (i.e., that of inferring the parameters of the DMC model). We then describe a particle marginal Metropolis–Hastings (PMMH) algorithm (Andrieu et al., 2010) that can be used to perform such inference.

2.1. PPI network and DMC model

Consider an undirected graph G without multiple edges and self-loops, where the nodes represent proteins and the edges represent interactions between those proteins. Such a graph is called a PPI network, and as in Li et al. (2013), we denote the vertex set by $V(G)$, the edge set by $E(G)$, and the number of nodes in G by $|V(G)|$. All nodes that are adjacent to a node v (not including v itself) comprise the neighborhood of v, and that neighborhood is denoted by $N_G(v)$.

We assume that G evolved from a seed graph G_0 via a series of duplication, mutation, and homodimerization steps under a DMC model. Under the DMC model, at each time step t, the graph G_t evolves from G_{t-1} by the following processes in order:

1. The anchor node u_t is chosen uniformly at random from $V(G_{t-1})$, and a duplicate node v_t is added to G_{t-1} and connected to every member of $N_{G_{t-1}}(u_t)$. This is the duplication step, and it yields an intermediary graph denoted G^*_{t-1}.
2. For each $w \in N_{G^*_{t-1}}(u_t)$, we uniformly choose one of the two edges in $\{(u_t, w), (v_t, w)\} \subseteq E(G^*_{t-1})$ at random and delete it with probability $(1-p)$. This is the mutation step, and the parameter p is henceforth referred to as the mutation parameter.
3. The anchor node u_t and the duplicate node v_t are connected with probability p_c to finally obtain G_t. This is the homodimerization step, and the parameter p_c is henceforth known as the homodimerization parameter.

The DMC model is Markovian, and we denote the transition density at time t (which encompasses the three aforementioned steps) by $p_{\mathcal{M}}(G_t \mid G_{t-1})$, where $\mathcal{M} := (p, p_c)$. If we assign to a seed graph some prior density $p_{\mathcal{M}}(G_0)$, then the density of the observed graph G will be

$$p_{\mathcal{M}}(G) = \sum_{\mathcal{H}\setminus\{G_n\}} \left[p_{\mathcal{M}}(G_0) \prod_{t=1}^{n} p_{\mathcal{M}}(G_t \mid G_{t-1}) \right], \tag{1}$$

where $G = G_n$, $n = |V(G)| - |V(G_0)|$, and $\mathcal{H} = (G_0, G_1, \ldots, G_n = G)$ denotes the collection of growth histories. In this work, a seed graph will always be the graph consisting of two connected nodes; thus, $|V(G_0)| = 2$ and $p_{\mathcal{M}}(G_0) = 1$. Note that we are summing over all possible growth histories by which G can evolve from a seed graph. Also note that a growth history $\mathcal{H}$ induces a unique sequence of duplicate nodes, $\theta(\mathcal{H}) = (v_1, \ldots, v_n)$ (Li et al., 2013).

2.2. *Bayesian inference*

In practice, one will not have access to the parameters (p, p_c), and they must be inferred given G. Thus, in the Bayesian setting, our objective is to consider the posterior density

$$\pi(\mathcal{M} \mid G) \propto p(\mathcal{M}) p_{\mathcal{M}}(G), \tag{2}$$

where $p(\mathcal{M})$ is some proper prior for (p, p_c) that we assume can easily be computed (at least pointwise up to a normalizing constant).

The total number of growth histories grows exponentially with n (Li et al., 2013), and so any computations involving (1), and thus (2) [e.g., the evaluation of $p_{\mathcal{M}}(G)$], could potentially become very expensive. In the following sections, we reformulate the inference problem in the same manner as in Li et al. (2013) to alleviate this issue.

2.3. *Duplication history*

As in Li et al. (2013), let Γ be a binary forest, that is, a collection of rooted binary trees. The authors of Li et al. (2013) describe a scheme that encodes the duplication history of a growth history $\mathcal{H}$ within a series of duplication forests, $(\Gamma_0, \Gamma_1, \ldots, \Gamma_n)$, where each forest Γ_t corresponds to a graph G_t. We describe that scheme here.

