stats

MDPI

Article

Copula Approximate Bayesian Computation Using Distribution Random Forests

George Karabatsos [1,2]

[1] Department of Mathematics, Statistics and Computer Science, University of Illinois at Chicago, 1040 W. Harrison St. (MC 147), Chicago, IL 60607, USA; georgek@uic.edu or gkarabatsos1@gmail.com
[2] Department of Educational Psychology in Statistics and Measurement, University of Illinois at Chicago, 1040 W. Harrison St. (MC 147), Chicago, IL 60607, USA

Abstract: Ongoing modern computational advancements continue to make it easier to collect increasingly large and complex datasets, which can often only be realistically analyzed using models defined by intractable likelihood functions. This *Stats* invited feature article introduces and provides an extensive simulation study of a new approximate Bayesian computation (ABC) framework for estimating the posterior distribution and the maximum likelihood estimate (MLE) of the parameters of models defined by intractable likelihoods, that unifies and extends previous ABC methods proposed separately. This framework, `copulaABCdrf`, aims to accurately estimate and describe the possibly skewed and high-dimensional posterior distribution by a novel multivariate copula-based meta-*t* distribution based on univariate marginal posterior distributions that can be accurately estimated by distribution random forests (`drf`), while performing automatic summary statistics (covariates) selection, based on robustly estimated copula dependence parameters. The `copulaABCdrf` framework also provides a novel multivariate mode estimator to perform MLE and posterior mode estimation and an optional step to perform model selection from a given set of models using posterior probabilities estimated by `drf`. The posterior distribution estimation accuracy of the ABC framework is illustrated and compared with previous standard ABC methods through several simulation studies involving low- and high-dimensional models with computable posterior distributions, which are either unimodal, skewed, or multimodal; and exponential random graph and mechanistic network models, each defined by an intractable likelihood from which it is costly to simulate large network datasets. This paper also proposes and studies a new solution to the simulation cost problem in ABC involving the posterior estimation of parameters from datasets simulated from the given model that are smaller compared to the potentially large size of the dataset being analyzed. This proposal is motivated by the fact that, for many models defined by intractable likelihoods, such as the network models when they are applied to analyze massive networks, the repeated simulation of large datasets (networks) for posterior-based parameter estimation can be too computationally costly and vastly slow down or prohibit the use of standard `ABC` methods. The `copulaABCdrf` framework and standard `ABC` methods are further illustrated through analyses of large real-life networks of sizes ranging between 28,000 and 65.6 million nodes (between 3 million and 1.8 billion edges), including a large multilayer network with weighted directed edges. The results of the simulation studies show that, in settings where the true posterior distribution is not highly multimodal, `copulaABCdrf` usually produced similar point estimates from the posterior distribution for low-dimensional parametric models as previous `ABC` methods, but the copula-based method can produce more accurate estimates from the posterior distribution for high-dimensional models, and, in both dimensionality cases, usually produced more accurate estimates of univariate marginal posterior distributions of parameters. Also, posterior estimation accuracy was usually improved when pre-selecting the important summary statistics using `drf` compared to `ABC` employing no pre-selection of the subset of important summaries. For all `ABC` methods studied, accurate estimation of a highly multimodal posterior distribution was challenging. In light of the results of all the simulation studies, this article concludes by discussing how the `copulaABCdrf` framework can be improved for future research.

check for updates

Citation: Karabatsos, G. Copula Approximate Bayesian Computation Using Distribution Random Forests. *Stats* **2024**, *7*, 1002–1050. https://doi.org/10.3390/stats7030061

Academic Editor: Wei Zhu

Received: 9 August 2024
Revised: 11 September 2024
Accepted: 13 September 2024
Published: 17 September 2024

Copyright: © 2024 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

Keywords: Bayesian analysis; maximum likelihood; intractable likelihood

1. Introduction

Statistical models defined by intractable likelihood functions are important for analyzing complex and large datasets from many scientific fields. The broad field of ABC provides alternative algorithms for estimating the approximate posterior distribution or MLE of the parameters of any model defined by a likelihood that may be intractable, either because it does not have an explicit form, because the dataset being analyzed is large, or because the model is high-dimensional (i.e., the model has many parameters).

