Journal of the American Statistical Association

ISSN: 0162-1459 (Print) 1537-274X (Online) Journal homepage: https://www.tandfonline.com/loi/uasa20

Likelihood-Based Inference for Partially Observed Epidemics on Dynamic Networks

Fan Bu, Allison E. Aiello, Jason Xu & Alexander Volfovsky

To cite this article: Fan Bu, Allison E. Aiello, Jason Xu & Alexander Volfovsky (2020): Likelihood-Based Inference for Partially Observed Epidemics on Dynamic Networks, Journal of the American Statistical Association, DOI: 10.1080/01621459.2020.1790376

To link to this article: https://doi.org/10.1080/01621459.2020.1790376

JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION
2020, VOL. 00, NO. 0, 1–17: Theory and Methods
https://doi.org/10.1080/01621459.2020.1790376

Likelihood-Based Inference for Partially Observed Epidemics on Dynamic Networks

Fan Bu[a], Allison E. Aiello[b], Jason Xu*[a], and Alexander Volfovsky*[a]

[a]Department of Statistical Science, Duke University, Durham, NC; [b]Department of Epidemiology, Gillings School of Global Public Health, University of North Carolina at Chapel Hill, Chapel Hill, NC

ABSTRACT

We propose a generative model and an inference scheme for epidemic processes on dynamic, adaptive contact networks. Network evolution is formulated as a link-Markovian process, which is then coupled to an individual-level stochastic susceptible-infectious-recovered model, to describe the interplay between the dynamics of the disease spread and the contact network underlying the epidemic. A Markov chain Monte Carlo framework is developed for likelihood-based inference from partial epidemic observations, with a novel data augmentation algorithm specifically designed to deal with missing individual recovery times under the dynamic network setting. Through a series of simulation experiments, we demonstrate the validity and flexibility of the model as well as the efficacy and efficiency of the data augmentation inference scheme. The model is also applied to a recent real-world dataset on influenza-like-illness transmission with high-resolution social contact tracking records. Supplementary materials for this article are available online.

ARTICLE HISTORY
Received October 2019
Accepted June 2020

KEYWORDS
Bayesian data augmentation; Conditional simulation; Contact networks; Continuous-time Markov chains; Mobile healthcare; Stochastic susceptible-infectious-removed model

1. Introduction

The vast majority of epidemiological models, such as the well-known susceptible-infectious-recovered (SIR) model, rely on compartmentalizing individuals according to their disease status (Kermack and McKendrick 1927). Classically, such models describe population-level behavior under a "random mixing" assumption that an infectious individual can spread the disease homogeneously to any susceptible individual (Kermack and McKendrick 1927; Bailey 1975; Anderson and May 1992). In the last two decades, an alternative assumption—that the disease is transmitted through links in a contact network—has gradually gained popularity. It has been found that the contact network structure can fundamentally impact the behavior of epidemic processes (Edmunds, O'Callaghan, and Nokes 1997; Wallinga, Edmunds, and Kretzschmar 1999; Edmunds et al. 2006; Mossong et al. 2008; Volz and Meyers 2008; Melegaro et al. 2011; Cui, Zhang, and Feng 2019); on the other hand, the network structure can in turn be influenced by disease status of individuals as well (Bell et al. 2006; Eames et al. 2010; Funk, Salathé, and Jansen 2010; Van Kerckhove et al. 2013).

This growing interest—alongside technological advances in mobile data—has spurred efforts on collecting high-resolution data that inform the dynamics of the contact network (Vanhems et al. 2013; Barrat et al. 2014; Voirin et al. 2015; Aiello et al. 2016; Kiti et al. 2016; Ozella et al. 2018). Data of this type has most recently been collected by Ministry of Health, State of Israel (2020) and Korea Centers for Disease Control and Prevention (2020). However, there is a gap between the demand to analyze such emergent data and available methods: a recent review by Britton (2020) outlines possible approaches to inference for epidemic models on networks and calls to attention considerable challenges—in particular, key terms such as transition probabilities that appear in central quantities such as likelihood expressions are unavailable. Accounting for the relationship between disease spread and the underlying contact network during inference is crucial to accurately estimating parameters describing the inherent properties of the disease, and moreover has direct practical implications on epidemic control and intervention. Policies such as quarantine or suppression are naturally described as changes to the contact network, and yet modeling approaches are largely restricted to prospective simulations and/or analysis based on static networks in lieu of direct inference from modern data. Recent examples can be seen in analyses for COVID-19 (Ferguson et al. 2020) and for MERS-CoV transmissions (Yang and Jung 2020).

To address this methodological gap, we develop a statistical model that describes the mutual interplay between SIR-type epidemics and an underlying dynamic network, together with a tractable inferential framework to fit such models to modern time-resolved datasets. In particular, we propose a stochastic generative model that can be fit to data using likelihood-based methods. We present a Bayesian data augmentation scheme that accommodates partial observations such as missing recovery times that are common in real data while quantifying uncertainty in estimated parameters.

The majority of existing work on epidemic processes over networks adopt a deterministic approach based on ordinary differential equation (ODE) models (WHOW Group 2006; Volz and Meyers 2007; Shaw and Schwartz 2008; Volz 2008; Van

CONTACT Jason Xu *jason.q.xu@duke.edu* Department of Statistical Science, Duke University, Durham, NC 27708.
*Joint last authors.
Color versions of one or more of the figures in the article can be found online at *www.tandfonline.com/r/JASA*.
Supplementary materials for this article are available online. Please go to *www.tandfonline.com/r/JASA*.

Segbroeck, Santos, and Pacheco 2010; Kiss et al. 2012; Tunc, Shkarayev, and Shaw 2013; Ogura and Preciado 2017). Such approaches do not provide a measure of uncertainty, and do not offer probabilistic interpretations. Our framework builds upon previous stochastic (and likelihood-based) methods that either do not consider network dynamics (Britton and O'Neill 2002; Dong, Pentland, and Heller 2012; Fan et al. 2015; Fan, Li, and Heller 2016), or do not model the contact network at all (Cauchemez et al. 2006; Hoti et al. 2009; Britton 2010; Fintzi et al. 2017; Ho et al. 2018). Moreover, the proposed inference scheme accommodates partially observed data, with a focus on unknown recovery times in this article. Handling missing data even *without* network constraints is already challenging, and often requires simplifying model assumptions (Finkenstädt and Grenfell 2000; Cauchemez and Ferguson 2008) or computationally intensive simulation-based inference (Andrieu, Doucet, and Holenstein 2010; He, Ionides, and King 2010; Ionides et al. 2015; Pooley, Bishop, and Marion 2015).

The rest of the article is organized as follows: Section 2 reviews background on epidemic models (focusing on the stochastic SIR model) and dynamic network processes. Section 3 formulates the generative model and derives maximum likelihood estimators as well as Bayesian posterior distributions based on the complete data likelihood. Section 4 describes a Bayesian inference scheme that deals with incomplete observations on individual recovery times. Sections 5 and 6 present experiment results on simulated datasets and a real-world dataset. Finally Section 7 provides further discussions.

2. Background

2.1. Compartmental Epidemiological Models

Compartmental models divide individuals into non-overlapping subsets according to their disease statuses. In classical models, the changes in these subpopulations over time are described by ODEs (Hethcote 2000). One widely used model is the SIR model, which assumes three disease statuses—susceptible (S), infectious (I), and recovered (or removed, R). On a closed population of N individuals (with N sufficiently large), the dynamics of the deterministic SIR model can be expressed as

$$\frac{dS(t)}{dt} = -\beta S(t)I(t), \quad \frac{dI(t)}{dt} = \beta S(t)I(t) - \gamma I(t), \quad (1)$$

where $S(t)$ and $I(t)$ refer to the number of susceptible and infectious individuals at time t, respectively, and the number of recovered individuals satisfies $R(t) = N - [S(t) + I(t)]$. The parameters describe the epidemic mechanistically: here β is interpreted as the rate of disease transmission *per contact between* an S individual and an I individual, and γ is the rate of recovery for an I individual.

By setting the growth rate of infection to be proportional to $S(t)I(t)$, the model in (1) implicitly assumes that any two members can interact with each other. This assumption is easily violated in reality, where an individual only maintains contact with a limited number of others. Moreover, the differential equations can only account for the average, expected behavior of the process, but the transmission of an infectious disease exhibits randomness and uncertainty by nature.

To account for the underlying network structure of a population as well as the random nature of an epidemic process, we adopt an *individual-level*, *stochastic* variation of the SIR model, similar to that used in Auranen et al. (2000). An individual of status S (susceptible) at time t (> 0) changes disease status to I (infectious) at time $t + h$ (h is an infinitesimal quantity) with a probability that is dependent on both the infection rate β *and* his/her contacts at time t. An infectious individual at time t becomes a member of the R (recovered) subpopulation at time $t + h$ with a probability determined by the recovery rate γ. Specifically, for any susceptible individual p_1 and infectious individual p_2 in the population at t, conditioned on the current overall state of the process, $\mathcal{Z}_t$, then

$$\Pr(p_1 \text{ gets infected by } p_2 \text{ at } t + h \mid \mathcal{Z}_t) = \beta h + o(h) \quad (2)$$

if p_1 and p_2 are in contact at t, and

$$\Pr(p_2 \text{ recovers at } t + h \mid \mathcal{Z}_t) = \gamma h + o(h). \quad (3)$$

2.2. Basic Network Concepts

A network, or a graph, is a two-component set, $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$, where $\mathcal{V}$ is the set of N nodes and $\mathcal{E}$ is the set of links. A network can be represented by its "adjacency matrix," $\mathbf{A}$, where $\mathbf{A}_{ij} = 1$ indicates there is a link from node i to j. Since most infectious diseases can be transmitted in both directions through a contact, we assume that the adjacency matrix is *symmetric*, $\mathbf{A}_{ij} = \mathbf{A}_{ji}$.

A special network structure is the fully connected network (or the complete graph), $\mathcal{K}_N$, and its adjacency matrix $\mathbf{A}$ satisfies $\mathbf{A}_{ij} = 1$ for any $i \neq j$. This network structure corresponds to the widely adopted "random mixing" assumption in epidemiological models, which, as stated before, may be restrictive and unrealistic. Therefore, in the rest of the article, we instead consider *arbitrary* network structures underlying the population.

2.3. Temporal and Adaptive Networks

Interactions between individuals are dynamic in nature, and such dynamics are important when modeling epidemic processes (Masuda and Holme 2017; also as demonstrated later in Section 5.1). We consider a continuous-time link-Markovian model for temporal networks (Clementi et al. 2010; Ogura and Preciado 2016). Following the symmetric network assumption stated above, two individuals i and j ($i < j$) who are not in contact at time t form a link at time $t + h$ ($h \ll 1$) with probability αh, where α is the link activation rate. Similarly, if there is an edge between i and j at time t, then the edge is deleted at time $t + h$ with probability ωh, where ω is some link termination rate.

If, instead, individuals establish and dissolve their social links with rates that vary according to their disease statuses, then the evolution of the network is coupled to the epidemic process and thus becomes *adaptive*. This mechanism can be described via instantaneous rates of single-link changes. For any two individuals i and j, their corresponding entry in the adjacency matrix is modeled as a $\{0, 1\}$-valued Markov process, $\mathbf{A}_{ij}(t), t > 0$. Suppose that at time t, i is of status A, j is of status B,[1] then for

[1] Here $A, B \in \{S, I, R\}$, and we only consider $i < j$ since the network is assumed symmetric.

an infinitesimal quantity h,

$$Pr(\mathbf{A}_{ij}(t+h)=\mathbf{A}_{ji}(t+h)=1 \mid \mathbf{A}_{ij}(t)=0)=\alpha_{AB}h+o(h); \tag{4}$$

$$Pr(\mathbf{A}_{ij}(t+h)=\mathbf{A}_{ji}(t+h)=0 \mid \mathbf{A}_{ij}(t)=1)=\omega_{AB}h+o(h). \tag{5}$$

Here $\alpha_{AB}(=\alpha_{BA})$ is the activation rate for an A-B type link, and similarly, $\omega_{AB}(=\omega_{BA})$ is the termination rate for an A-B type link.

3. Epidemic Processes Over Adaptive Networks

3.1. The Generative Model

In this subsection, we lay out a stochastic data generative process (referred to as the "generative model") for the joint evolution of an individualized SIR process on a networked population and the dynamics of the contact network. In contrast to the ODE literature and existing network models described in Section 1, the key feature of the model is the *interplay* between epidemic progression and network adaptation. On one hand, transmission of infection depends on the existence of susceptible-infectious links, which may change through time; on the other hand, network links temporally update in a manner that in turn depends on individual disease status.

We formulate this complex process as a continuous-time Markov chain that comprises all individual-level events described in Section 2. The joint evolution of the individual Poisson processes described by (2)–(5) can be described via a competing risks construction. By the Markov property, the time until each type of event has an exponential waiting time, and thus the time to next event in the joint process remains exponentially distributed (Guttorp and Minin 2018). Events occur stochastically and are of one of the four types:

- *Infection:* The disease is transmitted through a link between an S (susceptible) and an I (infectious) individual (S-I link) with rate β;
- *Recovery:* Each I individual recovers with rate γ independently;
- *Link activation:* A link is formed at rate $\alpha_{AB}(=\alpha_{BA})$ between an individual of status A and another of status B who are not connected, where $A, B \in \{S, I, R\}$;
- *Link termination:* An existing link is removed at rate $\omega_{AB}(=\omega_{BA})$ between an individual of status A and another of status B, where $A, B \in \{S, I, R\}$.

This formulation will allow for joint inference of both disease spread and network evolution. As illustrated in the next subsection, inference is straightforward when all the events are fully observed. Furthermore, this formulation implies a relatively simple generative process at the population level. Conditioned on the current state of the process $\mathcal{Z}_t$ at time $t\,(>0)$, the very next event of the entire process is the *earliest* event that occurs among the four competing processes by the superposition property:

- *Infection:* An infection occurs with rate $\beta SI(t)$, where $SI(t)$ is the number of S-I links at time t;
- *Recovery:* A recovery occurs with rate $\gamma I(t)$, where $I(t)$ is the number of infectious individuals at time t;
- *Link activation:* An A-B link is established with rate $\alpha_{AB}M^d_{AB}(t)$, where $M^d_{AB}(t)$ is the number of disconnected A-B pairs at time t;
- *Link termination:* An A-B link is dissolved with rate $\omega_{AB}M_{AB}(t)$, where $M_{AB}(t)$ is the number of connected A-B pairs at time t.