Consider a trivial forest Γ_0, whose only two isolated trees each consist of a single node. Each of those isolated nodes will correspond to a node within the seed graph G_0. To build Γ_1 from Γ_0, one replaces an anchor node u_1 from Γ_0 with a subtree, $\{u_1, v_1\}$, consisting of two leaves (v_1 is the duplicate node including G_1 but not G_0). This process continues until one builds the series of forests $(\Gamma_0, \Gamma_1, \ldots, \Gamma_n = \Gamma)$ to correspond to $\mathcal{H}$.

As highlighted in Li et al. (2013), the duplication forest Γ (corresponding to G) is uniquely determined by $\mathcal{H}$ and a list of anchor nodes, $\pi = (u_1, \ldots, v_n)$. The important thing to emphasize here is that given the duplication forest Γ and G, one now only needs to infer the duplication nodes sequence $\theta(\mathcal{H}) = (v_1, v_2, \ldots, v_n)$ to reconstruct the complete growth history $\mathcal{H}$. For instance, at the first step backward, knowledge of $\Gamma_n = \Gamma$ together with v_n uniquely identifies the anchor node u_n, thus one can reconstruct G_{n-1} and Γ_{n-1}; this is then repeated for the remaining backward steps. Thus, given $\theta(\mathcal{H})$, one can construct the growth history $\mathcal{H}$ backward-in-time using the backward operators defined in section 2.4 of Li et al. (2013), which constructs G_{t-1}, Γ_{t-1} deterministically given (G_t, Γ_t, v_t), for $t = n, n-1, \ldots, 1$. An example of a growth history is given in Figure 1.

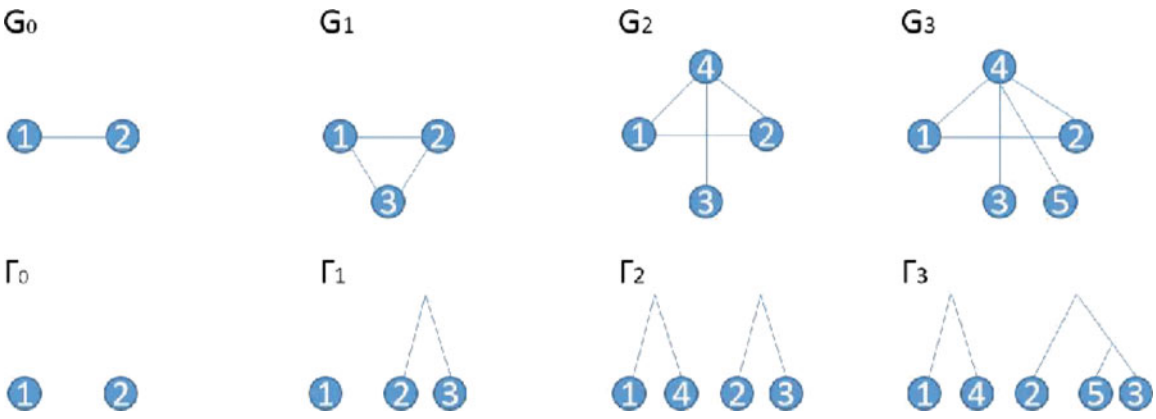


FIG. 1. An example growth history for a network together with the corresponding history of the duplication forest. In this example, $(u_1, u_2, u_3) = (2, 1, 3)$ and $(v_1, v_2, v_3) = (3, 4, 5)$.

2.4. Bayesian inference given the duplication history

Now suppose that in addition to G, a practitioner is given Γ corresponding to G. Our new objective—and the primary inference problem with which this work is concerned—is to consider the posterior density $\pi(\mathcal{M}|G, \Gamma)$. Notice that we have the joint distribution:

$$\pi(\mathcal{M}, \{G, \Gamma\}, \theta) = p(\mathcal{M}) p_{\mathcal{M}}(G_0^\theta, \Gamma_0^\theta) \prod_{t=1}^{n} p_{\mathcal{M}}(G_t^\theta, \Gamma_t^\theta \mid G_{t-1}^\theta, \Gamma_{t-1}^\theta)$$