For the model chosen to analyze a given observed dataset, represented by a set of observed data summary statistics (also referred to as observed summaries) that are ideally sufficient (at least approximately), the original rejection (vanilla) ABC algorithm [1,2], here referred to as `rejectionABC`, obtains samples from the approximate posterior distribution of the model parameters from the subsample of many samples from the prior distribution, where, for each prior sample, the summary statistics of a pseudo-dataset (or *pseudo-data summaries*) drawn from the model's exact likelihood conditionally on the prior sample of the model parameters are within a chosen small $\epsilon \geq 0$ in (e.g., Euclidean) distance to the observed data summary statistics. Then, from this subsample, estimates of marginal posterior means, medians, density and marginal densities, and estimates of the density and mode and MLE of all the parameters can be obtained. A central object of ABC is the *Reference Table*, consisting of many rows, where each table row contains a prior sample of the model parameters and the corresponding summaries of a pseudo-dataset drawn from the model's exact likelihood, conditional on this prior sample. The applicability of `rejectionABC` hinges on the ability to efficiently sample from the model's likelihood. A main motivation for `rejectionABC` is that it produces exact samples from the posterior distribution, as $\epsilon \to 0$ when the summary statistics are sufficient [3]. To explain, consider any parametric model with parameters $\theta \in \Theta \subseteq \mathbb{R}^d$ $(1 \leq d < \infty)$, a prior distribution $\pi(\theta)$ over Θ, and likelihood $f(x \mid \theta)$ for any given dataset x with summary statistics $s(x)$. Further, if s is (minimally) sufficient for x, then the posterior distribution $\pi(\theta \mid x) \propto f(x \mid \theta)\pi(\theta)$ satisfies $\pi(\theta \mid x) = \pi(\theta \mid s(x))$ for all $\pi(\theta)$ ([4] §5.1.4). Then, for any random dataset $Y \sim f(\cdot \mid \theta)$, the posterior $\pi(\theta \mid s(Y))$ converges in distribution to the exact posterior $\pi(\theta \mid x)$ as the distance δ between $s(Y)$ and $s(x)$ vanishes, and $\delta(s(Y), s(x)) = \epsilon \to 0$, as then $s(Y)$ becomes the sufficient statistics $s(x)$ of x. Therefore, when ϵ is small and, in more typical scenarios, when sufficient statistics are unavailable, the `rejectionABC` produces samples from the approximate posterior distribution and the summary statistics are approximately sufficient for x at best for the given dataset and the model applied to analyze it. Often in `rejectionABC` practice, the tolerance ϵ is chosen as a small q% (e.g., 1%) quantile of the simulated distances to define a q%-nearest-neighbor `rejectionABC` [5]. In ABC, an asymptotically efficient estimation of posterior expectations relies on the number of summary statistics to equal the number of parameters of the given model, where each parameter corresponds to a summary statistic being an approximate MLE of the model parameter [6]. Therefore, in scenarios where the number of summary statistics (summaries) exceeds the number of parameters of the given model, semiautomatic `rejectionABC` [7] is often applied, which pre-selects the summary statistics that are most important and equal in number to the number of parameters of the given model (here, referred to as `rejectionABCselect`) using any one of several available methods to select important summaries for ABC data analysis (e.g., [8]).

The original `rejectionABC` method (and variants) quintessentially represents ABC but, for the purposes of obtaining reliably accurate posterior inferences, it can be computationally inefficient or prohibitive to apply easily because it can require generating a huge number of prior samples, especially for models with more than a few parameters and/or from which it is computationally expensive to simulate pseudo-datasets from. Also, the

appropriate selection of relevant summary statistics becomes important in typical scenarios where they are not sufficient. Further, the choices of distance measure and small ϵ are also important tuning parameters for the rejection (vanilla) ABC algorithm, which can impact the level of posterior inference accuracy. Therefore, more recent research has developed various Monte Carlo algorithms that can more efficiently sample the ABC posterior distribution, with possibly less dependence on the tolerance, distance measure, and summary statistics tuning parameters of `rejectionABC`. Essentially all ABC methods can each be viewed as defining a specific approximate model likelihood providing a surrogate to the exact intractable likelihood of the given model, which measures the closeness of the summary statistics of the observed dataset to the same summaries of a random dataset simulated from the model's exact likelihood, conditionally on any proposed set of model parameters. The ABC field has had many reviews due to its wide theoretical scope and applicability (e.g., [9–17]).

The main research objective of this paper is to introduce and study a novel ABC framework, `copulaABCdrf`, for estimating the posterior distribution and MLE of the parameters of models defined by intractable likelihood functions from the given generated reference table. The `copulaABCdrf` framework unifies and extends `copulaABC` [18–21], ABC random forests (`abcrf`) [22], and ABC- and other simulation-based (`AMLE`) methods for calculating the approximate MLE for intractable likelihood models [23–30]. Notably, `copulaABCdrf` provides a single framework to perform all the tasks of inferences from a possibly high-dimensional posterior distribution (including estimation of the posterior mean, median, mode, univariate marginal posteriors), MLE estimation, automatic summary statistics selection, and model selection. In contrast, nearly all the other proposed ABC methods, typically `rejectionABC` methods, were proposed to perform only a small subset of these tasks in a non-unified manner. Therefore, a related objective of this paper is to compare the results of `copulaABCdrf` with other existing `rejectionABC` methods in terms of accuracy in estimating quantities from the posterior distribution. While some `rejectionABC` methods were already mentioned, others are reviewed later in due course.