We may interpret this generative model as a generalization of two simpler models. If we set $\alpha_{AB} \equiv \alpha$ and $\omega_{AB} \equiv \omega$ for any status A and B, the coupled process reduces to a *decoupled* process, where network evolution is *independent* of individual disease status. Moreover, if we fix $\alpha \equiv \omega \equiv 0$, the process is further reduces to an SIR process over a *static* network.

Here we assume that the population size N is fixed, and that at $t = 0$, the initial network $\mathcal{G}_0$ as well as $I(0)$ initial infection cases are observed. We summarize a list of model parameters and notation in Table 1.

Table 1. Table of parameters and notation.

Parameter	Description
β	Infection rate
γ	Recovery rate
α	Link activation rate for a currently disconnected pair
ω	Link termination rate for a currently connected pair
α_{AB}	Link activation rate for a currently disconnected A-B pair
ω_{AB}	Link termination rate for a currently connected A-B pair
Notation	**Description**
N	Total population size (assumed to remain fixed throughout the process)
$T_{\max}$	Maximum observation time
$\mathcal{Z}_t$	State of the process at time t (including the epidemic status of every individual and the social network structure at time t)
$\mathcal{G}_t$	Social network structure (a graph) at time t
$S(t), I(t)$	Numbers of susceptible/infectious individuals in the population at time t
$H(t)$	Number of healthy (not infectious) individuals in the population at time t
$I_k(t)$	Number of infectious individuals in person k's neighborhood at time t
$SI(t)$	Number of S-I links in the network at time t
$M(t)$	Total number of edges in the network at time t
$M_{AB}(t)$	Number of A-B links at time t
$M^d_{AB}(t)$	Number of disconnected A-B pairs at time t
n_E, n_R	Counts of infection events and recovery events in the process
n_N	Count of network events in the process (each event is the activation or termination of a single link)
C, D	Total counts of link activation/termination
C_{AB}, D_{AB}	Counts of link activation/termination events for A-B pairs

3.2. Complete Data Likelihood and Parameter Inference

3.2.1. Derivation of Complete Data Likelihood

Suppose n_E infection events and n_R recovery events are observed in total. Let i_k be the infection time for individual k ($k = 1, 2, \ldots, n_E$), r_k be k's recovery time (if $r_k > T_{\max}$, k's recovery is not observed), and without loss of generality, set $i_1 = 0$. Recall that the widely used "random mixing" assumption in classical epidemiological models is equivalent to assuming that the contact network is a complete graph, $\mathcal{K}_N$, and the individual-

based complete data likelihood under this assumption is[2]

$$\begin{aligned}\mathcal{L}(\beta,\gamma) &= p(\text{epidemic events}|\beta,\gamma)\\ &= \gamma^{n_R}\prod_{k=2}^{n_E}[\beta I(i_k)]\\ &\quad\times\exp\left(-\int_0^{T_{\max}}[\beta S(u)I(u)+\gamma I(u)]\,du\right).\end{aligned}$$

To account for the contact network, let $\mathcal{G}_t$ be an arbitrary network, and begin by assuming that the *entire network process* $\{\mathcal{G}_t : 0 < t < T_{\max}\}$ is fully observed. Explicitly accounting for the number of infectious contacts per individual at the time of infection as well as the total number of *S*-*I* links in the system, the complete data likelihood becomes

$$\begin{aligned}\mathcal{L}(\beta,\gamma|\mathcal{G}) &= \gamma^{n_R}\prod_{k=2}^{n_E}[\beta I_k(i_k)]\\ &\quad\times\exp\left(-\int_0^{T_{\max}}[\beta SI(u)+\gamma I(u)]\,du\right).\end{aligned}\tag{6}$$

Here $I_k(i_k)$ denotes the number of *infectious neighbors of* person k at his time of infection i_k and $SI(t)$ denotes the number of *S*-*I* links in the system at time t. We see that the dynamic nature of the network is implicitly subsumed into the terms $I_k(i_k)$'s and $SI(u)$. To clarify this point, note that the same likelihood holds for a *static* network $\mathcal{G}$. As neighborhoods $I_k(i_k)$ are fixed in the static case, one could further simplify (6) using a constant $I_k(i_k) = I_k$ for all times i_k.

Equation (6) serves as a point of departure toward network dynamics. As a stepping stone, we first consider the simpler *decoupled* case in which the network and epidemic evolve *independently*. Here the edge activation rate α and deletion rate ω, as well as total number of activated and terminated edges denoted C and D, do not depend on disease status. Given an initial network $\mathcal{G}_0$, the network process likelihood can be easily written as

$$\begin{aligned}&\mathcal{L}(\alpha,\omega|\mathcal{G}_0) = p(\text{network events}|\alpha,\omega,\mathcal{G}_0)\\ =&\alpha^C\omega^D\prod_{\ell=1}^{n_N}\left[\left(\frac{N(N-1)}{2}-M(s_\ell)\right)^{1-D_\ell}M(s_\ell)^{D_\ell}\right]\\ &\times\exp\left(-\alpha\frac{N(N-1)}{2}T_{\max}+(\alpha-\omega)\int_0^{T_{\max}}M(u)du\right).\end{aligned}\tag{7}$$

Here s_ℓ is the time of the ℓth network event, and $D_\ell = 1$ if this event is a *link termination* and otherwise $D_\ell = 0$. By independence, the complete data likelihood in this decoupled case is simply a product of Equations (6) and (7):

$$\begin{aligned}&\mathcal{L}(\beta,\gamma,\alpha,\omega|\mathcal{G}_0) = p(\text{epidemic events}|\beta,\gamma,\mathcal{G}_t)\\ &\quad\cdot p(\text{network events}|\alpha,\omega,\mathcal{G}_0)\\ &= \beta^{n_E-1}\gamma^{n_R}\alpha^C\omega^D\\ &\quad\times\prod_{k=2}^{n_E}[I_k(i_k)]\prod_{\ell=1}^{n_N}\left[\left(\frac{N(N-1)}{2}-M(s_\ell)\right)^{1-D_\ell}M(s_\ell)^{D_\ell}\right]\\ &\quad\times\exp\left(-\int_0^{T_{\max}}[\beta SI(u)+\gamma I(u)+(\omega-\alpha)M(u)]\,du\right.\\ &\quad\left.-\alpha\frac{N(N-1)}{2}T_{\max}\right).\end{aligned}\tag{8}$$

Finally, we allow link activation and termination to be dependent on individual disease status, yielding an *adaptive* network. We introduce some notation; it is natural to assume that the *S* and *R* populations behave identically from the perspective of the network process:

$$\alpha_{RA}\equiv\alpha_{SA},\ \text{and}\ \omega_{RA}\equiv\omega_{SA},\ \forall A\in\{S,I,R\}.$$

Let $H(t) = R(t)+S(t) = N-I(t)$ denote the number of such "healthy" individuals at time t. Naturally the term "*H*-*H* link" represent an *S*-*S* link, an *S*-*R* link, or an *R*-*R* link, and the term "*H*-*I* link" represents an *S*-*I* link or *R*-*I* link. We also define $g(p,t)$ as the indicator function of infectiousness, that is, $g(p,t) = 1$ if person p is infected at time t and $g(p,t) = 0$ otherwise.

Denote the ordered epidemic and network events together as $\{e_j = (t_j, p_{j1}, p_{j2})\}_{j=1}^n$, with $n = n_E+n_R+n_N$. Here t_j ($j = 1,2,\ldots,n$) denote the event times and $t_1 = 0$ is the infection time of the first patient. If e_j is a network event, p_{j1} and p_{j2} are the two individuals getting connected or disconnected, and if e_j is an epidemic event, let p_{j1} be the person getting infected or recovered and set $p_{j2} = 0$. Furthermore let event type indicators F_j, C_j, D_j take the value 1 only if e_j is an infection, a link activation, and a link deletion, respectively, and 0 otherwise.

The contribution of all network events to the complete data likelihood is in essence of the same form as (7), except that for every activation or termination event the link type has to be considered. Then the likelihood component of the adaptive network process is

$$\begin{aligned}&\alpha_{SS}^{C_{HH}}\alpha_{SI}^{C_{HI}}\alpha_{II}^{C_{II}}\omega_{SS}^{D_{HH}}\omega_{SI}^{D_{HI}}\omega_{II}^{D_{II}}\prod_{j=2}^{n}\tilde{M}(t_j)\\ &\quad\times\exp\left(-\int_0^{T_{\max}}\left[\tilde{\alpha}^T\mathbf{M}_{\max}(t)+(\tilde{\omega}-\tilde{\alpha})^T\mathbf{M}(t)\right]dt\right),\end{aligned}$$

where

$$\begin{aligned}\tilde{M}(t_j) =&\left[(\alpha_{SS}M_{HH}^d(t_j))^{C_j}(\omega_{SS}M_{HH}(t_j)^{D_j}\right]^{(1-g(p_{j1},t_j))(1-g(p_{j2},t_j))}\\ &\times\left[(\alpha_{SI}M_{HI}^d(t_j))^{C_j}(\omega_{SI}M_{HI}(t_j))^{D_j}\right]^{|g(p_{j1},t_j)-g(p_{j2},t_j)|}\\ &\times\left[(\alpha_{II}M_{II}^d(t_j))^{C_j}(\omega_{II}M_{II}(t_j))^{D_j}\right]^{g(p_{j1},t_j)g(p_{j2},t_j)},\end{aligned}\tag{9}$$

$$\tilde{\alpha} = (\alpha_{SS},\alpha_{SI},\alpha_{II})^T,\tag{10}$$

$$\tilde{\omega} = (\omega_{SS},\omega_{SI},\omega_{II})^T,\tag{11}$$

$$\mathbf{M}_{\max}(t) = \left(\frac{H(t)(H(t)-1)}{2}, H(t)I(t), \frac{I(t)(I(t)-1)}{2}\right)^T,\tag{12}$$

$$\mathbf{M}(t) = (M_{HH}(t_j), M_{HI}(t), M_{II}(t))^T.\tag{13}$$

[2]Note that this expression differs from the *population-level* complete likelihood (Becker and Britton 1999). When epidemic events are tied to individuals, that is, a recovery time is associated to a specific infection event, we must use the individual-based likelihood instead (similar to that in Auranen et al. (2000)).

Therefore, given the initial network structure $\mathcal{G}_0$ and one infectious case at time 0, the complete data likelihood of the coupled process can be expressed as[3]

$$\begin{aligned}\mathcal{L}(\beta,\gamma,\tilde{\alpha},\tilde{\omega}|\mathcal{G}_0) &= p(\text{epidemic events, network events}|\beta,\gamma,\tilde{\alpha},\tilde{\omega},\mathcal{G}_0)\\ &= \gamma^{n_R}\beta^{n_E-1}\alpha_{SS}^{C_{HH}}\alpha_{SI}^{C_{HI}}\alpha_{II}^{C_{II}}\omega_{SS}^{D_{HH}}\omega_{SI}^{D_{HI}}\omega_{II}^{D_{II}}\\ &\quad\times\prod_{j=2}^{n}\left[\tilde{M}(t_j)\left(I_{p_{j1}}(t_j)\right)^{F_j}\right]\\ &\times\exp\left(-\int_0^{T_{\max}}\left[\beta SI(t)+\gamma I(t)+\tilde{\alpha}^T\mathbf{M}_{\max}(t)\right.\right.\\ &\quad\left.\left.+(\tilde{\omega}-\tilde{\alpha})^T\mathbf{M}(t)\right]dt\right).\end{aligned}\tag{14}$$

3.2.2. Inference Given Complete Event Data

The likelihood function (14) will be used toward inference under missing data, but immediately suggests straightforward procedures when the process is fully observed. Given the complete event data $\{e_j\}_{j=1}^n$ and the initial conditions of the process $\mathcal{G}_0$ and $I(0)$, the only unknown quantities in (14) are the model parameters $\Theta = \{\beta,\gamma,\alpha_{SS},\alpha_{SI},\alpha_{II},\omega_{SS},\omega_{SI},\omega_{II}\}$. The following Theorems state results on maximum likelihood estimation as well as Bayesian estimation.

Theorem 3.1 (Maximum likelihood estimation). Following the likelihood function in (14), given $\mathcal{G}_0$ and complete event data $\{e_j\}$, the MLEs of the model parameters are given as follows:

$$\begin{aligned}\hat{\beta} &= \frac{n_E-1}{\sum_{j=1}^n SI(t_j)(t_j-t_{j-1})},\\ \hat{\gamma} &= \frac{n_R}{\sum_{j=1}^n I(t_j)(t_j-t_{j-1})},\\ \hat{\alpha}_{SS} &= \frac{C_{HH}}{\sum_{j=1}^n\left[\frac{H(t_j)(H(t_j)-1)}{2}-M_{HH}(t_j)\right](t_j-t_{j-1})},\\ \hat{\omega}_{SS} &= \frac{D_{HH}}{\sum_{j=1}^n M_{HH}(t_j)(t_j-t_{j-1})},\\ \hat{\alpha}_{SI} &= \frac{C_{HI}}{\sum_{j=1}^n\left[H(t_j)I(t_j)-M_{HI}(t_j)\right](t_j-t_{j-1})},\\ \hat{\omega}_{SI} &= \frac{D_{HI}}{\sum_{j=1}^n M_{HI}(t_j)(t_j-t_{j-1})},\\ \hat{\alpha}_{II} &= \frac{C_{II}}{\sum_{j=1}^n\left[\frac{I(t_j)(I(t_j)-1)}{2}-M_{II}(t_j)\right](t_j-t_{j-1})},\\ \hat{\omega}_{II} &= \frac{D_{II}}{\sum_{j=1}^n M_{II}(t_j)(t_j-t_{j-1})}.\end{aligned}$$

The above results can be directly obtained by setting all partial derivatives of the log-likelihood to zero. The detailed proof is provided in Supplement S2.

[3]The likelihood derived here can be slightly modified to describe an SIS-type epidemic instead; see Supplement S1.