where $\theta = (v_1, \ldots, v_n)$ is a sequence of duplication nodes compatible with the observed G, Γ, and $G_0^\theta, \Gamma_0^\theta, \ldots, G_n^\theta, \Gamma_n^\theta$ the corresponding reconstructed history. We are thus interested in the parameter posterior:

$$\begin{aligned} &\pi(\mathcal{M} \mid G, \Gamma) \propto p(\mathcal{M}) p_{\mathcal{M}}(G, \Gamma), \\ &p_{\mathcal{M}}(G, \Gamma) = \sum_{\theta | G, \Gamma} \left[p_{\mathcal{M}}(G_0^\theta, \Gamma_0^\theta) \prod_{t=1}^{n} p_{\mathcal{M}}(G_t^\theta, \Gamma_t^\theta | G_{t-1}^\theta, \Gamma_{t-1}^\theta) \right], \end{aligned} \tag{3}$$

The density $p_{\mathcal{M}}(G_0, \Gamma_0)$ is typically a trivial term that can be ignored in practice. As the duplication forest Γ limits the number of allowable anchor-and-duplicate node pairs, one can see that the number of possible growth histories is reduced.

2.5. Methods

We will now present an SMC algorithm that can sample the latent growth histories from the DMC model given the fixed parameters $\mathcal{M} := (p, p_c)$. We then show that this algorithm can be employed within a PMMH algorithm, as in Andrieu et al. (2010), to sample from the posterior (3) and infer $\mathcal{M}$ (and even $\theta|G, \Gamma$).

An SMC algorithm simulates a collection of N samples (or, particles) sequentially along the index t via importance sampling and resampling techniques to approximate a sequence of probability distributions of increasing state-space, which are known pointwise up to their normalizing constants. In this work, we use the SMC methodology to sample from the posterior distribution of the latent duplication history:

$$p_{\mathcal{M}}(\theta \mid G, \Gamma) \propto p_{\mathcal{M}}(G_0^\theta, \Gamma_0^\theta) \prod_{t=1}^{n} p_{\mathcal{M}}(G_t^\theta, \Gamma_t^\theta \mid G_{t-1}^\theta, \Gamma_{t-1}^\theta)$$

backward along the index t via Algorithm 1 in the Appendix. The technique provides an unbiased estimate of the normalizing constant (Theorem 7.4.2 of Del Moral, 2004), $p_{\mathcal{M}}(G, \Gamma)$:

$$\hat{p}_{\mathcal{M}}(G, \Gamma) = \prod_{t=0}^{n-1} \left[\frac{1}{N} \sum_{i=1}^{N} W_t^i \right], \tag{4}$$

where each W_t^i is an unnormalized importance weight computed in Algorithm 1. Note that under assumptions on the model, if $N > cn$ for some $c < \infty$, then the relative variance of the estimate is $\mathcal{O}(n/N)$ (see Cérou et al., 2011). It is remarked that, as in Wang et al. (2014), one could also use the discrete particle filter (Fearnhead, 1998), with a possible improvement over the SMC method detailed in Algorithm 1 (see Wang et al., 2014, for some details).

This SMC can be employed within a PMMH algorithm to target the posterior of $\mathcal{M}$ in (3). One can think of the deduced method as an MCMC algorithm running on the marginal $\mathcal{M}$-space, but with the SMC unbiased estimate $\hat{p}_{\mathcal{M}}(G, \Gamma)$ replacing the unknown likelihood $\hat{p}_{\mathcal{M}}(G, \Gamma)$. More analytically, we can consider all random variables involved in the method and write down the equilibrium distribution in the enlarged state space, with $\mathcal{M}$-marginal the target posterior $p_{\mathcal{M}}(G, \Gamma)$. Following Andrieu et al. (2010) and letting ϕ_t^i denote a sample (G_t, Γ_t) at time t, the extended equilibrium distribution is written as:

$$\pi^N\left(l, \mathcal{M}, a_{1:n-1}^{1:N}, \phi_{0:n-1}^{1:N} \mid G, \Gamma\right) = \frac{\pi\left(\mathcal{M}, \phi_{0:n-1}^{l} | G, \Gamma\right)}{N^n} \cdot \frac{\Psi_{\mathcal{M}}(a_{1:n-1}^{1:N}, \phi_{0:n-1}^{1:N})}{q_{\mathcal{M}}(v_n^{a_{n-1}^l}) \left(\prod_{t=1}^{n-1} w_t^{a_t^l} q_{\mathcal{M}}(v_t^{a_{t-1}^l}) \right)}, \tag{5}$$

where $\Psi_{\mathcal{M}}$ is the probability of all the variables associated to Algorithm 1, with a_k^l, $l \in \{1,..., N\}$ and the ϕ's being the simulated variables at each step of Algorithm 1.