To elaborate, `copulaABCdrf` combines `copulaABC`'s ability to approximate high-dimensional posterior distributions based on the estimation of one-dimensional marginal posterior distributions (unlike `abcrf` and `AMLE` methods) while separately estimating the dependence structure of the full multivariate posterior; with `abcrf`'s ability to estimate these one-dimensional marginal distributions (of the dependent variables) using random forests (`rf`), conditionally on observed data summaries, while automatically selecting relevant summary statistics (covariates) for the given model and dataset under analysis (unlike nearly all `copulaABC` and `AMLE` methods); and the ability to calculate the approximate MLE (unlike `abcrf` and `copula ABC`). In particular, `copulaABCdrf`, unlike `abcrf`, employs a more modern distribution random forest (`drf`) ([31] §4.1) that more directly targets the estimation of the one-dimensional marginal distributions of the model parameters from the reference table. A benefit of using `drf` for ABC is that it allows the data analyst to avoid the inefficient accept–reject procedure used in the original `rejectionABC` algorithm while not requiring the analyst to pre-select summary statistics and the tolerance and distance measure tuning parameters of `rejectionABC`. Besides `copulaABCdrf`, other methods in the literature use generative models to perform this multivariate density regression task of estimating the full joint posterior distribution without using copulas. They include autoregressive flow models [32] and conditional generative adversarial models [33], which have demonstrated to accurately estimate low-dimensional multimodal posterior distributions. Meanwhile, `copulaABC` has shown the ability to accurately estimate a skewed unimodal posterior distribution of a 250-parameter Gaussian model [18,19].

The `copulaABCdrf` framework aims to estimate (at least approximately) the posterior distribution of model parameters of possibly high dimension $d = \dim(\theta)$ using a multivariate copula-based distribution estimated from a constructed reference table. It achieves this while accounting for any skewness in the posterior and for the fact that the posterior converges to a multivariate Gaussian distribution as the data sample size increases, even

for mis-specified models, under regularity conditions [34]. For the purposes of estimating the posterior mode and MLE in practical applications, it should be easy to compute the probability density function (pdf) of and sample from such a posterior distribution estimate.

According to Sklar's (1959) [35] theorem, a continuous d-variate random parameter vector θ can be described by a joint (posterior) cumulative distribution function (cdf), called a copula, which can be uniquely represented by its univariate marginal densities (distributions) and a copula density function describing its dependence structure. If the random vector θ has any discrete-valued parameters, then these discrete random variables can be continued using jittering methods [36,37], which maintain a unique representation of the joint copula-based posterior cdf of θ while ensuring that no information is lost and preserving the same dependence relationship among the model parameters.

Approximating a possibly skewed high-dimensional posterior distribution requires a suitable choice of multivariate copula-based distribution family. Choices are limited because it is difficult to construct high-dimensional copulas ([38] p. 105). Meanwhile, skewness can be introduced into a multivariate copula-based distribution, either (1) by using skewed univariate margins (e.g., [39–41]); (2) by using a skewed copula density function to account for asymmetric dependencies in the multivariate distribution, such as a skewed, grouped Student-t, or mixture of K Gaussian or t copula density functions (e.g., [42–46]), or using a covariate-dependent copula density function [21,47,48]; or (3) by combining both skewed marginals and skewed densities, but this can unnecessarily increase the complexity of the copula model [41]. In high dimensions, estimating the parameters of skewed copulas is challenging and cumbersome [49], and, arguably, the mixture copulas are over-parameterized, with $Kd(d-1)/2$ correlation matrix parameters, while the number of correlation parameters grows quadratically with d, and it is not easy to reliably estimate K correlation matrices [50,51]. In addition, the `copulaABCdrf` framework also aims to estimate a multivariate posterior mode and the MLE as a mode of a posterior-to-prior pdf ratio, while the mode is robust to the shapes of the tails of the distribution.

Therefore, all things considered, `copulaABCdrf` employs a novel multivariate meta-t [52] copula-based distribution approximation of the posterior that allows for multivariate skewness by using skewed marginals that are flexibly modeled and estimated by `drfs`, respectively. This multivariate distribution is defined by univariate marginals, which are (covariate-dependent) `drf` regressions, along with a multivariate t parametric copula that is independent of covariates, extending previous related copula models [53,54]. The dependence structure of the meta-t distribution is defined by the density of a multivariate t distribution with (posterior) degrees of freedom and scale matrix parameters, which accounts for symmetric tail dependence and inherits the robust properties of the t distribution by inversely weighting each observation by its Mahalanobis distance from the data distribution center [55].

The `drf` targets the whole distribution of the dependent variable more appropriately than the mean functions of the dependent variable achieved by the earlier `rf` and `abcrf` algorithms. The capability of `drf` to automatically select summary statistics enables the new ABC framework to avoid the potential issues involved with pre-selecting summary statistics [56] and to avoid having to introduce an extra prior distribution to perform this selection, as performed in other ABC methods. The original `copulaABC` method [18,19] was based on a Gaussian copula density, defined by tails that are thinner than those of the more-robust t distribution (but the correlation matrix parameters of the meta-Gaussian can be robustly estimated using robust correlations, as carried out in [57]), and defined by a corresponding meta-Gaussian distribution with univariate densities estimated from the reference table by smooth kernel density estimation methods, conditionally on bandwidth parameters and summary statistics that need to be non-automatically pre-selected by the data-analyst.