Theorem 3.2 (Bayesian inference with conjugate priors). Under Gamma priors

$$\beta\sim\text{Ga}(a_\beta,b_\beta),\quad \gamma\sim\text{Ga}(a_\gamma,b_\gamma),\quad \alpha_{\cdot\cdot}\sim\text{Ga}(a_\alpha,b_\alpha),\quad \omega_{\cdot\cdot}\sim\text{Ga}(a_\omega,b_\omega)$$

and given initial network $\mathcal{G}_0$ and complete data $\{e_j\}$, the posterior distributions of model parameters under likelihood (14) are given by

$$\begin{aligned}&\beta|\{e_j\}\sim\text{Ga}(a_\beta+(n_E-1),b_\beta+(n_E-1)/\hat{\beta}),\\ &\gamma|\{e_j\}\sim\text{Ga}\left(a_\gamma+n_R,b_\gamma+n_R/\hat{\gamma}\right),\\ &\alpha_{SS}|\{e_j\}\sim\text{Ga}\left(a_\alpha+C_{HH},b_\alpha+C_{HH}/\hat{\alpha}_{SS}\right),\\ &\omega_{SS}|\{e_j\}\sim\text{Ga}\left(a_\omega+D_{HH},b_\omega+D_{HH}/\hat{\omega}_{SS}\right),\\ &\alpha_{SI}|\{e_j\}\sim\text{Ga}\left(a_\alpha+C_{HI},b_\alpha+C_{HI}/\hat{\alpha}_{SI}\right),\\ &\omega_{SI}|\{e_j\}\sim\text{Ga}\left(a_\omega+D_{HI},b_\omega+D_{HI}/\hat{\omega}_{SI}\right),\\ &\alpha_{II}|\{e_j\}\sim\text{Ga}\left(a_\alpha+C_{II},b_\alpha+C_{II}/\hat{\alpha}_{II}\right),\\ &\omega_{II}|\{e_j\}\sim\text{Ga}\left(a_\omega+D_{II},b_\omega+D_{II}/\hat{\omega}_{II}\right),\end{aligned}\tag{15}$$

where $\hat{\beta},\hat{\gamma},\hat{\alpha}_{SS},\hat{\alpha}_{SI},\hat{\alpha}_{II},\hat{\omega}_{SS},\hat{\omega}_{SI},\hat{\omega}_{II}$ are the MLEs defined in Theorem 3.1.

Equation (14) is consistent with the general form of likelihood of a continuous-time Markov chain with Exponentially distributed dwell times. Applying the Gamma-Exponential conjugacy leads to the posterior distributions in (15).

3.2.3. Relaxing the Closed Population Assumption

Above, the host population was implicitly assumed to be closed by fixing N. If in reality the observed population of size N is a subset of a larger unobserved population, then it is possible for an individual to get infected by an external source. This is not reflected in the likelihood above as there is no corresponding S-I link within the observed population, so we introduce an "external infection" rate ξ describing the rate for each susceptible individual to contract the disease from an external source. In other words, ξ can be thought of as the constant rate for any S individual to enter status I independently of interaction with infectious members in the observed population. In this scenario, the complete data likelihood becomes

$$\begin{aligned}\mathcal{L}(\beta,\xi,\gamma,\tilde{\alpha},\tilde{\omega}|\mathcal{G}_0) &= \gamma^{n_R}\alpha_{SS}^{C_{HH}}\alpha_{SI}^{C_{HI}}\alpha_{II}^{C_{II}}\omega_{SS}^{D_{HH}}\omega_{SI}^{D_{HI}}\omega_{II}^{D_{II}}\\ &\quad\prod_{j=2}^{n}\left[\tilde{M}(t_j)\left(\beta I_{p_{j1}}(t_j)+\xi\right)^{F_j}\right]\\ &\quad\times\exp\left(-\int_0^{T_{\max}}\left[\beta SI(t)+\xi S(t)+\gamma I(t)+\tilde{\alpha}^T\mathbf{M}_{\max}(t)\right.\right.\\ &\quad\left.\left.+(\tilde{\omega}-\tilde{\alpha})^T\mathbf{M}(t)\right]dt\right).\end{aligned}$$

The MLEs for $\{\gamma,\tilde{\alpha},\tilde{\omega}\}$ remain unchanged, and though there is no longer a closed-form solution to the MLEs for β and ξ, numerical solutions can be easily obtained as detailed in Supplement S3. In particular, if we have information on which infection cases are caused by internal sources (described by β) and which are caused by external sources (described by ξ), we can directly obtain the MLEs (and posterior distributions) for

all the parameters. In this case, estimation for all parameters except β and ξ remains unchanged. When there is missingness in recovery times, the Bayesian inference procedure proposed in the next section can still be carried out with only minor adaptations.

4. Inference With Partial Epidemic Observations

Though likelihood-based inference is straightforward when all events are observed, complete event data are rarely collected in real-world epidemiological studies. Even in epidemiological studies with very comprehensive observations (Aiello et al. 2016), there still exists some degree of missingness in the exact individual recovery times. In these data, infection times are recorded when a study subject reports symptoms, but recoveries are aggregated at a coarse time scale rather than immediately recorded when a subject becomes disease-free.

Incomplete observations on epidemic paths have long been a major challenge for inference, even when assuming a randomly mixing population or a simple, fixed network structure. With the exact times of infections and/or recoveries unknown, it essentially requires integrating over all possible individual disease episodes to obtain the *marginal* likelihood of the parameters comprised of finite-time transition probabilities. While this can be done for some classes of processes (Xu et al. 2015; Xu and Minin 2015; Ho, Crawford, and Suchard 2018; Tavaré 2018), in most cases it is computationally intractable. Instead, our strategy is to bypass the need for direct marginalization through data augmentation. This entails treating the unknown quantities in data as latent variables, iteratively imputing their values, and then estimating parameters given the current setting of latent variables. By *augmenting* the data via latent variables, the parameter estimation step makes use of the computationally tractable complete-data likelihood. This class of methods has proven successful for related problems based on individual-level data (Auranen et al. 2000; Höhle and Jørgensen 2002; Cauchemez et al. 2006; Hoti et al. 2009; Tsang et al. 2019), population-wide prevalence counts (Fintzi et al. 2017), and observations on a structured but static population (Neal and Roberts 2004; O'Neill 2009; Tsang et al. 2019), but has not yet been designed for an epidemic process coupled with a *dynamic* network. The time-varying nature of social interactions imposes complex constraints on the data augmentation. Though network dynamics complicate the design of data-augmented samplers, the information they provide on possible infection sources and transmission routes allow us to exploit additional structure, effectively reducing the size of the latent space.

We derive a data augmentation method specifically designed to enable inference under missing recovery times. The algorithm utilizes the information presented by the dynamic contact structure. In contrast to existing methods (such as Fintzi et al. (2017) and Hoti et al. (2009)), it is able to efficiently impute unobserved event times in parallel instead of updating individual trajectories one by one. We focus on the case when recovery times are missing, directly motivated by the case study data in Section 6, but note that the proposed framework applies to other sources of missing data; see Section 7 for discussion.

4.1. Method Overview

4.1.1. Problem Setting

Throughout the observation period $(0, T_{\max}]$, suppose $\{(u_\ell, v_\ell]\}_{\ell=1}^{L}$ ($u_\ell < v_\ell$ and $v_\ell \le u_{\ell+1}$) is the collection of *disjoint* time intervals in which a certain number of recoveries occur, but the exact times of those recoveries are unknown. That is, for each $\ell = 1, 2, \ldots, L$, some individuals are reported as infectious up to time u_ℓ, and they are reported as healthy again starting from time v_ℓ. Within one particular interval $(u_\ell, v_\ell]$, let $n_E^{(\ell)}$ be the number of infections, and $n_R^{(\ell)}$ the number of recoveries for which the *exact times are known*, so the number of *unknown* recovery times for this interval is $R_\ell = I(u_\ell) - I(v_\ell) + n_E^{(\ell)} - n_R^{(\ell)}$. Denote these recovery times by latent variables $\mathbf{r}_\ell = \{r_{\ell,1}, \ldots, r_{\ell,R_\ell}\}$. Our goal is to conduct inference despite the absence of all the exact recovery times $\mathbf{r} = \{\mathbf{r}_\ell\}_{\ell=1:L}$ in the observed data.

Further assume that we have a health status report (indicating ill or healthy) of each individual periodically during $(0, T_{\max}]$. Access to such information is usually granted in epidemiological studies where every study subject gives updates on health statuses through regular surveys (e.g., weekly surveys).

4.1.2. Inference Scheme

We propose to address the problem of missing recovery times through data-augmented Markov chain Monte Carlo. Given an initial guess of parameter values $\Theta^{(0)}$ and the observed data $\mathbf{x} = \{e_j\} \cup \{\text{health status reports}\} \cup \mathcal{Z}_0$, at each iteration s the algorithm samples a set of values for the missing recovery times $\mathbf{r}^{(s)} = \{\mathbf{r}_\ell^{(s)}\}_{\ell=1:L}$ from their probability distribution conditioned on $\mathbf{x}$ and the current draw of parameter values. It then samples a new set of parameter values $\Theta^{(s)}$ from their posterior distributions conditioned on the augmented data. In summary, for $s = 1 : S$ where S is the maximum iteration count:

1. *Data augmentation:* Draw $\mathbf{r}^{(s)} = \{\mathbf{r}_\ell^{(s)}\}_{\ell=1:L}$ from the joint conditional distribution
$$p(\mathbf{r} | \Theta^{(s-1)}, \mathbf{x}, \mathbf{r}^{(s-1)}). \tag{16}$$
2. *Parameter update:* Combine $\mathbf{x}$ and $\mathbf{r}^{(s)}$ to form the augmented, complete data. Sample parameters $\Theta^{(s)} \mid \mathbf{x}, \mathbf{r}^{(s)}$ according to (15).

4.2. Data Augmentation via Endpoint-Conditioned Sampling

In the inference scheme stated above, the data augmentation step (Step 1) is challenging because (16) describes the distribution of missing recovery times conditioned on both historical events *and* future events. Thus, drawing from (16) amounts to sampling unobserved event times from a continuous-time Markov process with a series of fixed endpoints (Hobolth and Stone 2009), a challenging task. Even though (3) suggests that, in *forward simulations*, the time it takes for an infectious person to recover only depends on the recovery rate γ, when recovery times need to be inferred *retrospectively*, there are additional constraints imposed by the observed data. First, an individual q cannot recover before a certain time point t if it is observed that at time t the person is still ill. More subtly, if another

individual p gets infected during his contact with q, then the associated recovery time for q cannot leave p without a possible infection source. The first condition is easy to satisfy. The second constraint is much more complicated due to the network dynamics, which a simple forward simulation approach would fail to effectively accommodate.

We tackle the challenge in data augmentation by first simplifying the expression of (16) and then stating an efficient sampling algorithm.

Lemma 4.1. (16) can be simplified into the following expression:

$$\prod_{\ell=1:L} p\left(\mathbf{r}_\ell | \gamma^{(s-1)}, \{e_j\}_{t_j \in (u_\ell, v_\ell]}, \mathcal{Z}_{u_\ell}\right), \tag{17}$$

where $\mathcal{Z}_t$ is the state of the process at time t, including the epidemic status of each individual and the social network structure.

Proof. Consider the joint density of the complete data given parameter values $\Theta^{(s-1)}$.

$$\begin{aligned}
&p\left(\mathbf{x}, \{\mathbf{r}_\ell\}_{\ell=1:L} | \Theta^{(s-1)}\right) \\
&= \prod_{\ell=1:L} \left[p\left(\{e_j\}_{t_j \in (u_\ell, u_{\ell+1}]}, \mathbf{r}_\ell | \mathcal{Z}_{u_\ell}, \Theta^{(s-1)}\right)\right] \\
&\quad \times p\left(\{e_j\}_{t_j \le u_1 \text{ or } t_j > v_L} | \mathcal{Z}_0, \mathcal{Z}_{v_L}, \Theta^{(s-1)}\right) \\
&= \prod_{\ell=1:L} \left[p\left(\{e_j\}_{t_j \in (u_\ell, v_\ell]} | \mathbf{r}_\ell, \mathcal{Z}_{u_\ell}, \Theta^{(s-1)}\right) p\left(\mathbf{r}_\ell | \mathcal{Z}_{u_\ell}, \gamma^{(s-1)}\right)\right] \\
&\quad \times \left[\prod_{\ell=1:L} p\left(\{e_j\}_{t_j \in (v_\ell, u_{\ell+1}]} | \mathcal{Z}_{v_\ell}, \Theta^{(s-1)}\right)\right] \\
&\quad \times p\left(\{e_j\}_{t_j \le u_1 \text{ or } t_j > v_L} | \mathcal{Z}_0, \mathcal{Z}_{v_L}, \Theta^{(s-1)}\right).
\end{aligned}$$

Examining all terms concerning $\mathbf{r}_\ell$ for each ℓ indicates that, conditioned on $\gamma^{(s-1)}$, $\{e_j\}_{t_j \in (u_\ell, v_\ell]}$, and $\mathcal{Z}_{u_\ell}$, the distribution of $\mathbf{r}_\ell$ does *not* depend on $\{\mathbf{r}_{\ell'}\}_{\ell' \neq \ell}$. Thus,

$$p(\mathbf{r} | \Theta^{(s-1)}, \mathbf{x}, \mathbf{r}^{(s-1)}) = \prod_{\ell=1:L} p\left(\mathbf{r}_\ell | \gamma^{(s-1)}, \{e_j\}_{t_j \in (u_\ell, v_\ell]}, \mathcal{Z}_{u_\ell}\right).$$

□

The lemma above suggests that imputation of missing recovery times inside an interval $(u, v]$ only depends on the events that occur in $(u, v]$, the state of the process at the start of the interval, $\mathcal{Z}_u$, and the value of recovery rate γ. Further, imputation on disjoint intervals can be conducted separately and in parallel.

Now consider sampling recovery times within any interval $(u, v]$. Let $\mathcal{Q}$ denote the group of individuals who recover at unknown times during $(u, v]$, and for each $q \in \mathcal{Q}$, let q's exact recovery time be r_q ($\in (u, v]$). Similarly, let $\mathcal{P}$ denote the group of individuals who get infected during $(u, v]$; for $p \in \mathcal{P}$, let p's infection time be i_p, $\mathcal{N}_p(i_p)$ be the set of p's contacts at time i_p, and $\mathcal{I}(i_p)$ be the set of *known* infectious individuals at time i_p (i.e., $\mathcal{I}(i_p)$ excludes any individual who may have recovered before i_p).