A PMMH algorithm (see Algorithm 2) samples from (5), and one can remove the auxiliary variables from the samples to obtain draws for the parameters from (3). Furthermore, one could even save the sampled growth histories with particle index l to obtain draws from the joint posterior $\pi(\mathcal{M}, \theta \mid G, \Gamma)$. However, in this work, we are primarily interested in the inference of $\mathcal{M}$.

3. RESULTS

The variance of the estimate (4) plays a crucial role in the performance of Algorithm 2, as (4) is used to compute the acceptance probability within the PMMH algorithm. Thus, we first tested the variability of (4) as computed by Algorithm 1 to understand how the variance changes with $|V(G)|$. We then ran Algorithm 2 to sample from the posterior (3) and infer $\mathcal{M}$ for a given pair of observations (G, Γ). We present the details of those experiments below.

3.1. Variance of $\hat{p}_{\mathcal{M}}(G, \Gamma)$

We simulated a graph G and a forest Γ from the DMC model with the parameters set as ($p = 0.7$, $p_c = 0.7$), where $|V(G)| = 40$. We saved each pair (G_t, Γ_t) for $1 \le t \le 40$, and we ran Algorithm 1 50 times per pair (with $N = |V(G_t)| * 5$) to compute 50 unbiased estimates of $p_{\mathcal{M}}(G_t, \Gamma_t)$for $1 \le t \le 40$. In the top of Appendix Figure A1 in Appendix A, we plot the relative variance of the estimate (or, the variance divided by the square of the expected value) per each value of $|V(G_t)|$. We repeated the experiment two more times, with $N = |V(G_t)| * 10$ and $N = |V(G_t)| * 20$, and the associated output is also presented in Appendix Figure A1.

As remarked above, if $N > cn$ for some $c < \infty$, then the relative variance of $\hat{p}_{\mathcal{M}}(G_t, \Gamma_t)$ is $\mathcal{O}(n/N)$. Appendix Figure A1 confirms that the variance increases linearly, and that increasing the value of N with $|V(G)|$ (at least linearly) will help to control the variance. However, the plots show that the relative variance is still high, which means that N will have to be large to ensure satisfactory performance of the PMMH in practice.

3.2. Parameter inference

We separately simulated a graph G and a forest Γ from the DMC model with the parameters again set as ($p = 0.7$, $p_c = 0.7$), and we set $|V(G)| = 15$. Given only (G, Γ), we inferred (p, p_c) with each parameter having a uniform prior on the interval [0.1, 0.9]. We set the number of particles within Algorithm 1 to be $N = 2{,}000$, and we ran the PMMH algorithm to obtain 10,000 samples from the extended target.