Further, for scenarios where multiple models are considered to analyze the same dataset, `copulaABCdrf` employs an optional step that can be used to perform model selection based on posterior model probabilities that are estimated using `drf` (conditionally on the

observed data summaries). This provides a simple minor extension of the previously validated `rf`-based approach to model selection in ABC [58,59] by directly estimating the M-category multinomial distribution of the model index dependent variable, among the M models compared, instead of estimating M separate classification `rf`s for binary (0 or 1 valued) indices of models.

Finally, the `copulaABCdrf` framework calculates an estimate of the multivariate posterior mode and MLE of the parameters θ of the (selected) model. The MLE is the multivariate mode of the ratio of the meta-t posterior pdf estimate to the prior pdf, or the posterior mode under a proper uniform prior supporting the MLE. The new multivariate mode estimator maximizes a meta-t posterior density estimate (or finds the MLE by maximizing the ratio of the posterior density estimate to the prior density) over parameter samples of points from the reference table (or over an extra set of parameter samples of points drawn from the meta-t posterior density estimate if necessary), which avoids a possibly costly grid search involved in multivariate mode estimation by global maximization. Similarly, other "indirect" multivariate mode estimation methods [60–64] use the data sample points to estimate the mode from a nonparametric kernel, k-nearest-neighbor, or parametric density estimate, while, respectively, relying on choice of bandwidth, or k, or on a successful normal transformation of the data. The `copulaABCdrf` framework employs a novel semi-parametric approach to posterior mode and MLE estimation because it is copula-based and thus applicable to high dimensions as shown by previous `copulaABC` research [19] using Gaussian copulas and not using `rf` or summary statistics selection algorithms. Other multivariate mode estimators seem only applicable to lower dimensions (see [65] for a review). `copulaABCdrf` estimates the MLE (mode) while performing automatic summary statistics (variable) selection via `drfs` for the meta-t marginals.

Next, Section 2 describes the `copulaABCdrf` framework. Then, Section 3 investigates this framework through several simulation studies, which compare the results of `copulaABCdrf` with the results from existing `rejectionABC` methods. The first three simulation studies consider models with low-dimensional, high-dimensional, and multimodal joint posterior distributions, being either analytically or simulation-based computable posterior distributions, in order to allow for their direct comparisons with posterior distributions estimated under different ABC methods. The subsequent simulation studies and real data analyses in Section 3 focus on models defined by intractable likelihoods. Most models defined by intractable likelihoods have the two properties that (1) it is possible to rapidly simulate datasets from the model's likelihood, in which case, ABC or synthetic likelihood methods ([66,67] and references therein) can be used to estimate the (approximate) posterior distribution and MLE of the model parameters; and/or (2) the MLE or another point estimator and its sampling distribution can be computed for the model, in which case, the approximate posterior distribution of the model parameters can be computed using certain ABC or bootstrap methods (e.g., [15,68–71]).

Therefore, the rest of Section 3 focuses on models defined by intractable likelihoods that do not possess these two properties, at least in large data scenarios. They include exponential random graph models (ERGMs) and mechanistic models for network datasets, ubiquitous in scientific fields. For such a model, the likelihood is intractable for a large network dataset, and it is too costly or prohibitive to simulate large network datasets from the likelihood and to compute the MLE or other point estimator of the model parameters from large networks. Along these lines, another contribution of this paper is a new solution to the simulation cost problem in ABC, which involves constructing a reference table by simulating datasets from the exact model likelihood (given proposed model parameters, resp.) that are smaller compared to the potentially large size of the original dataset being analyzed by the model. This approach can be useful in settings where it is computationally costly to simulate data from the model, including the modeling of very large networks. Using these models, the `copulaABCdrf` framework is illustrated through the analyses of three large real-life networks, of sizes ranging between twenty thousand and over six million nodes, including a large multilayer network with weighted directed edges. In these

large data scenarios, using `copulaABCdrf` and `rejectionABC` methods, posterior inferences for these models are achieved by employing summary statistics that are invariant to the size of the network. This allows the reference table to be constructed by simulating summary statistics based on networks of a smaller tractable size compared to the size of the original observed large network dataset being analyzed, while summary statistics of the latter large network can be efficiently computed once using subsampling methods when necessary. As shown, the `copulaABCdrf` framework provides practical methods for analyzing widely available complex and large datasets. Finally, Section 4 ends with conclusions with potential future research directions based on the results of this paper.