Proposition 4.1 (Data augmentation regulated by contact information (DARCI)). Following the notation stated above, given a recovery rate γ, the state of the process at time u, $\mathcal{Z}_u$, and all the observed events in the interval $(u, v]$, $\{e_j\}_{t_j \in (u,v]}$, one can sample $\{r_q\}$ from the conditional distribution $p\left(\{r_q\} | \gamma, \{e_j\}_{t_j \in (u,v]}, \mathcal{Z}_u\right)$ in the following steps:

1. Initialize a vector LB of length $|\mathcal{Q}|$ with $\text{LB}_q = u$ for every $q \in \mathcal{Q}$; then for any $p \in \mathcal{P}$ such that $p \in \mathcal{Q}$, further set $\text{LB}_p = u$;
2. Arrange the set $\mathcal{P}$ in the order of $\{p_1, p_2, \ldots, p_{|\mathcal{P}|}\}$ such that $i_{p_1} < i_{p_2} < \cdots < i_{p_{|\mathcal{P}|}}$, and for each $p \in \mathcal{P}$ (chosen in the arranged order), examine the "potential infectious neighborhood"
$$\mathcal{I}_p = \mathcal{N}_p(i_p) \cap \left(\mathcal{I}(i_p) \cup \mathcal{Q}\right).$$
If $\mathcal{I}_p \subset \mathcal{Q}$ (i.e., potential infection sources are all members of $\mathcal{Q}$), then randomly and uniformly select one $q \in \mathcal{I}_p$, and set $\text{LB}_q = i_p$.
3. Draw recovery times $r_q \overset{\text{ind}}{\sim} \text{TEXP}(\gamma, \text{LB}_q, v)$, where $\text{TEXP}(\gamma, s, t)$ is a truncated Exponential distribution with rate γ and truncated on the interval (s, t).

Intuitively, this procedure enables a draw of recovery times that are "consistent with" the observed data. To achieve this goal, an imputed recovery cannot occur in a way that leaves a to-be-infected individual without any infectious neighbor at the time of infection, nor take place before the corresponding individual gets infected. Effectively there is a "lower bound" for each missing recovery time conditioned on the observed data, particularly the dynamic contact structure. An illustration of the DARCI algorithm is provided in Figure 1.

Combining Lemma 4.1 and Proposition 4.1 enables exact sampling from the conditional distribution (17) in the data augmentation step: for each $\ell = 1, 2, \ldots, L$, applying the DARCI algorithm to the interval $(u_\ell, v_\ell]$ gives an updated set of missing recovery times, $\mathbf{r}_\ell = \{r_{\ell,i}\}_{i=1:R_\ell}$. This allows us to carry out MCMC sampling using a simple Gibbs sampler.

5. Simulation Experiments

In this section, we present results of a series of experiments with simulated datasets. In all experiments, we employ a forward simulation procedure that can be seen as a variation of Gillespie's algorithm (Gillespie 1976) to sample realizations of the network epidemic from our generative model. The input consists of the parameter values $\Theta = \{\beta, \gamma, \tilde{\alpha}, \tilde{\omega}\}$,[4] an arbitrary initial network $\mathcal{G}_0$, the number of infectious cases at onset $I(0)$, and the observation time length $T_{\max}$. The output is the complete collection of all events $\{e_j = (t_j, p_{j1}, p_{j2}, F_j, C_j, D_j)\}$ that occur within the time interval $(0, T_{\max}]$. Associated with each event e_j is a timestamp (t_j), labels of the individuals involved (p_{j1}, p_{j2}), and the event-type indicator F_j, C_j, or D_j.

The steps of the simulation procedure are detailed as follows:

[4] Here $\tilde{\alpha} = (\alpha_{SS}, \alpha_{SI}, \alpha_{II})^T$ and $\tilde{\omega} = (\omega_{SS}, \omega_{SI}, \omega_{II})^T$, as defined in (10) and (11).

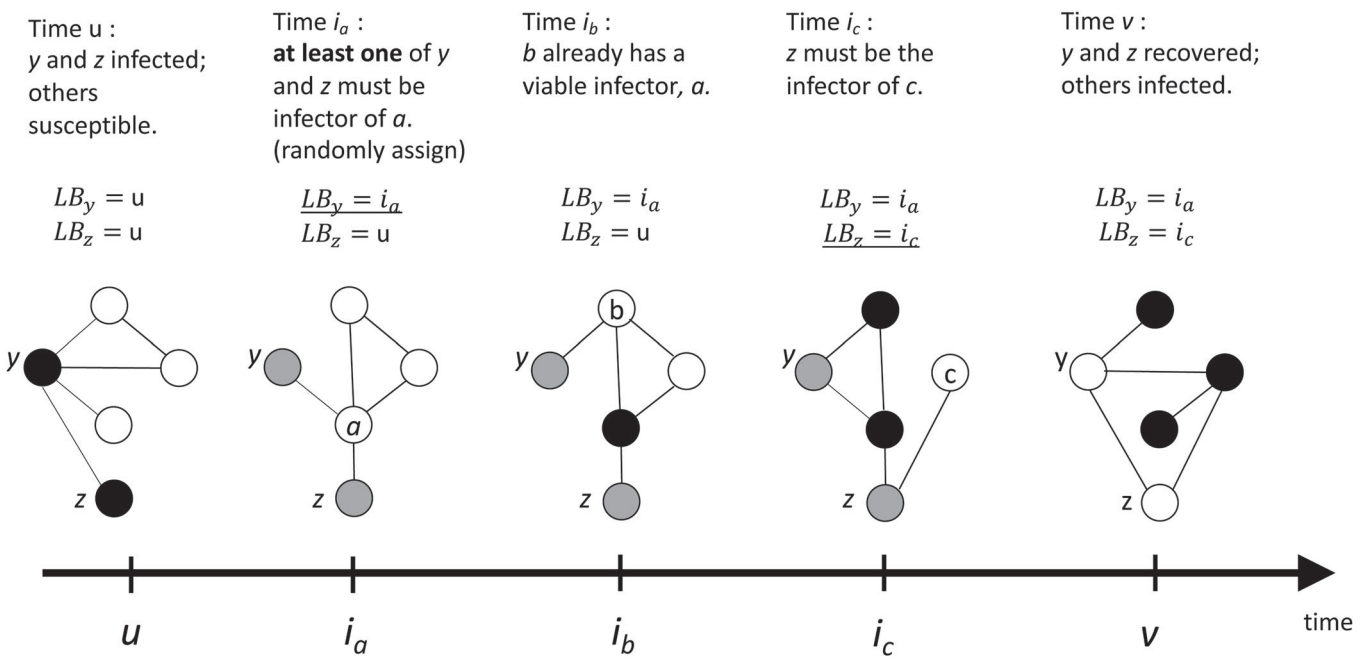

Figure 1. Illustration of DARCI on a $N = 5$ population. Each circle represents an individual and each solid line represents a link. Disease status is color-coded: dark = infectious, gray = unknown (possibly infectious or recovered), and white = healthy (susceptible or recovered). Individuals *y* and *z* are known to be infectious at time *u* but are recovered by time *v*, and individuals *a*, *b*, and *c* are known to get infected at time points i_a, i_b, and i_c, respectively. For each person $p \in \{a, b, c\}$, the DARCI algorithm inspects *p*'s contacts at infection time i_p, and updates "lower bounds" (*LB*) of *y* and *z*'s recovery times to ensure that *p* has an infector. The lower bounds that are updated in each step of the example above are underlined. For example, at time i_a, one of *y* and *z* has to be *a*'s infector, so DARCI randomly selects *one* of *y* and *z* (in this example it's *y*) and postpones his recovery time until after i_a.

1. Initialization: Randomly select which $I(0)$ individuals are infectious (then the rest of the population are all susceptible). Set $t_{\text{cur}} = 0$.
2. Iterative update: While $t_{\text{cur}} < T_{\max}$, do:
   (a) Bookkeeping: Summarize the following statistics at t_{cur}: 1) $SI(t_{\text{cur}})$, the number of S-I links in the population; 2) $\mathbf{M}_{\max}(t_{\text{cur}})$, the possible number of links of each type defined in (12); 3) $\mathbf{M}(t_{\text{cur}})$, the number of existing links of each type defined in (13). Then set $\mathbf{M}^d(t_{\text{cur}}) = \mathbf{M}_{\max}(t_{\text{cur}}) - \mathbf{M}(t_{\text{cur}})$.
   (b) Next event time: Compute the instantaneous rate of the occurrence of any event, $\Lambda(t_{\text{cur}}) = \beta SI(t_{\text{cur}}) + \gamma I(t_{\text{cur}}) + \tilde{\alpha}^T \mathbf{M}^d(t_{\text{cur}}) + \tilde{\omega}^T \mathbf{M}(t_{\text{cur}})$, and draw $\Delta t \sim \text{Exponential}(\Lambda(t_{\text{cur}}))$.
   (c) Next event type: Sample $Z \sim \text{Multinomial}(\tilde{\lambda}(t_{\text{cur}}))$, where

$$\tilde{\lambda}(t_{\text{cur}}) = \left(\frac{\beta SI(t_{\text{cur}})}{\Lambda(t_{\text{cur}})}, \frac{\gamma I(t_{\text{cur}})}{\Lambda(t_{\text{cur}})}, \frac{\tilde{\alpha}^T \mathbf{M}^d(t_{\text{cur}})}{\Lambda(t_{\text{cur}})}, \frac{\tilde{\omega}^T \mathbf{M}(t_{\text{cur}})}{\Lambda(t_{\text{cur}})} \right)^T.$$

   Then do one of the following based on the value of Z:
   If $Z = 1$ (infection), uniformly pick one *S-I* link and infect the *S* individual in this link.
   If $Z = 2$ (recovery), uniformly pick one *I* individual to recover.
   If $Z = 3$ (link activation), randomly select $Y \in \{H\text{-}H, H\text{-}I, I\text{-}I\}$ with probabilities proportional to $\tilde{\alpha} \circ \mathbf{M}^d(t_{\text{cur}})$, and uniformly pick one de-activated "Y link" to activate.
   If $Z = 4$ (link termination), randomly select $Y \in \{H\text{-}H, H\text{-}I, I\text{-}I\}$ with probabilities proportional to $\tilde{\omega} \circ \mathbf{M}(t_{\text{cur}})$, and uniformly pick one existing "Y link" to terminate.
   (d) Replace t_{cur} by $t_{\text{cur}} + \Delta t$, record relevant information about the sampled event, and repeat from (a).

In Step 2 (c), "$\circ$" refers to the Hadamard product (entrywise product) for two vectors.

5.1. Experiments With Complete Observations

In this subsection, we first demonstrate the insufficiency of analyzing disease spread without considering the network structure or its dynamics. Then we validate our claims on maximum likelihood estimation and Bayesian inference given complete event data (Theorems 3.1 and 3.2). Finally, we show that the model estimators can detect simpler models such as the decoupled process and the static network process. Unless otherwise stated, throughout this section we set the initial network $\mathcal{G}_0$ as a random Erdős–Rényi graph[5] (undirected) with edge probability $p = 0.1$, let $I(0) = 1$ individual to get infected at onset, and choose the ground-truth parameters as

$$\beta = 0.03, \gamma = 0.12; \tilde{\alpha}^T = (0.005, 0.001, 0.005), \quad \tilde{\omega}^T = (0.05, 0.1, 0.05). \tag{18}$$

These settings are chosen to produce simulated datasets with a population size and event counts that are comparable to our real data example.

[5]We note that the form of the initial network does not necessarily predict the behavior of the epidemic. Specifically, asymptotic qualities, such as the Poisson degree distribution of Erdős–Rényi graphs, do not hold when the network dynamically reacts to an epidemic, detailed empirically in Supplement S4.

For Bayesian inference, we adopt the following Gamma priors for the parameters:

$$\beta \sim \text{Ga}(1, 1/0.02), \gamma \sim \text{Ga}(1, 1/0.1); \alpha_{\cdot\cdot} \sim \text{Ga}(1, 1/0.004), \quad \omega_{\cdot\cdot} \sim \text{Ga}(1, 1/0.06). \tag{19}$$

We intentionally choose prior means different from the true parameter values; experiments show that inference is insensitive to prior specifications as long as a reasonable amount of data is available. For each parameter, 1000 posterior samples are drawn after a 200-iteration burn-in period.

5.1.1. The Danger of Neglecting Networks or Network Dynamics

Adopting the "random mixing" assumption about an actually networked population can lead to severe under-estimation of the infection rate. Erroneous estimation can also happen if contacts are in fact dynamic but are mistaken as static during inference. Table 2 displays the MLEs of the infection rate β (*under true value 0.05*, chosen to generate nontrivial epidemics to illustrate inference) obtained by methods under three different assumptions regarding the network structure (assuming a dynamic network, assuming a static network, and assuming random mixing without any network). The population size is $N = 50$, and results are summarized over 50 different simulated datasets.

These results make clear that neglecting the effects of the contact network, even when the quantity of interest is the disease transmission rate, is dangerously misleading. Resulting estimates that are far from the truth with significantly under-estimated uncertainty measures. Incorporating the initial network structure statically throughout the process helps—the 95% confidence interval now includes the truth—but disregarding the time-evolution of the network remains a noticeable model misspecification leading to biased inference.

Table 2. Maximum likelihood estimates of β, the per link infection rate (true value 0.05), using dynamic network information, the initial static network, and no network structure (random mixing), respectively.

Method	Dynamic network	Static network	No network
Estimate	0.0540	0.0278	0.00219
Standard deviation	0.0158	0.0081	0.000821
2.5% quantile	0.0230	0.00825	0.000614
97.5% quantile	0.0817	0.0553	0.00425

NOTE: The standard deviations as well as the 2.5% and 97.5% quantiles of the estimates are obtained from outcomes across 50 different simulated datasets on a $N = 50$ population.

5.1.2. Validity and Efficacy of Parameter Estimation

Complete event data are generated using the simulation procedure stated in above, and maximum likelihood estimates (MLEs) as well as Bayesian estimates are obtained for parameters $\Theta = \{\beta, \gamma, \tilde{\alpha}, \tilde{\omega}\}$. Here we set the population size as $N = 100$ and the infection rate as $\beta = 0.03$ while keeping the other parameter values the same as stated in (18).