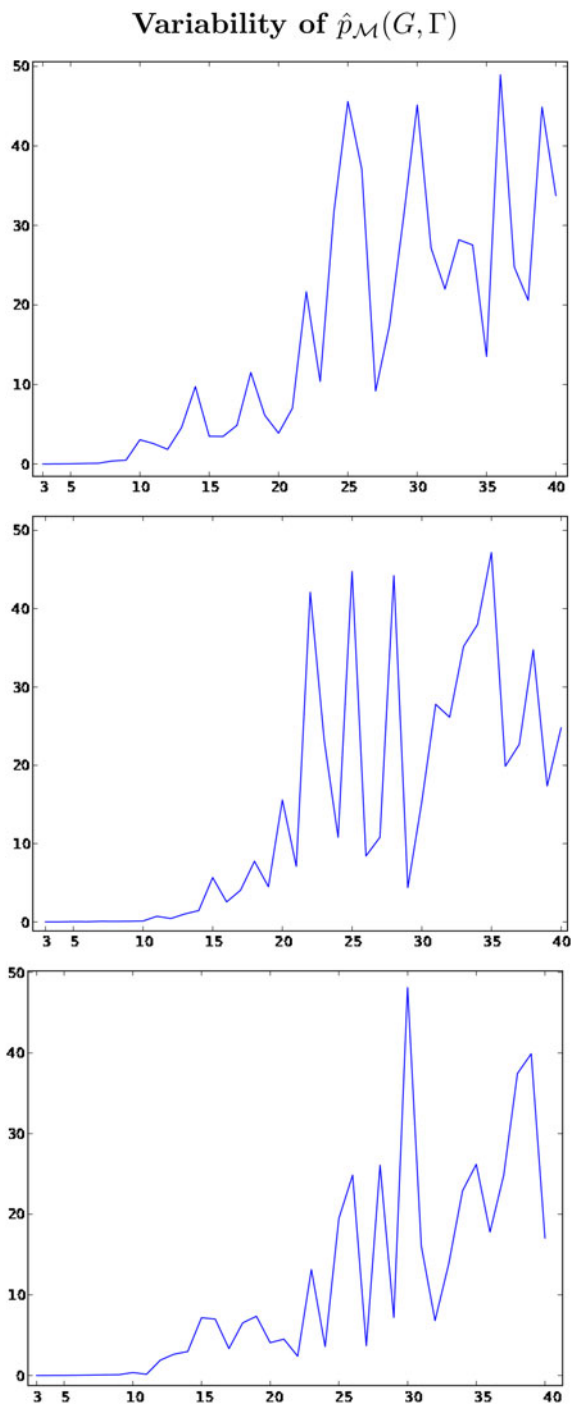
Appendix Figure A2 in Appendix A illustrates good mixing of the PMMH algorithm and accurate inference of the parameters (p, p_c). The trace plots show that the algorithm is not sticky, and the auto-correlation functions give evidence to an approximate independence between samples. The posterior densities are also interesting, in that they are clearly different from the uniform priors, and they show that the PMMH algorithm spends a majority of the computational time sampling the true parameter values.