2. Methods: The `copulaABCdrf` Framework

The proposed `copulaABCdrf` framework is summarized in Algorithm 1, shown below. The algorithm can be applied to any set of $M \geq 1$ models. It outputs an estimate of the ABC posterior distribution $\pi^{\text{ABC}}_{\widehat{m}}(\theta \mid s(x))$, a probability density function (pdf) based on corresponding cumulative distribution function (cdf) $\Pi^{\text{ABC}}_{\widehat{m}}$, the selected best predicted model, $\widehat{m} \in \{m\}_{m=1}^{M}$ among $M \geq 1$ models compared, the model's prior distribution (pdf) $\pi_{\widehat{m}}(\theta)$, and its approximate likelihood of the generic form $\int f_{\widehat{m},s}(s(x) + \varepsilon\nu \mid \theta)K(\nu)\mathrm{d}\nu$ [6] based on summary statistics vector s and some kernel density function $\mathcal{K}$.

Optional Step 2 of Algorithm 1, used when a set of $M > 1$ Bayesian models is considered by the data analyst, performs model selection based on `drf`-based [31] estimates of the M posterior model probabilities, conditionally on observed data summaries $s(x)$. As mentioned in Section 1, Step 2 provides a simple minor extension to a previously validated random forest approach to model selection in ABC [58,59] by directly targeting the full M-category multinomial distribution dependent variable over the M models being compared in the given model selection problem, instead of running M separate classification `rfs` for binary (0- or 1-valued) indices of models (resp.) as the previous random forest approach to model selection did. It has been discussed in the literature (e.g., [72]) that summary statistics can be insufficient for model comparison. Tree-based models can alleviate this issue when enough suitably chosen summary statistics are used for model selection [58].

Step 3 of Algorithm 1 aims to estimate the posterior distribution of possibly high-dimensional model parameters at least approximately well using a tractable multivariate copula-based distribution based on Sklar's (1959) theorem.

To explain Sklar's theorem in terms of Algorithm 1, let

$$\Pi^{\text{ABC}}(\theta \mid s(x)) \equiv \Pi_x(\theta_1, \ldots, \theta_d) = \Pi_x(\theta)$$

be the cdf of the approximate posterior distribution (approximate because the summaries s are not necessarily sufficient), with corresponding univariate posterior marginal cdfs $\Pi_{x,1}(\theta_1), \ldots, \Pi_{x,d}(\theta_d)$ and pdfs $\pi_{x,1}(\theta_1), \ldots, \pi_{x,d}(\theta_d)$.

Sklar's (1959) theorem implies that there exists a copula C (cdf) such that, for all $(\theta_1, \ldots, \theta_d) \in \mathbb{R}^d$, the joint posterior cdf can be represented by

$$\Pi_x(\theta_1, \ldots, \theta_d) = C\{\Pi_{x,1}(\theta_1), \ldots, \Pi_{x,d}(\theta_d)\}.$$

If the $\Pi_{x,k}$ (cdfs) are continuous for all $k = 1, \ldots, d$ (with corresponding pdfs $\pi_{x,k}$), then C is unique, with uniform distributed $\Pi_{x,k}(\theta_k) \sim \mathcal{U}_{(0,1)}$ over the random θ_k (for $k = 1, \ldots, d$) by the probability integral transform, and the copula C is 'margin-free' in the sense that it is invariant under increasing transformations of the margins ([38] Theorem 2.4.3). Then, for an arbitrary continuous multivariate distribution (e.g., [73] §13.2), its copula C (cdf) can be determined from the transformation

$$C(u_1, \ldots, u_d) = \Pi_x(\Pi^{-1}_{x,1}(u_1), \ldots, \Pi^{-1}_{x,d}(u_d)), \text{ for } u_1, \ldots, u_d \in [0,1],$$

with corresponding copula density (pdf) $c(u_1,\ldots,u_d) = \frac{\partial^d C(u_1,\ldots,u_d)}{\partial u_1,\ldots,\partial u_d}$ (for $u_1,\ldots,u_d \in [0,1]$), where the $\Pi_{x,k}^{-1}$ (for $k = 1,\ldots,d$) are inverse marginal distribution functions. Also, the posterior distribution pdf of θ is given by

$$\pi_x(\theta_1,\ldots,\theta_d) = c\{\Pi_{x,1}(\theta_1),\ldots,\Pi_{x,d}(\theta_d)\}\prod_{k=1}^{d}\pi_{x,k}(\theta_k) \text{ for } \theta_1,\ldots,\theta_d \in \mathbb{R}\cup\{\infty,-\infty\}.$$

As mentioned, the ABC framework (Algorithm 1) employs a novel multivariate meta-t [52] copula-based distribution approximation of the posterior that allows for multivariate skewness by using skewed marginals that are modeled flexibly and accurately estimated by `drfs`, respectively. This multivariate distribution has marginals that are (covariate-dependent) regressions and a multivariate t parametric copula that is independent of covariates, extending previous related copula models [53,54] through the use of `drf`. The meta-t distribution is defined by a t copula cdf $C_{\nu,\rho}$ (copula density pdf, $c_{\nu,\rho}$) with degrees of freedom ν and $d \times d$ scale matrix parameters ρ, which is a correlation matrix if the density is that of a normal distribution.