Figure 2 shows the results of maximum likelihood estimation in one simulated dataset. The MLEs for each parameter

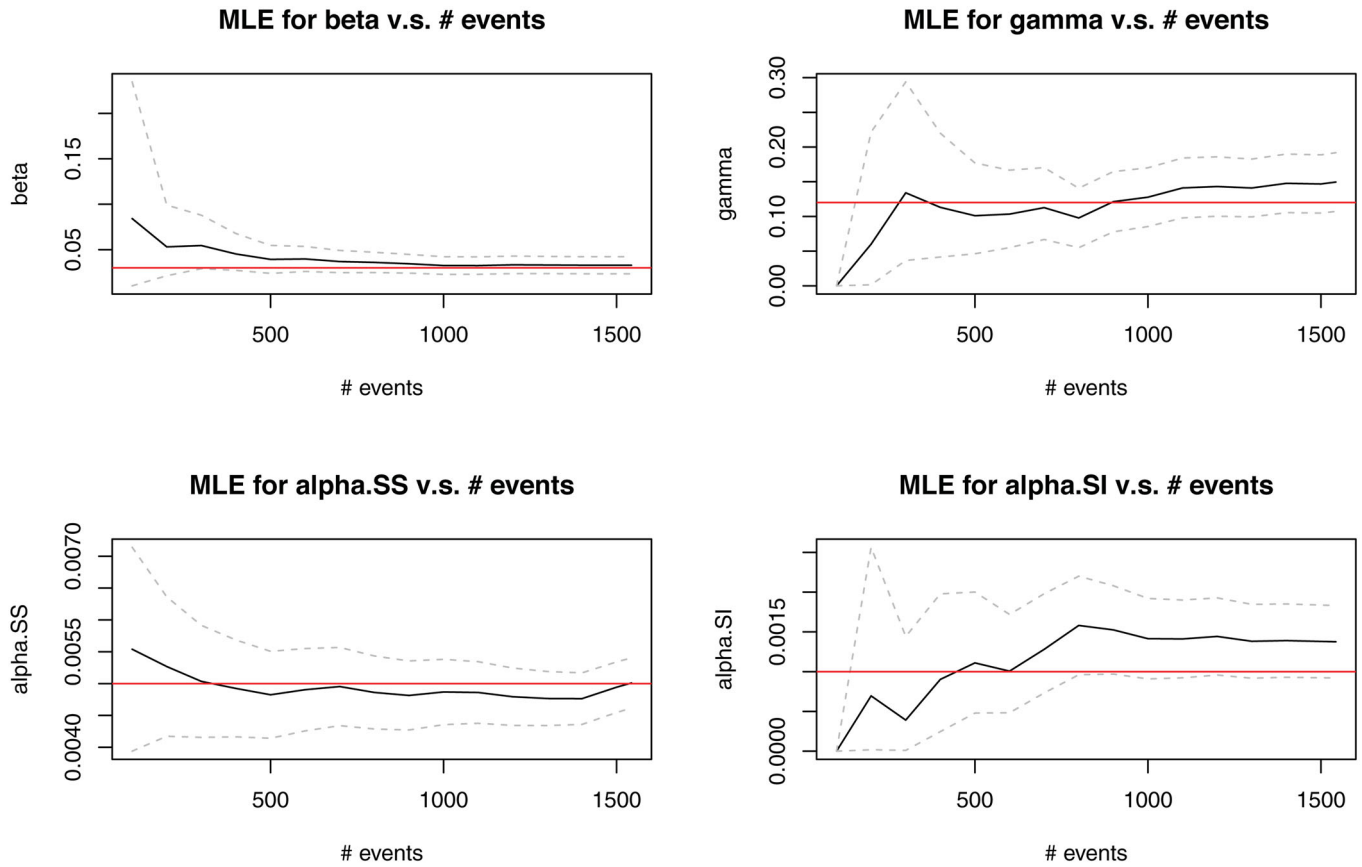

Figure 2. MLEs versus number of events used for inference. Dashed gray lines show the lower and upper bounds for 95% frequentist confidence intervals, and red lines mark the true parameter values. Results are presented for $\beta, \gamma, \alpha_{SS}$ and α_{SI}. In this realization $n_E = n_R = 48, C_{HH} = 621, C_{HI} = 35, C_{II} = 13, D_{HH} = 573, D_{HI} = 189, D_{II} = 17$.

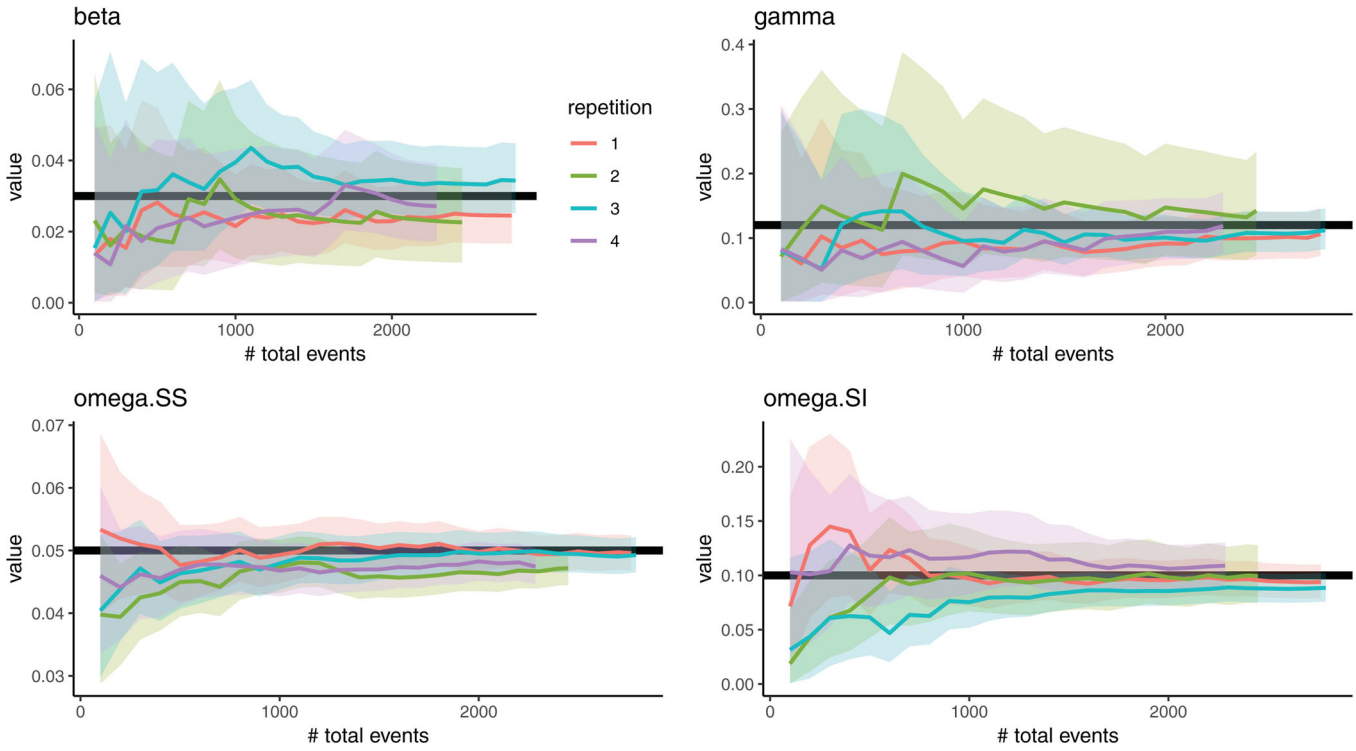

Figure 3. Posterior sample means versus number of total events used for inference. True parameter values are marked by *bold dark* horizontal lines, along with 95% credible bands. Results are presented for four different complete datasets and for parameters $\beta, \gamma, \omega_{SS}$, and ω_{SI}.

(dark solid line) are computed using various numbers of events, and are compared with the true parameter value (red horizontal lines). The lower and upper bounds for 95% confidence intervals are also calculated (dashed gray lines). Only the MLEs for parameters $\beta, \gamma, \alpha_{SS}$, and α_{SI} are shown, but results for all parameters are included in Supplement S5. Estimation is relatively accurate even when observation ends earlier than the actual process (thus leaving later events unobserved). When more events are available for inference, accuracy is improved and the uncertainty is reduced.

Figure 3 presents the posterior sample means (solid lines) and 95% credible bands (shades) for each parameter inferred using various numbers of events, with the true parameter values marked by bold, dark horizontal lines. The results are shown for 4 different simulated datasets (each dataset represented by a distinct color) and for parameters $\beta, \gamma, \omega_{SS}$, and ω_{SI} (complete results are in Supplement S5). When more events are utilized in inference, the posterior means tend to be closer to the true parameter values, while the credible bands gradually narrow down.

It is worth noting that the proposed inferential framework is capable of handling large-scale networks as well as arbitrary network structures. Additional results with larger values of N and different configurations of $\mathcal{G}_0$ are provided in Supplement S5.

5.1.3. Assessing Model Flexibility

Our proposed framework is a generalization of epidemics over networks that evolve independently (the "decoupled" process), which in turn generalize epidemic processes over fixed networks (the "static network" process). Thus, our model class contains these simpler models: if events are generated from the decoupled process, we would expect all the link activation and termination rates to be estimated as the same. Likewise, if the true network process is static, then we expect all link rates to be estimated as zero.

To confirm this, experiments are conducted on complete event datasets generated from the two simpler models. Here we only show select results of Bayesian inference on datasets generated from static network epidemic processes (Figure 4) and relegate other results to Supplement S5. We see that information from a moderate number of events is sufficient to accurately estimate the epidemic parameters β and γ and learn the static nature of the network—note how quickly the posterior credible bands for α_{SI} and ω_{SI} shrink toward zero.

5.2. Experiments With Incomplete Observations

Upon validating the model and inference framework, we now assess the performance of our proposed inference scheme in the more realistic setting where epidemic observations are incomplete. In this subsection, we first verify that the MCMC sampling scheme in Section 4 is able to retrieve the parameter values despite missing recovery times in the observed data. Then we compare our DARCI algorithm (Proposition 4.1) with two baselines and show that it produces posterior samples of higher quality and with higher efficiency.

5.2.1. Simulating Partially Observed Data

We first generate complete event data using the simulation procedure stated earlier in this section, and then randomly discard $\eta \times 100\%$ of the *exact* recovery times and treat them

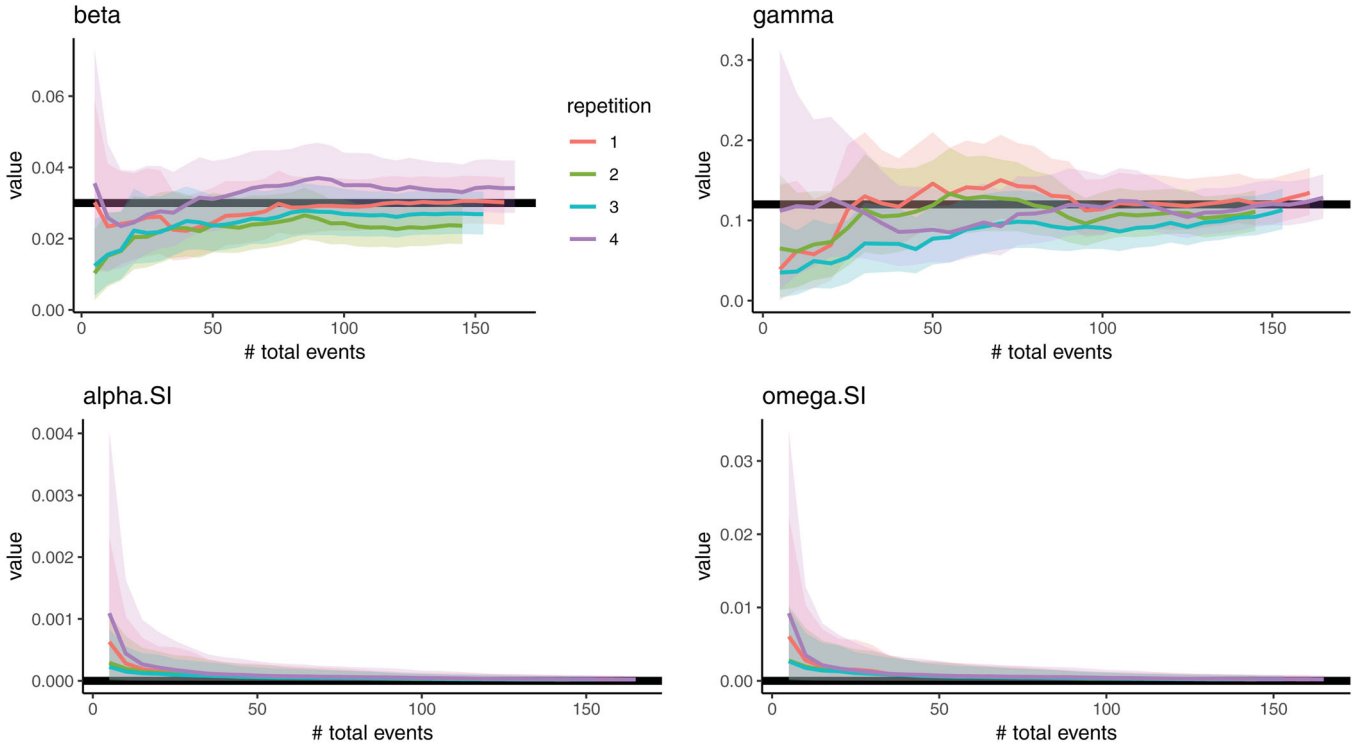

Figure 4. Posterior sample means versus number of total events used for inference, with data generated from *static* network epidemic processes. True parameter values are marked by *bold dark* horizontal lines, along with 95% credible bands. Results are presented for four different complete datasets and for parameters β, γ, α_{SI}, and ω_{SI}. With a moderate number of events, the epidemic-related parameters, β and γ, are accurately estimated, and the posteriors for the edge rates quickly shrink toward zero (the truth).

as unknown. Meanwhile, a periodic status report (as described in Section 4.1) is produced every 7 time units throughout the entire process so that individual disease statuses are informed at a coarse resolution. If one regards 1 time unit as *a day*, the periodical disease statuses correspond to *weekly* reports.