4. DISCUSSION

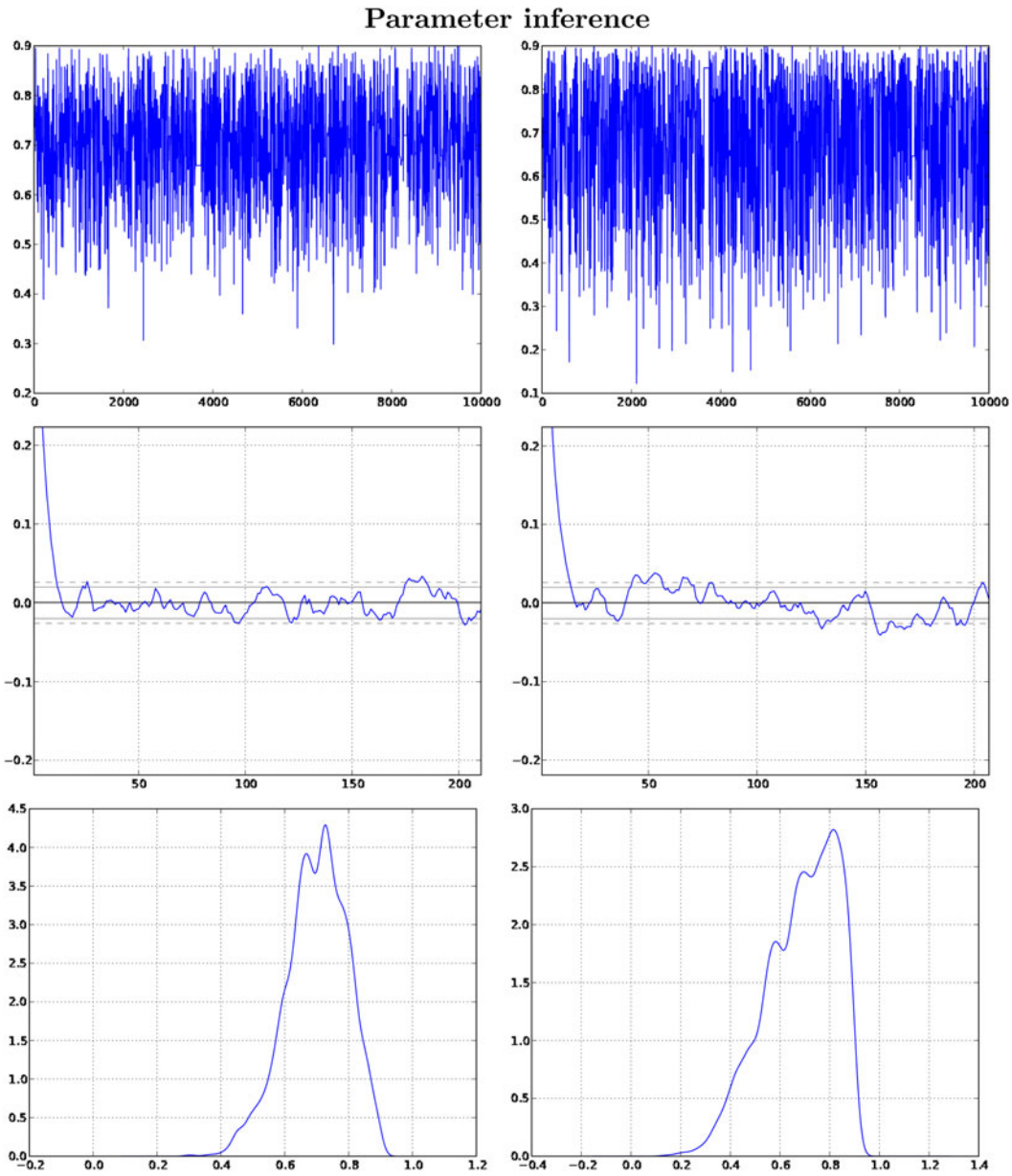
We have introduced a Bayesian inferential framework for the DMC model, where, as in Li et al. (2013), one assumes the pair (G, Γ) is observed and the parameters (p, p_c) are unknown. We then described how an SMC algorithm can be used to simulate growth histories and ultimately be employed within PMMH to target the posterior distribution of the parameters (3), thereby opening up the possibility of performing Bayesian inference on the DMC model.

Numerical tests demonstrated that Algorithm 1 can have a high variability when $|V(G)|$ is large and N is not sufficiently high, and this limits the scope of the inference problem, which can be tackled using the complete Algorithm 2. However, the proposals used in the example experiments within Algorithms 1 and 2 are naive, as the method simply chooses a candidate duplicate node v_t^i at random from all permitted nodes given the current G_t^i, Γ_t^i. It is reasonable to assume that more sophisticated proposal densities could reduce the variance of the SMC and/or improve the mixing of the PMMH, thereby allowing one to perform inference when $|V(G_t)|$ is large and N is smaller. This could be explored in a future work.

5. APPENDIX A



APPENDIX FIG. A1. All plots illustrate the relative variance of $\hat{p}_{\mathcal{M}}(G_t, \Gamma_t)$, per $|V(G_t)|$ on the horizontal axis; the relative variance is the variance divided by the square of the expected value. In the top plot, the number of SMC particles used to compute each $\hat{p}_{\mathcal{M}}(G_t, \Gamma_t)$ is $|V(G_t)| * 5$. In the middle and bottom plots, that number is $|V(G_t)| * 10$ and $|V(G_t)| * 20$, respectively. Recall that the seed graph, G_0, has two nodes, and note that we did not compute $\hat{p}_{\mathcal{M}}(G_0, \Gamma_0)$ because it is trivial.



APPENDIX FIG. A2. Plots associated with p and p_c are at left and right respectively. The top figures are trace plots, with PMMH iteration running along the horizontal axes and parameter value running along the verticals. The middle figures are plots of the autocorrelation functions (with lag running along the horizontal axes), and at the bottom we present the parameter posterior densities.