As an aside, when the random vector θ contains any discrete variables, the joint cdf C is not identifiable under Sklar's theorem (e.g., [74]). Therefore, to maintain direct use of Sklar's theorem using the t-copula, Step 1(e) of Algorithm 1 can easily apply jittering [36,37] to continue each (of any) discrete model parameter (integer-valued without loss of generality) while ensuring that no information is lost and preserving the same dependence relationship among all model parameters. Each (of any) discrete (integer) model parameter $\theta_k^* \in \theta$ with posterior cdf $\Pi_{x,k}^*(\theta_k)$ is continued ("jittered") into a continuous parameter $\theta_k = \theta_k^* - U_k$ with uniform random variable $U_k \sim \mathcal{U}_{(0,1)}$ (with $U_k \perp \theta_k^*$ and $U_k \perp U_l$ for $k \neq l$) and posterior cdf

$$\Pi_{x,k}(\theta_k) = \Pi_{x,k}^*(\lfloor\theta_k\rfloor) + (\theta_k - \lfloor\theta_k\rfloor)\Pr(\vartheta_k^* = \lfloor\theta_k + 1\rfloor)$$

and pdf $\pi_{x,k}(\theta_k) = \Pr(\vartheta_k^* = \lfloor\theta_k + 1\rfloor)$, where $\lfloor z\rfloor = \text{floor}(z)$. The parameters of $\Pi_{x,k}$ and $\pi_{x,k}$ are exactly those of $\Pi_{x,k}^*$, and θ_k^* can be recovered from θ_k as $\theta_k^* = \lfloor\theta_k + 1\rfloor$.

Step 3 of Algorithm 1 performs a two-stage semiparametric estimation [75] of the meta-t posterior distribution parameters, $\{\nu,\rho,\{\pi_k,\Pi_k\}_{k=1}^d\}$. The first stage calculates nonparametric marginal posterior cdfs and pdfs, $\{\widehat{u}_k^{(j)} = \widehat{\Pi}_{x,k}(\theta_k^{(j)}), \widehat{\pi}_{x,k}(\theta_k^{(j)})\}_{j=1}^N$, conditionally on $s(x)$, from `drfs` trained on columns of the reference table $\{\theta_k^{(j)}, s(y^{(j)})\}_{j=1}^N$ for $k = 1,\ldots,d$ (resp.), while performing automatic summary statistics (covariate) selection. The second stage employs an expectation–maximization algorithm [49,76] to calculate the MLE $(\widehat{\nu},\widehat{\rho})$ of the copula density parameters of the meta-t pdf based on these estimated cdfs and pdfs for the subset of the N rows of $\{(\widehat{u}_k^{(j)} = \widehat{\Pi}_{x,k}(\theta_k^{(j)}, \widehat{\pi}_{x,k}(\theta_k^{(j)}) \mid k = 1,\ldots,d)\}_{j=1}^N$, for which $0 < \widehat{u}_k^{(j)} < 1$ for all $k = 1,\ldots,d$, as this MLE is calculable only for this subset, which can be performed using the `nvmix` R software package [76] (R version 4.4.1 was used).

The `drf` is trained (estimated) on the reference table $\{\theta_k^{(j)}, s(y^{(j)})\}_{j=1}^N$ (the training dataset), such that prior parameter samples $\theta_k^{(j)}$ (dependent variable) are regressed on the summary statistics covariate vectors $s(y^{(j)})$ while targeting the conditional distribution of the dependent variable for each model parameter indexed by $k = 1,\ldots,d$. Then, based on each of these d trained `drfs`, the posterior distribution (cdfs $\Pi_{x,k}$ and pdfs $\pi_{x,k}$, for $k = 1,\ldots,d$) can be accurately predicted (estimated) [31], conditionally on the summary statistics of the original observed dataset, $s(x)$. All of this is carried out while performing automatic summary statistics (variable) selection from a potentially larger set of summary statistics and accounting for the uncertainty in variable selection, without requiring the user to pre-select the subset of summary statistics relevant to the intractable likelihood model under consideration.

Algorithm 1 (`copulaABCdrf`).

Estimate the posterior probability density function:
$$\pi^{\text{ABC}}_{\widehat{m}}(\theta \mid s(x)) \propto \pi_{\widehat{m}}(\theta) \int f_{\widehat{m},s}(s(x) + \varepsilon\nu \mid \theta)\mathcal{K}(\nu)\mathrm{d}\nu,$$
for the selected (best) model $\widehat{m}$ among $M \geq 1$ models considered.