5.2.2. Efficacy of the Inference Scheme

We validate the method outlined in Section 4.1 through experiments on an example dataset, where the settings and parameters are the same as those in (18) and the population size is fixed at $N = 100$. In this particular realization, there are 26 infection cases spanning over approximately 37 days (less than 6 weeks), and there are 767 and 893 instances of social link activation and termination, respectively.[6]

First set $\eta = 50$, that is, randomly select 50% of exact recovery times to be taken as missing. Figure 5 plots 1000 consecutive MCMC samples (after a 200-iteration burn-in) for each parameter $\{\beta, \gamma, \alpha_{SS}, \omega_{SS}\}$, as well as the 2.5% and 97.5% quantiles of the posterior samples (gray, dashed lines) compared with the true parameter value (red horizontal line). We can see that for every parameter, the 95% sample credible interval covers the true parameter value, suggesting that the proposed inference scheme is able to estimate parameters from incomplete data reasonably well.

We then set $\eta = 100$, discarding all exact recovery times. Figure 6 presents outcomes of the inference algorithm in this case. Understandably, parameter estimation is affected by the total unavailability of exact recovery times, but the drop in accuracy is marginal. Moreover, the credible bands are slightly wider, reflecting increased uncertainty with more missingness.

5.2.3. Efficiency of the DARCI Algorithm

We compare performance of the data augmentation algorithm stated in Proposition 4.1 with two conventional sampling methods:

1. Rejection sampling: Carry out Step 1 of the inference scheme via rejection sampling. For $\ell = 1 : L$, keep proposing recovery times $\mathbf{r}_\ell^* = \{r_{\ell,i}^*\}_{i=1:R_\ell} \overset{\text{iid}}{\sim} \text{TEXP}(\gamma^{(s-1)}, u_\ell, v_\ell)$ until the proposed $\mathbf{r}_\ell^*$ are compatible with the observed event data in $(u_\ell, v_\ell]$. We label this method by "*Reject*."
2. Metropolis–Hastings: Modify Step 1 of the inference scheme into a Metropolis–Hastings step. For $\ell = 1 : L$, propose $\mathbf{r}_\ell^* = \{r_{\ell,i}^*\}_{i=1:R_\ell} \overset{\text{iid}}{\sim} \text{TEXP}(\gamma^{(s-1)}, u_\ell, v_\ell)$, and accept them as $\mathbf{r}_\ell^{(s)}$ with probability

$$\min\left(1, \frac{p\left(\mathbf{x}, \mathbf{r}_\ell^*, \{\mathbf{r}_{\ell'}^{(s-1)}\}_{\ell' \neq \ell} | \Theta^{(s-1)}\right) \times \text{pTEXP}\left(\mathbf{r}_\ell^{(s-1)}; \gamma^{(s-1)}, u_\ell, v_\ell\right)}{p\left(\mathbf{x}, \{\mathbf{r}_{\ell'}^{(s-1)}\}_{\ell'=1:L} | \Theta^{(s-1)}\right) \times \text{pTEXP}\left(\mathbf{r}_\ell^*; \gamma^{(s-1)}, u_\ell, v_\ell\right)}\right),$$

which equals to 1 when the proposed $\mathbf{r}_\ell^*$ are consistent with the observed event data in $(u_\ell, v_\ell]$ and 0 otherwise. If the proposal is not accepted, then set $\mathbf{r}_\ell^{(s)} = \mathbf{r}_\ell^{(s-1)}$. We label this method by "*MH*."

[6] The event time scales in the example dataset are chosen to be comparable to, though not exactly the same as, those in the real-world data used in Section 6.

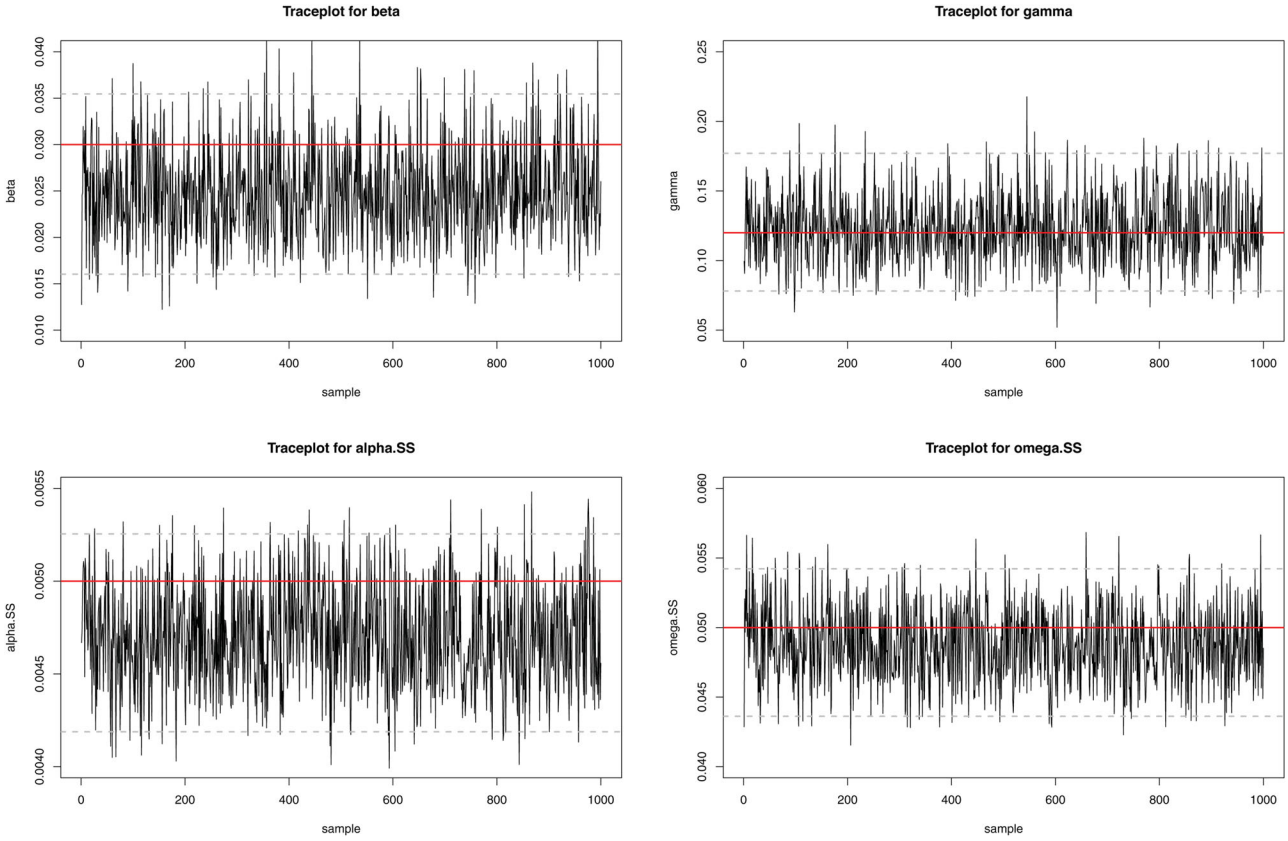

Figure 5. Inference results for parameters $\beta, \gamma, \alpha_{SS}, \omega_{SS}$ with 50% recovery time missingness. The uncertainty in exact recovery times does affect the estimation of the type-dependent edge rates, but not detrimentally (all the true parameter values fall into the 95% credible intervals of the posteriors).

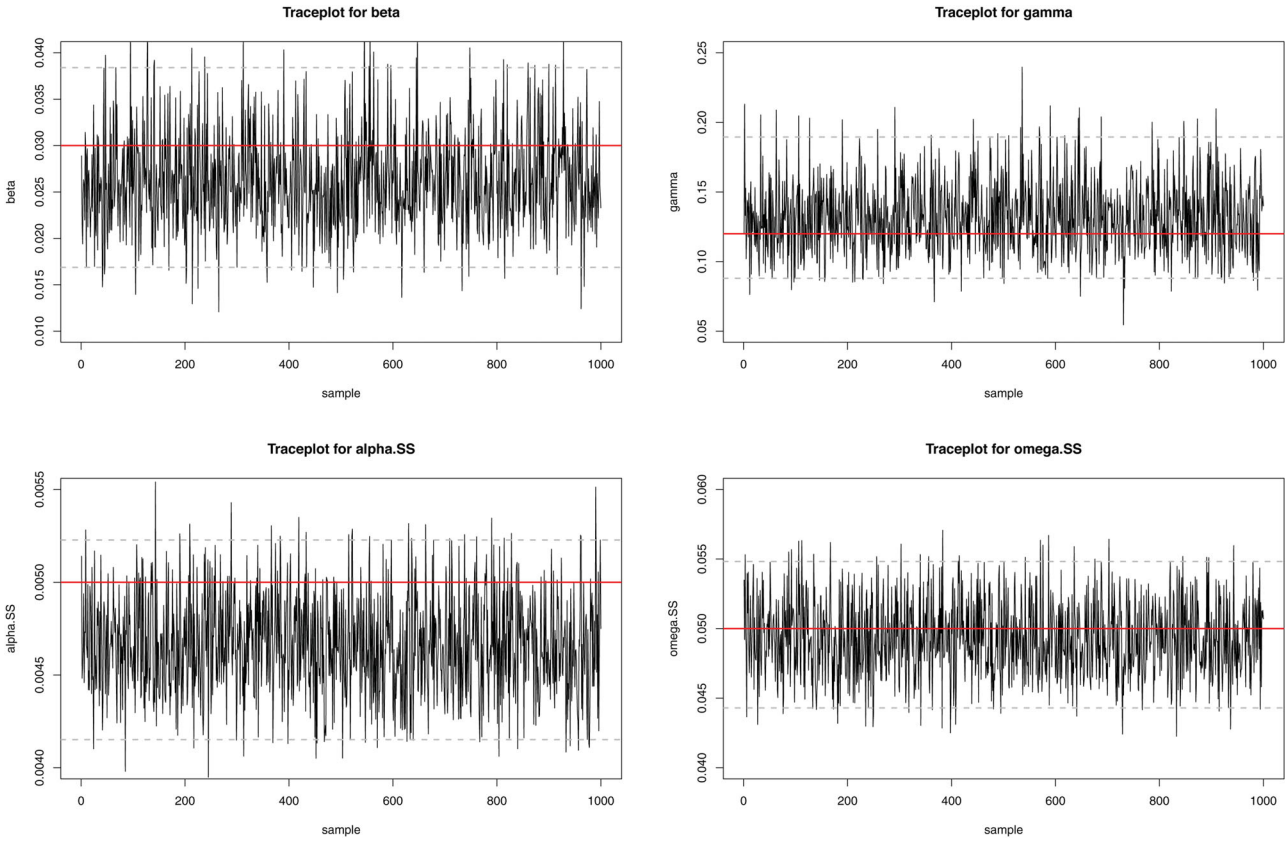

Figure 6. Inference results for parameters $\beta, \gamma, \alpha_{SS}, \omega_{SS}$ with complete (100%) missingness in exact recovery times. The *y*-axis scale for each plot is the same as that in Figure 5 for easier comparison. Wider credible intervals are seen for β and γ, but still, all the true parameter values fall into the 95% credible intervals of the posteriors, suggesting the capability of the inference algorithm to estimate parameters even when there is uncertainty in individual epidemic histories.

Table 3. MCMC diagnostics for three data augmentation sampling methods, labeled as "*DARCI*," "*Reject*," and "*MH*."

Method	Statistic	β	γ	α_{SS}	α_{SI}	α_{II}	ω_{SS}	ω_{SI}	ω_{II}
DARCI	ESS	1000.00	1000.00	1000.00	1000.00	1000.00	1000.00	1000.00	1000.00
	Z-score	−0.90	−0.20	−0.56	−1.32	−0.22	0.84	−0.02	−1.24
	Pr(> |*Z*|)	0.37	0.84	0.58	0.19	0.82	0.40	0.99	0.22
Reject	ESS	1000.00	1160.17	1000.00	955.29	1000.00	1000.00	926.63	1000.00
	Z-score	0.48	−1.01	0.44	0.28	1.08	−2.18	−0.16	0.31
	Pr(> |*Z*|)	0.63	0.31	0.66	0.78	0.28	0.03	0.87	0.76
MH	ESS	566.43	1000.00	1000.00	1000.00	538.12	907.14	729.33	1000.00
	Z-score	−1.25	−1.83	−0.48	−0.59	−2.09	−0.24	−1.52	−0.57
	Pr(> |*Z*|)	0.21	0.07	0.63	0.55	0.04	0.81	0.13	0.57

NOTE: "ESS" stands for "effective sample size." The "*Z*-score" is the test statistic for MCMC convergence proposed by Geweke (1991), and the two-sided *p*-value for each standard *Z*-score is also computed. Samples acquired by *MH* tend to have higher autocorrelations and thus smaller effective sample sizes.

The "*MH*" method employs the same principle as existing agent-based data augmentation methods (Cauchemez et al. 2006; Hoti et al. 2009; Fintzi et al. 2017) that propose candidates of individual disease histories and accept them with probabilities computed through evaluating the likelihood (or an approximation of it) and the proposal density. In our case the implementation of "*MH*" is actually simpler and computationally lighter because the proposal is conditioned on known infection times and the current posterior draw of γ. The acceptance step is reduced to inspecting compatibility with known data, thus avoiding the intensive computation of likelihood evaluation.

Although the three methods give the same inference results since they all sample from the same posterior distributions, our data augmentation algorithm (labeled by "*DARCI*") is more efficient than the others in two aspects. First, by drawing a new sample of recovery times from the conditional distribution in (16) in each iteration, the resulting Markov chain exhibits lower autocorrelation, which leads to better mixing and fewer iterations needed to achieve a certain effective sample size. This is especially so when compared with "*MH*." Second, the DARCI algorithm typically draws from the conditional distribution (16) much more efficiently than "*Reject*," because it parses out a configuration of lower bounds for imputing missing recovery times while accounting for the constraints imposed by contagion spreading and the dynamics of social links.

Three MCMC samplers are run using the three methods on the dataset showcased above. Again 1000 consecutive samples are retained for each parameter after a 200-iteration burn-in period in each case. For each parameter, we calculate the effective sample size (ESS), the Geweke *Z*-score (Geweke 1991), and the two-sided *p*-value for the *Z*-scores of the resulting chains. These results are presented in Table 3. Among the three methods, "*MH*" suffers the most from the correlation between two successive samples, while "*DARCI*" seems to produce high-quality MCMC samples.