6. APPENDIX B: ALGORITHM SUMMARIES

Algorithm 1 Sequential Monte Carlo (SMC)

- Step 0: Input an observed graph $G = G_n$ and a corresponding observed forest $\Gamma = \Gamma_n$, where G is not a seed graph.
- Step 1: Set $t = n$. For $i \in \{1, \ldots, N\}$, sample a subtree with two nodes uniformly at random from Γ_t^i, and choose one of the two nodes uniformly as the proposed duplicate node v_t^i (thus the other will be the anchor node). Using the backward operators defined in [10, section 2.4], construct each $(G_{t-1}^i, \Gamma_{t-1}^i)$ from the subtrees and (G_t^i, Γ_t^i). For $i \in \{1, \ldots, N\}$, compute the unnormalized weight

$$W_{t-1}^i = \frac{p_{\mathcal{M}}(G_t^i, \Gamma_t^i \mid G_{t-1}^i, \Gamma_{t-1}^i)}{q_{\mathcal{M}}(v_t^i)},$$

  where $q\mathcal{M}$ is the density of the proposal mechanism used to sample $\{u_t^i\}$.
- Step 2: If $\{G_{t-1}^{1:N}\}$ are not seed graphs, then set $t = t - 1$ and continue to Step 3. Otherwise, the algorithm terminates.
- Step 3: For $i \in \{1, \ldots, N\}$, sample $a_t^i \in \{1, \ldots, N\}$ from a discrete distribution on $\{1, \ldots, N\}$ with j^{th} probability $w_t^j \propto W_t^j$. The sample $\{a_t^{1:N}\}$ are the indices of the resampled particles. Set all normalized weights equal to N^{-1}.
- Step 4: For $i \in \{1, \ldots, N\}$, sample a subtree with two nodes uniformly at random from the resampled forest $\Gamma_t^{a_t^i}$, and select uniformly one of the two nodes as the proposed duplicate node v_t^i. Construct $(G_{t-1}^i, \Gamma_{t-1}^i)$ from v_t^i, $(G_t^{a_t^i}, \Gamma_t^{a_t^i})$. For $i \in \{1, \ldots, N\}$, compute the unnormalized weight

$$W_{t-1}^i = \frac{p_{\mathcal{M}}(G_t^{a_t^i}, \Gamma_t^{a_t^i} \mid G_{t-1}^i, \Gamma_{t-1}^i)}{q_{\mathcal{M}}(v_t^i)}.$$

  Return to Step 2.

Algorithm 2 Particle Marginal Metropolis—Hastings (PMMH)

- Step 0: Set $r = 0$. Sample $\mathcal{M}^{(r)} \sim p(\cdot)$. All remaining random variables can be sampled from their full conditionals defined by the target (5):
  - Sample $\phi_{0:n-1}^{(r),1:N}, a_{1:n-1}^{(r),1:N} \sim \Psi_{\mathcal{M}^{(r)}}(\cdot)$ via Algorithm 1.
  - Choose a particle index $l^{(r)} \propto W_0^{(r),l^{(r)}}$.

  Finally, calculate $\hat{p}_{\mathcal{M}^{(r)}}(G, \Gamma)$ via (4).
- Step 1: Set $r = r + 1$. Sample $\mathcal{M}^* \sim q(\cdot \mid \mathcal{M})$. All remaining random variables can be sampled from their full conditionals defined by the target (5):
  - Sample $\phi_{0:n-1}^{*,1:N}, a_{1:n-1}^{*,1:N} \sim \Psi_{\mathcal{M}^*}(\cdot)$ via Algorithm 1.
  - Choose a particle index $l^* \propto W_0^{*,l^*}$.

  Finally, calculate $\hat{p}_{\mathcal{M}^*}(G, \Gamma)$ via (4).
- Step 2: With acceptance probability

$$1 \wedge \frac{\hat{p}_{\mathcal{M}^*}(G, \Gamma) q(\mathcal{M} \mid \mathcal{M}^*)}{\hat{p}_{\mathcal{M}^{(r-1)}}(G, \Gamma) q(\mathcal{M}^* \mid \mathcal{M})},$$

  set $(l^{(r)}, \mathcal{M}^{(r)}, \phi_{0:n-1}^{(r),1:N}, a_{1:n-1}^{(r),1:N}) = (l^*, \mathcal{M}^*, \phi_{0:n-1}^{*,1:N}, a_{1:n-1}^{*,1:N})$. Otherwise, set

  $(l^{(r)}, \mathcal{M}^{(r)}, \phi_{0:n-1}^{(r),1:N}, a_{1:n-1}^{(r),1:N}) = (l^{(r-1)}, \mathcal{M}^{(r-1)}, \phi_{0:n-1}^{(r-1),1:N}, a_{1:n-1}^{(r-1),1:N})$.