Inputs: Dataset x of size n; number of iterations N_M; $M \geq 1$ models $\{m\}_{m=1}^{M}$ for dataset x; each model m with exact likelihood $f_m(\cdot \mid \theta)$, prior probability $\pi(m)$, prior distribution $\pi_m(\theta)$ for model parameters $\theta = (\theta_k)_{k=1}^{d(m)} \in \Theta^{(m)} \subseteq \mathbb{R}^{d(m)}$, candidate summary statistics $s(\cdot)$, size $n_{\text{sim}} > 0$ of data y to simulate $y \sim f_m(\cdot \mid \theta)$.

(Step 1) Construct an initial Reference Table.
for $j = 1, \ldots, N_M$ do (using parallel computing cores if necessary)
- (a) If $M > 1$, draw model index, $m(j) \sim \pi(m)$. If $M = 1$, set $m(j) \equiv \widehat{m} \equiv 1$.
- (b) Draw a parameter sample, $\theta^{(j)} \sim \pi_{m(j)}(\theta)$.
- (c) Draw a sample dataset, $y^{(j)} \sim f_{m(j)}(x \mid \theta^{(j)})$ of size n_{sim} ($\leq n$).
- (d) Calculate summary statistics, $s(y^{(j)})$.
- (e) For each discrete $\theta_k^{*(j)} \in \theta^{(j)}$, draw $U_k^{(j)} \sim \mathcal{U}_{(0,1)}$, set $\theta_k^{(m(j))} \equiv \theta_k^{*(m(j))} - U_k^{(j)}$.

end for
Output Reference Table: $\{m(j), \theta^{(j)}, s(y^{(j)})\}_{j=1}^{N_M}$; or $\{\theta^{(j)}, s(y^{(j)})\}_{j=1}^{N}$ if $M = 1$, $N_M \equiv N$.

(Step 2) If $M > 1$, train `drf` on $\{m(j), s(y^{(j)})\}_{j=1}^{N_M}$, regressing m on s (selected automatically by `drf`). Then based on trained `drf`, estimate the posterior probabilities $\widehat{\Pi}_x(m)$ of models $\{m\}_{m=1}^{M}$ conditional on $s(x)$. Select the model $\widehat{m} \in \{m\}_{m=1}^{M}$ with the highest posterior probability. Reduce Reference Table to $\widehat{m}$: $\{\theta^{(j)}, s(y^{(j)})\}_{j=1}^{N} \equiv \{\theta^{(j)}, s(y^{(j)}) : m(j) = \widehat{m}\}_{j=1}^{N}$.

(Step 3) Estimate parameters $\{\nu, \rho, \{\pi_k, \Pi_k\}_{k=1}^{d}\}$ of meta-t posterior pdf:
$$\pi^{\text{ABC}}_{\widehat{m},\nu,\rho,\pi}(\theta \mid x) = c_{\nu,\rho}(u) \prod_{k=1}^{d} \pi_{x,k}(\theta_k)$$
$$= \frac{t_{d,\nu,\rho}(T_\nu^{-1}(u_1), \ldots, T_\nu^{-1}(u_d))}{\prod_{k=1}^{d} t_\nu(T_\nu^{-1}(u_k))} \prod_{k=1}^{d} \pi_{x,k}(\theta_k), \text{ for } u \in (0,1)^d;$$
$c_{\nu,\rho}(u)$ is t-copula pdf; t_ν is pdf (T_ν cdf) of standard univariate t distribution; $t_{d,\nu,\rho}$ is the d-variate t pdf; with degrees of freedom ν and $d \times d$ scale matrix ρ; univariate marginal posterior pdfs $\pi_{x,k}(\theta_k)$ and cdfs $u_k = \Pi_{x,k}(\theta_k)$ for $k = 1, \ldots, d$. Find estimates $\{\widehat{\nu}, \widehat{\rho}, \{\widehat{\pi}_k, \widehat{\Pi}_k\}_{k=1}^{d}\}$ using 2-stage semiparametric estimation:
(1) For $k = 1, \ldots, d$, train `drf` on $\{\theta_k^{(j)}, s(y^{(j)})\}_{j=1}^{N}$, regressing θ_k on $s(y)$, with summary statistics (covariate) selection; use trained `drf` to predict estimates of posterior cdfs $\widehat{u}_k^{(j)} = \widehat{\Pi}_{x,k}(\theta_k^{(j)})$ and pdfs $\widehat{\pi}_{x,k}(\theta_k^{(j)})$ conditional on $s(x)$.
(2) Find: $(\widehat{\nu}, \widehat{\rho}) = \underset{(\nu,\rho)\in(0,\infty)\times\{\rho\}}{\arg\max} \prod_{j=1}^{N} c_{\nu,\rho}(\widehat{\Pi}_{x,1}(\theta_1^{(j)}), \ldots, \widehat{\Pi}_{x,d}(\theta_d^{(j)})) \prod_{k=1}^{d} \widehat{\pi}_{x,k}(\theta_k^{(j)})$.