We then compare "*DARCI*" and "*Reject*" in their running times (see Table 4). A dataset is simulated where there are different numbers of recoveries with unknown times within five time intervals. The two sampling methods are applied to draw a set of recovery times for each of those five intervals, and over multiple runs, the minimum and median times they take are recorded. Although the two methods draw samples from the same conditional distribution, "*DARCI*" tends to take less time than "*Reject*" in one iteration. Further results suggesting scalability to larger outbreaks are available in Supplement S5.

Table 4. Comparison between the two sampling methods ("*DARCI*" and "*Reject*") for imputing missing recovery times.

		Min time		Median time	
Interval	#(To recover)	*Reject*	*DARCI*	*Reject*	*DARCI*
1	1	227 μsec	224 μsec	484 μsec	245 μsec
2	8	285 μsec	287 μsec	563 μsec	319 μsec
3	15	163 μsec	161 μsec	279 μsec	181 μsec
4	2	138 μsec	138 μsec	153 μsec	156 μsec
5	1	133 μsec	133 μsec	146 μsec	147 μsec

NOTE: Overall, the DARCI algorithm is more efficient, especially when the number of missing recovery times is relatively large (e.g., Interval 3), or there are special constraints on viable recovery times (e.g., Interval 1, where the observed events suggest that the recovery cannot occur until half way through the time interval).

6. Influenza-Like-Illnesses on a University Campus

In this section, we apply the proposed model and inference scheme to a real-world dataset on the transmission of influenza-like illnesses among students on a university campus.

6.1. Data Overview

The data we analyze in this section were collected in a 10-week network-based epidemiological study, eX-FLU (Aiello et al. 2016). The study was originally designed to investigate the effect of social intervention on respiratory infection transmission. 590 university students enrolled in the study and were asked to respond to weekly surveys on influenza-like illness (ILI) symptoms and social interactions. 103 individuals further participated in a sub-study in which each study subject was provided a smartphone equipped with an application, iEpi. The application pairs smartphones with other nearby study devices via Bluetooth, recording individual-level social interactions at 5-min intervals.

The sub-study using iEpi was carried out from January 28, 2013 to April 15, 2013 (from week 2 until after week 10). Between weeks 6 and 7, there was a one-week spring break (March 1 to March 7), during which the volume of recorded social contacts dropped noticeably. In our experiments, we use data collected on the $N = 103$ sub-study participants from January 28 to April 4 (week 2 to week 10), and treat the two periods before and after the spring break as two separate and independent observation periods ($T_{\max} = 31$ days for period 1 and $T_{\max} = 28$ days for period 2).

Summary statistics of the data are provided in Table 5. Overall, infection instance counts peaked in the middle of each

Table 5. Summary statistics of the real data (processed) by week: number of new infection cases (top row), maximum network density (middle row), and minimum network density (bottom row).

Week	Wk 2	Wk 3	Wk 4	Wk 5	Wk 6
# (infections)	0	3	5	4	2
Max. density	0.0053	0.2048	0.0040	0.0038	0.0044
Min. density	0.0000	0.0000	0.0000	0.0002	0.0000
Week	(break)	Wk 7	Wk 8	Wk 9	Wk 10
# (infections)	N.A.	1	3	5	1
Max. Density	N.A.	0.0032	0.0032	0.0032	0.0023
Min. Density	N.A.	0.0000	0.0000	0.0000	0.0000

NOTE: No new infection cases took place in week 2, but two participants were already ill at the beginning of the week. The dynamic network remained sparse throughout the duration of the sub-study, except for one instance in week 3—the unusually high network density only occurred on the night of February 4, possibly due to a large-scale on-campus social event.

Table 6. Posterior sample means and 95% credible intervals of select parameters (first 3 columns) obtained by the Bayesian inference scheme modified from that in Section 4.

Parameter	Posterior mean	2.5% quantile	97.5% quantile	Multi-SD
β (internal infection)	0.0695	0.0247	0.1500	0.0074
ξ (external infection)	0.00331	0.00208	0.00494	1.797×10^{-5}
γ (recovery)	0.294	0.186	0.428	0.0108
α_{SS} (*S-S* link activation)	0.0514	0.0499	0.0529	0.0002
ω_{SS} (*S-S* link termination)	38.26	33.55	40.62	0.2522
α_{SI} (*S-I* link activation)	0.130	0.0785	0.194	0.0097
ω_{SI} (*S-I* link termination)	53.5	22.5	231.7	31.4092

NOTE: Inference is carried out jointly on the two periods before and after the spring break. The final column summarizes the standard deviations of posterior sample means across 10 versions of data, in which infection times are sampled randomly (and randomly) between 0 to 3 days prior to symptom onset.

observation period and dropped at the end, and the dynamic social network was quite sparse; more activity (in both the epidemic process and network process) was observed in the weeks before the spring break. Further details on data cleaning and preprocessing are provided in Supplement S6.

6.2. Analysis

Since the 103 individuals are sub-sampled from the 590 study participants, which are also sub-sampled from the university campus population, we treat the real data as observed on an open population. Following the parameterization introduced at the end of Section 3.2, we include the parameter ξ to denote the rate of infection from an external source for each susceptible individual. Every infectious individual that came into contact with any infectives within 3 days prior to the onset of symptoms is regarded as an internal case (governed by parameter β); otherwise the infection is labeled as an external case (governed by parameter ξ). This enables the inference procedure stated at the end of Section 4.

The data collected during the two observation periods are considered as independent realizations of the same adaptive network epidemic processes. For each parameter, samples drawn in the first 500 iterations are discarded and then every other sample is retained in the next 2000 iterations, resulting in 1000 posterior samples. Table 6 summarizes the posterior sample means and the lower and upper bounds of 95% sample credible intervals for a selection of parameters. The output from one chain is presented here; repeated runs (with different initial conditions and random seeds) yield similar results.

The data provide symptom onset times for flu-like illnesses, which can serve as proxies for the actual infection times, but the former are on average 2 days later than the latter (U.S. Centers for Disease Control and Prevention (CDC) 2018). To address this issue, we assume that the real infection times may be somewhere between 0 and 3 days prior to symptom onset (see Supplement S5.1) and thus randomly sample the latent infection times to generate multiple data versions for inference. Results are similar across different versions of data, suggesting that inference is robust to this choice of handling possible latency periods (this is summarized by the "Multi-SD" column of Table 6).

Our findings suggest that flu-like symptoms spread slowly but recoveries are made rather quickly. For instance, if a susceptible person maintains *one* infectious contact, then he has a probability of approximately 6.71% to contract infection through such contact *within one day*, and yet it takes (on average) a little more than 3 days for someone to no longer feel ill after infection. The external infection force is nonnegligible: given the number of susceptibles in the population (typically about 100), the population-wide external infection rate is approximately $100 \times 0.0033 = 0.33$, implying that an external ILI case is expected to occur every other three days. This is a reasonable estimate consistent with having observed 9 external infection cases within 28 days during the second period.

The inferred link rates reflect an interesting pattern in social interactions in this particular population: individuals are reluctant to establish contact and active contacts are broken off quickly—an average pair of healthy people initiate/restart their interaction after waiting 20 days and then end it after spending less than 40 minutes together. Moreover, it seems that on average a healthy-ill link is activated more frequently than a healthy-healthy link, but the former is also terminated faster—this might be because those students who fell ill in the duration of the study happen to be more socially proactive, but once their healthy social contacts realize they are sick and thus potentially infectious, the contact is cut short to avoid disease contraction.

It is also notable that the sample 95% credible intervals for β, α_{SI}, and ω_{SI} are relatively wide, indicating a high level of uncertainty in the estimation for these parameters. It is challenging to estimate the internal infection rate β because dataset contains only six cases of internal infection in total (five in period 1, one in period 2), providing limited information on the rate of transmission. Similar issues are present for the estimation of α_{SI} and ω_{SI}; since there were no more than five infectious individuals at any given time, network events related to them were few and far between. Moreover, since their exact recovery times are unknown, there is additional uncertainty associated with their exact disease statuses when they activated or terminated social links. Such measure of uncertainty, readily available through stochastic modeling and Bayesian inference, provides valuable insights into the amount of information the data contain and the level of confidence we possess when making conclusions and interpretations. The inference outcomes imply that, for example, the real data sufficiently inform the contact patterns among

healthy individuals in this population but are limited toward understanding how long a healthy person and a symptomatic person typically maintain their contact.

7. Discussion

This article has focused on enabling inference for partially observed epidemic processes on dynamic and adaptive networks. We formulated a continuous-time Markov process model to describe the epidemic-network interplay and derived its complete data likelihood. This leads to the design of conditional sampling techniques that enable data augmented inference methods to accommodate missingness in individual recovery times.

There are several limitations and natural extensions of our model. First, we address the issue of a latency period here by using a sensitivity analysis of the symptom onset time. As infectiousness is the focus of inference, we prefer this approach over, for instance, modeling an additional compartment (i.e., an SEIR model) in favor of model parsimony. We note that the latter is possible by extending our proposed likelihood framework, but introduces additional parameters that are often hard to identify without additional data directly informative of latency or modeling assumptions regarding the latency period (e.g., noninfectious or less infectious when latent).

A second extension of the model pertains to the handling of other missing data types. Motivated by our case study in which infection-related events (symptom onset) are updated daily, whereas recoveries are only provided at a much coarser resolution (in weekly summaries), our current method focuses on imputing missing recovery times. While it is common in real-world data to focus on new cases (WHO 2004, 2020), one may be provided with such incidence data at coarser time resolution so that infection times must also be imputed. Our framework applies in principle to such settings where missing infection times should be accounted for explicitly. One may derive analogous conditional samplers to DARCI, or at worst incur a computational tradeoff. Even without access to sampling from the exact conditional distributions, we can replace the Gibbs step to impute infection times by a Metropolis proposal within each iteration of the MCMC scheme (Gibson and Renshaw 1998; Britton and O'Neill 2002).

Our contributions leverage network information to avoid a common model misspecification, but the current methods are limited to scenarios where such information is completely informed. If network dynamics are only partially observed—for instance link events are missing or exist only on a weekly survey basis—our proposed methods do not immediately apply, but can be extended via further data augmentation over unobserved network event times. Doing so falls under the same likelihood-based framework, yet practical challenges related to mixing of the Markov chain may arise due to the increased latent space. Another viable strategy is to adopt a discrete time model for network evolution that can be seen either as an alternative or an approximation to the link-Markovian jump process we propose. These directions remain open for future work.

We have demonstrated that accounting for changes in the contact structure is critical to accurately estimate disease parameters such as infection and recovery rates. Because our model is a generalization of existing compartmental models, such rates are consistent with their definitions and interpretation in existing literature. Other quantities such as the basic reproductive number R_0, defined as the average number of infections caused by an infectious individual, do not translate as readily (Van Segbroeck, Santos, and Pacheco 2010; Tunc, Shkarayev, and Shaw 2013). This is the case when disease and network properties are conflated: the basic reproductive number depends on the product of infection rate β and number of contacts. The effect of interventions such as quarantine are often incorporated similarly, for instance by way of a change-point in β (Ho, Crawford, and Suchard 2018), yet such policies should naturally translate to changes in the contact network rather than the inherent infectivity of the disease. Because our model mechanistically describes the joint dynamics of the network and the disease spread, such phenomena can now be modeled explicitly in terms of network parameters rather than indirectly through disease parameters, leading to more accurate and interpretable inference. The proposed framework thus serves as a point of departure to further explore these promising extensions.

Supplementary Materials

Supplementary information: Supplementary proofs and derivations, inference details on open population epidemics, and more results from simulation experiments and real data experiments. (.pdf file: `supplement.pdf`)

Codes and examples: R codes for all simulation experiments, accompanied by example synthetic datasets. (Github repository: *https://github.com/fanbuduke17/CoEpiNet.git*)

Funding

This work was partially supported by NIH funding: R01 EB025021, NSF DMS 2030355 and DMS 1606177, and W911NF1810233. The eX-FLU data were supported by U01 CK000185.

ORCID

Jason Xu http://orcid.org/0000-0001-5472-3720
Alexander Volfovsky https://orcid.org/0000-0003-4462-1020

References

Aiello, A. E., Simanek, A. M., Eisenberg, M. C., Walsh, A. R., Davis, B., Volz, E., Cheng, C., Rainey, J. J., Uzicanin, A., Gao, H., and Osgood, N. (2016), "Design and Methods of a Social Network Isolation Study for Reducing Respiratory Infection Transmission: The eX-FLU Cluster Randomized Trial," *Epidemics*, 15, 38–55. [1,6,13]

Anderson, R. M., and May, R. M. (1992), *Infectious Diseases of Humans: Dynamics and Control*, Oxford: Oxford University Press. [1]

Andrieu, C., Doucet, A., and Holenstein, R. (2010), "Particle Markov Chain Monte Carlo Methods," *Journal of the Royal Statistical Society*, Series B, 72, 269–342. [2]

Auranen, K., Arjas, E., Leino, T., and Takala, A. K. (2000), "Transmission of Pneumococcal Carriage in Families: A Latent Markov Process Model for Binary Longitudinal Data," *Journal of the American Statistical Association*, 95, 1044–1053. [2,4,6]

Bailey, N. T. (1975), *The Mathematical Theory of Infectious Diseases and Its Applications* (2nd ed.), Bucks: Charles Griffin & Company Ltd. [1]

Barrat, A., Cattuto, C., Tozzi, A. E., Vanhems, P., and Voirin, N. (2014), "Measuring Contact Patterns With Wearable Sensors: Methods, Data Characteristics and Applications to Data-Driven Simulations of Infectious Diseases," *Clinical Microbiology and Infection*, 20, 10–16. [1]

Becker, N. G., and Britton, T. (1999), "Statistical Studies of Infectious Disease Incidence," *Journal of the Royal Statistical Society*, Series B, 61, 287–307. [4]

Bell, D., Nicoll, A., Fukuda, K., Horby, P., and Monto, A. (2006), "World Health Organization Writing Group. Non-Pharmaceutical Interventions for Pandemic Influenza, National and Community Measures," *Emerging Infectious Diseases*, 12, 88–94. [1]