  Return to the beginning of Step 1.

ACKNOWLEDGMENTS

This research was funded by the EPSRC grant ''Advanced Stochastic Computation for Inference from Tree, Graph and Network Models'' (Ref: EP/K01501X/1). A.J. was additionally supported by a Singapore

Ministry of Education Academic Research Fund Tier 1 grant (R-155-000-156-112) and is also affiliated with the Risk Management Institute and the Centre for Quantitative Finance at the National University of Singapore.

AUTHOR DISCLOSURE STATEMENT

The authors declare that no competing financial interests exist.

REFERENCES

Andrieu, C., Doucet, A., and Holenstein, R. 2010. Particle Markov chain Monte Carlo methods. *J. R. Stat. Soc. B*. 72, 269–342.

Cérou, F., Del Moral, P., and Guyader, A. 2011. A non-asymptotic variance theorem for unnormalized Feynman-Kac particle models. *Ann. Inst. Henri Poincare*. 47, 629–649.

Del Moral, P. 2004. *Feynman-Kac Formulae: Genealogical and Interacting Particle Systems with Applications*. Springer, New York.

Doucet, A., Godsill, S., and Andrieu, C. 2000. On sequential Monte Carlo sampling methods for Bayesian filtering. *Stat. Comput.* 10, 197–208.

Dutkowski, J., and Tiuryn, J. 2007. Identification of functional modules from conserved ancestral protein–protein interactions. *Bioinformatics*. 23, i149–i158.

Fearnhead, P. 1998. Sequential Monte Carlo methods in filter theory [D.Phil. thesis]. University of Oxford, Oxford.

Gibson, T.A., and Goldberg, D.S. 2009. Reverse engineering the evolution of protein interaction networks. *Pac. Symp. Biocomput.* 14, 190–202.

Gordon, N.J., Salmond, D.J., and Smith, A.F.M. 1993. Novel approach to nonlinear/non-Gaussian Bayesian state estimation. *IEE Proc. F.* 140, 107–113.

Kreimer, A., Borenstein, E., Gophna, U., and Ruppin, E. 2008. The evolution of modularity in bacterial metabolic networks. *Proc. Natl. Acad. Sci. USA*. 105, 6976–6981.

Li, S., Choi, K.P., Wu, T., and Zhang, L. 2013. Maximum likelihood inference of the evolutionary history of a PPI network from the duplication history of its proteins. *IEEE/ACM Trans. Comput. Biol. Bioinform.* 10, 1412–1421.

Patro, R., Sefer, E., Malin, J., et al. 2012. Parsimonious reconstruction of network evolution. Algorithms Mol. Biol. 7, 25.

Pereira-Leal, J.B., Levy, E.D., Kamp, C., and Teichmann, S.A. 2007. Evolution of protein complexes by duplication of homomeric interactions. *Genome Biol.* 8, R51.

Pinney, J., Amoutzias, G., Rattray, M., and Robertson, D. 2007. Reconstruction of ancestral protein interaction networks for the bZIP transcription factors. *Proc. Natl. Acad. Sci. USA*. 104, 20449–20453.

Vazquez, A., Flammini, A., Maritan, A., and Vespignani, A. 2003. Modeling of protein interaction networks. *ComplexUs*. 1, 38–44.

Wagner, A. 2001. The yeast protein interaction network evolves rapidly and contains few redundent duplicate genes. *Mol. Biol. Evol.* 18, 1283–1292.

Wang, J., Jasra, A., and De Iorio, M. 2014. Computational methods for a class of network models. *J. Comp. Biol*. 21, 141–161.

Address correspondence to:
Dr. Ajay Jasra
Department of Statistics & Applied Probability
National University of Singapore
6 Science Drive 2
Singapore 117546
Singapore

E-mail: staja@nus.edu.sg