Britton, T. (2010), "Stochastic Epidemic Models: A Survey," *Mathematical Biosciences*, 225, 24–35. [2]

——— (2020), "Epidemic Models on Social Networks—With Inference," *Statistica Neerlandica*. [1]

Britton, T., and O'Neill, P. D. (2002), "Bayesian Inference for Stochastic Epidemics in Populations With Random Social Structure," *Scandinavian Journal of Statistics*, 29, 375–390. [2,15]

Cauchemez, S., and Ferguson, N. M. (2008), "Likelihood-Based Estimation of Continuous-Time Epidemic Models From Time-Series Data: Application to Measles Transmission in London," *Journal of the Royal Society Interface*, 5, 885–897. [2]

Cauchemez, S., Temime, L., Valleron, A.-J., Varon, E., Thomas, G., Guillemot, D., and Boëlle, P.-Y. (2006), "*S. pneumonia* Transmission According to Inclusion in Conjugate Vaccines: Bayesian Analysis of a Longitudinal Follow-Up in schools," *BMC Infectious Diseases*, 6, 14. [2,6,13]

Clementi, A. E., Macci, C., Monti, A., Pasquale, F., and Silvestri, R. (2010), "Flooding Time of Edge-Markovian Evolving Graphs," *SIAM Journal on Discrete Mathematics*, 24, 1694–1712. [2]

Cui, J., Zhang, Y., and Feng, Z. (2019), "Influence of Non-Homogeneous Mixing on Final Epidemic Size in a Meta-Population Model," *Journal of Biological Dynamics*, 13, 31–46. [1]

Dong, W., Pentland, A., and Heller, K. A. (2012), "Graph-Coupled HMMs for Modeling the Spread of Infection," arXiv no. 1210.4864. [2]

Eames, K., Tilston, N., White, P., Adams, E., and Edmunds, W. (2010), "The Impact of Illness and the Impact of School Closure on Social Contact Patterns," *Health Technology Assessment*, 14, 267–312. [1]

Edmunds, W. J., Kafatos, G., Wallinga, J., and Mossong, J. (2006), "Mixing Patterns and the Spread of Close-Contact Infectious Diseases," *Emerging Themes in Epidemiology*, 3, 10. [1]

Edmunds, W. J., O'Callaghan, C., and Nokes, D. (1997), "Who Mixes With Whom? A Method to Determine the Contact Patterns of Adults That May Lead to the Spread of Airborne Infections," *Proceedings of the Royal Society B: Biological Sciences*, 264, 949–957. [1]

Fan, K., Eisenberg, M., Walsh, A., Aiello, A., and Heller, K. (2015), "Hierarchical Graph-Coupled HMMs for Heterogeneous Personalized Health Data," in *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ACM, pp. 239–248. [2]

Fan, K., Li, C., and Heller, K. (2016), "A Unifying Variational Inference Framework for Hierarchical Graph-Coupled HMM With an Application to Influenza Infection," in *Thirtieth AAAI Conference on Artificial Intelligence*. [2]

Ferguson, N. M., Laydon, D., Nedjati-Gilani, G., Imai, N., Ainslie, K., Baguelin, M., Bhatia, S., Boonyasiri, A., Cucunubá, Z., Cuomo-Dannenburg, G., and Dighe, A. (2020), "Impact of Non-Pharmaceutical Interventions (NPIs) to Reduce COVID19 Mortality and Healthcare Demand," *Imperial College COVID-19 Response Team*, 10, 77482. [1]

Finkenstädt, B. F., and Grenfell, B. T. (2000), "Time Series Modelling of Childhood Diseases: A Dynamical Systems Approach," *Journal of the Royal Statistical Society*, Series C, 49, 187–205. [2]

Fintzi, J., Cui, X., Wakefield, J., and Minin, V. N. (2017), "Efficient Data Augmentation for Fitting Stochastic Epidemic Models to Prevalence Data," *Journal of Computational and Graphical Statistics*, 26, 918–929. [2,6,13]

Funk, S., Salathé, M., and Jansen, V. A. (2010), "Modelling the Influence of Human Behaviour on the Spread of Infectious Diseases: A Review," *Journal of the Royal Society Interface*, 7, 1247–1256. [1]

Geweke, J. (1991), *Evaluating the Accuracy of Sampling-Based Approaches to the Calculation of Posterior Moments* (Vol. 196), Minneapolis, MN: Federal Reserve Bank of Minneapolis, Research Department. [13]

Gibson, G. J., and Renshaw, E. (1998), "Estimating Parameters in Stochastic Compartmental Models Using Markov Chain Methods," *Mathematical Medicine and Biology*, 15, 19–40. [15]

Gillespie, D. T. (1976), "A General Method for Numerically Simulating the Stochastic Time Evolution of Coupled Chemical Reactions," *Journal of Computational Physics*, 22, 403–434. [7]

Guttorp, P., and Minin, V. N. (2018), *Stochastic Modeling of Scientific Data*, Boca Raton, FL: Chapman and Hall/CRC. [3]

He, D., Ionides, E. L., and King, A. A. (2010), "Plug-and-Play Inference for Disease Dynamics: Measles in Large and Small Populations as a Case Study," *Journal of the Royal Society Interface*, 7, 271–283. [2]

Hethcote, H. W. (2000), "The Mathematics of Infectious Diseases," *SIAM Review*, 42, 599–653. [2]

Ho, L. S. T., Crawford, F. W., and Suchard, M. A. (2018), "Direct Likelihood-Based Inference for Discretely Observed Stochastic Compartmental Models of Infectious Disease," *The Annals of Applied Statistics*, 12, 1993–2021. [6,15]

Ho, L. S. T., Xu, J., Crawford, F. W., Minin, V. N., and Suchard, M. A. (2018), "Birth/Birth-Death Processes and Their Computable Transition Probabilities With Biological Applications," *Journal of Mathematical Biology*, 76, 911–944. [2]

Hobolth, A., and Stone, E. A. (2009), "Simulation From Endpoint-Conditioned, Continuous-Time Markov Chains on a Finite State Space, With Applications to Molecular Evolution," *The Annals of Applied Statistics*, 3, 1204. [6]

Höhle, M., and Jørgensen, E. (2002), *Estimating Parameters for Stochastic Epidemics*, [The Royal Veterinary and Agricultural University], Dina. [6]

Hoti, F., Erästö, P., Leino, T., and Auranen, K. (2009), "Outbreaks of *Streptococcus pneumonia* Carriage in Day Care Cohorts in Finland—Implications for Elimination of Transmission," *BMC Infectious Diseases*, 9, 102. [2,6,13]

Ionides, E. L., Nguyen, D., Atchadé, Y., Stoev, S., and King, A. A. (2015), "Inference for Dynamic and Latent Variable Models via Iterated, Perturbed Bayes Maps," *Proceedings of the National Academy of Sciences of the United States of America*, 112, 719–724. [2]

Kermack, W. O., and McKendrick, A. G. (1927), "A Contribution to the Mathematical Theory of Epidemics," *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 115, 700–721. [1]

Kiss, I. Z., Berthouze, L., Taylor, T. J., and Simon, P. L. (2012), "Modelling Approaches for Simple Dynamic Networks and Applications to Disease Transmission Models," *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 468, 1332–1355. [2]

Kiti, M. C., Tizzoni, M., Kinyanjui, T. M., Koech, D. C., Munywoki, P. K., Meriac, M., Cappa, L., Panisson, A., Barrat, A., Cattuto, C., and Nokes, D. J. (2016), "Quantifying Social Contacts in a Household Setting of Rural Kenya Using Wearable Proximity Sensors," *EPJ Data Science*, 5, 21. [1]

Korea Centers for Disease Control and Prevention (2020), "The Updates on COVID-19 in Korea," Public Press Release, avaiable at *https://www.cdc.go.kr/board/board.es?mid=a30402000000&bid=0030*. [1]

Masuda, N., and Holme, P. (2017), *Temporal Network Epidemiology*, Singapore: Springer. [2]

Melegaro, A., Jit, M., Gay, N., Zagheni, E., and Edmunds, W. J. (2011), "What Types of Contacts Are Important for the Spread of Infections? Using Contact Survey Data to Explore European Mixing Patterns," *Epidemics*, 3, 143–151. [1]

Ministry of Health, State of Israel (2020), "Press Releases," Public Resource of Israel Case Information, available at *https://www.health.gov.il/English/News_and_Events/Spokespersons_Messages/Pages/default.aspx*. [1]

Mossong, J., Hens, N., Jit, M., Beutels, P., Auranen, K., Mikolajczyk, R., Massari, M., Salmaso, S., Tomba, G. S., Wallinga, J., and Heijne, J. (2008), "Social Contacts and Mixing Patterns Relevant to the Spread of Infectious Diseases," *PLoS Medicine*, 5, e74. [1]

Neal, P. J., and Roberts, G. O. (2004), "Statistical Inference and Model Selection for the 1861 Hagelloch Measles Epidemic," *Biostatistics*, 5, 249–261. [6]

Ogura, M., and Preciado, V. M. (2016), "Stability of Spreading Processes Over Time-Varying Large-Scale Networks," *IEEE Transactions on Network Science and Engineering*, 3, 44–57. [2]

——— (2017), "Optimal Containment of Epidemics in Temporal and Adaptive Networks," in *Temporal Network Epidemiology*, eds. N. Masuda and P. Holme, Singapore: Springer, pp. 241–266. [2]

O'Neill, P. D. (2009), "Bayesian Inference for Stochastic Multitype Epidemics in Structured Populations Using Sample Data," *Biostatistics*, 10, 779–791. [6]

Ozella, L., Gesualdo, F., Tizzoni, M., Rizzo, C., Pandolfi, E., Campagna, I., Tozzi, A. E., and Cattuto, C. (2018), "Close Encounters Between Infants and Household Members Measured Through Wearable Proximity Sensors," *PLoS One*, 13, e0198733. [1]

Pooley, C., Bishop, S., and Marion, G. (2015), "Using Model-Based Proposals for Fast Parameter Inference on Discrete State Space, Continuous-Time Markov Processes," *Journal of the Royal Society Interface*, 12, 20150225. [2]

Shaw, L. B., and Schwartz, I. B. (2008), "Fluctuating Epidemics on Adaptive Networks," *Physical Review E*, 77, 066101. [1]

Tavaré, S. (2018), "The Linear Birth–Death Process: An Inferential Retrospective," *Advances in Applied Probability*, 50, 253–269. [6]

Tsang, T. K., Fang, V. J., Ip, D. K. M., Perera, R. A. P. M., So, H. C., Leung, G. M., Peiris, J. S. M., Cowling, B. J., and Cauchemez, S. (2019), "Indirect Protection From Vaccinating Children Against Influenza in Households," *Nature Communications*, 10, 106. [6]

Tunc, I., Shkarayev, M. S., and Shaw, L. B. (2013), "Epidemics in Adaptive Social Networks With Temporary Link Deactivation," *Journal of Statistical Physics*, 151, 355–366. [2,15]

US Centers for Disease Control and Prevention (CDC) (2018), "Key Facts About Influenza (Flu)," available at *https://www.cdc.gov/flu/about/keyfacts.htm*. [14]

Van Kerckhove, K., Hens, N., Edmunds, W. J., and Eames, K. T. (2013), "The Impact of Illness on Social Networks: Implications for Transmission and Control of Influenza," *American Journal of Epidemiology*, 178, 1655–1662. [1]

Van Segbroeck, S., Santos, F. C., and Pacheco, J. M. (2010), "Adaptive Contact Networks Change Effective Disease Infectiousness and Dynamics," *PLoS Computational Biology*, 6, e1000895. [2,15]

Vanhems, P., Barrat, A., Cattuto, C., Pinton, J.-F., Khanafer, N., Régis, C., Kim, B.-A., Comte, B., and Voirin, N. (2013), "Estimating Potential Infection Transmission Routes in Hospital Wards Using Wearable Proximity Sensors," *PLoS One*, 8, e73970. [1]

Voirin, N., Payet, C., Barrat, A., Cattuto, C., Khanafer, N., Régis, C., Kim, B.-A., Comte, B., Casalegno, J.-S., Lina, B., and Vanhems, P. (2015), "Combining High-Resolution Contact Data With Virological Data to Investigate Influenza Transmission in a Tertiary Care Hospital," *Infection Control & Hospital Epidemiology*, 36, 254. [1]

Volz, E. (2008), "SIR Dynamics in Random Networks With Heterogeneous Connectivity," *Journal of Mathematical Biology*, 56, 293–310. [1]

Volz, E., and Meyers, L. A. (2007), "Susceptible–Infected–Recovered Epidemics in Dynamic Contact Networks," *Proceedings of the Royal Society B: Biological Sciences*, 274, 2925–2934. [1]

——— (2008), "Epidemic Thresholds in Dynamic Contact Networks," *Journal of the Royal Society Interface*, 6, 233–241. [1]

Wallinga, J., Edmunds, W. J., and Kretzschmar, M. (1999), "Perspective: Human Contact Patterns and the Spread of Airborne Infectious Diseases," *TRENDS in Microbiology*, 7, 372–377. [1]

WHO (2004), "Cumulative Number of Reported Probable Cases of Severe Acute Respiratory Syndrome (SARS)," available at *https://www.who.int/csr/sars/country/en/*. [15]

——— (2020), "Coronavirus Disease (COVID-2019) Situation Reports," available at *https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-report*. [15]

WHOW Group (2006), "Nonpharmaceutical Interventions for Pandemic Influenza, National and Community Measures," *Emerging Infectious Diseases*, 12, 88. [1]

Xu, J., Guttorp, P., Kato-Maeda, M., and Minin, V. N. (2015), "Likelihood-Based Inference for Discretely Observed Birth–Death-Shift Processes, With Applications to Evolution of Mobile Genetic Elements," *Biometrics*, 71, 1009–1021. [6]

Xu, J., and Minin, V. N. (2015), "Efficient Transition Probability Computation for Continuous-Time Branching Processes via Compressed Sensing," in *Uncertainty in Artificial Intelligence: Proceedings of the…Conference. Conference on Uncertainty in Artificial Intelligence* (Vol. 2015), p. 952. [6]

Yang, C. H., and Jung, H. (2020), "Topological Dynamics of the 2015 South Korea MERS-CoV Spread-on-Contact Networks," *Scientific Reports*, 10, 4327. [1]