(Step 4) From estimated posterior distribution $\pi^{\text{ABC}}_{\widehat{m},\widehat{\nu},\widehat{\rho},\widehat{\pi}}(\theta \mid x)$, obtain posterior: mean, variance, quantiles of θ_k from `drf` cdf estimates $\widehat{\Pi}_{x,k}$ for $k = 1, \ldots, d$; posterior scale matrix estimate $\widehat{\rho}$ of θ from Step 3; and the mode and MLE by:
$\widehat{\text{Mode}}_x(\theta) = \underset{j=1,\ldots,N}{\arg\max}\, \pi^{\text{ABC}}_{\widehat{m},\widehat{\nu},\widehat{\rho},\widehat{\pi}}(\theta^{(j)})$, MLE $\widehat{\theta} = \underset{j=1,\ldots,N}{\arg\max}\, \pi^{\text{ABC}}_{\widehat{m},\widehat{\nu},\widehat{\rho},\widehat{\pi}}(\theta^{(j)})/\pi(\theta^{(j)})$;
or by: $\widehat{\text{Mode}}_x(\theta) = \underset{j=1,\ldots,N_+}{\arg\max}\, \pi^{\text{ABC}}_{\widehat{m},\widehat{\nu},\widehat{\rho},\widehat{\pi}}(\theta_+^{(j)})$, $\widehat{\theta} = \underset{j=1,\ldots,N_+}{\arg\max}\, \pi^{\text{ABC}}_{\widehat{m},\widehat{\nu},\widehat{\rho},\widehat{\pi}}(\theta_+^{(j)})/\pi(\theta_+^{(j)})$,
given N_+ additional draws $\{\theta_+^{(j)}\}_{j=1}^{N_+} \overset{\text{iid}}{\sim} \pi^{\text{ABC}}_{\widehat{m},\widehat{\nu},\widehat{\rho},\widehat{\pi}}$.

The `drf` is a weighted nearest neighbor method [77] that performs a locally adaptive estimation of the conditional distribution through the aggregation of dependent variable predictions of an ensemble of randomized flexible classification of regression trees

(CARTs) [78], respectively, where each CART is estimated from a random subsample of the training dataset (drawn without replacement), and each level of the CART's binary tree is constructed by splitting the training data points $\{\theta_k^{(j)}, s(y^{(j)})\}_{j=1}^N$ based on a covariate $s_k(y^{(j)}) \in s(y^{(j)})$ and its split point c both chosen in such a way that the distribution of the dependent responses $\theta_k^{(j)}$ for which $s_k(y^{(j)}) \leq c$ differs the most compared to the distribution for which $s_k(y^{(j)}) > c$, according to the maximal mean discrepancy (MMD) [79]. The MMD is a quickly computable two-sample test statistic that can detect a wide variety of distributional changes. This way, the dependent variable distribution in each of the resulting leaf nodes is as homogeneous as possible in order to define neighborhoods of relevant training data points for every covariate value s. Repeating this many (B) times with randomization induces a weighting function $w_{s(x)}(s(y^{(j)}))$ quantifying the relevance of each training covariate data point $s(y^{(j)})$ (for $j = 1, \ldots, N$) for a given test point $s(x)$. Specifically, the weight $w_{s(x)}(s(y^{(j)}))$ is the proportion of times out of B CART subsampling randomizations that $s(y^{(j)})$ ends up in the same terminal leaf node as $s(x)$.

The `drf` estimates of the posterior conditional cdfs are given by:

$$\widehat{\Pi}_{x,k}(\theta_k) = \sum_{j=1}^{N} w_{s(x),k}\{s(y^{(j)})\}\mathbf{1}(\theta_k^{(j)} \leq \theta_k) \tag{1a}$$

$$= \sum_{j=1}^{N} \left\{ \frac{1}{B} \sum_{b=1}^{B} \frac{\mathbf{1}[s(y^{(j)}) \in \mathcal{L}_b\{s(x)\}]}{|\mathcal{L}_b\{s(x)\}|} \right\} \mathbf{1}(\theta_k^{(j)} \leq \theta_k), \text{ for } k = 1, \ldots, d \tag{1b}$$

which can be obtained from N (training) points of the reference table $\{\theta_k^{(j)}, s(y^{(j)})\}_{j=1}^N$, where the $w_{s(x),k}(s(y^{(j)}))$ (for each $k = 1, \ldots, d$) are positive weights with $\sum_{j=1}^N w_{s(x),k}\{s(y^{(j)})\} = 1$; B is the number of randomized CARTs (T_b for $b = 1, \ldots, B$) of the ensemble; $\mathcal{L}_b\{s(x)\}$ and $|\mathcal{L}_b\{s(x)\}|$ are, respectively, the set and number of the training data points $(\theta^{(j)}, s(y^{(j)}))$ that end up in the same leaf as $s(x)$ in the CART T_b; and $\mathbf{1}(\cdot)$ is the indicator function. The corresponding `drf` pdf estimates (for $k = 1, \ldots, d$) can be computed by the following empirical histogram-type density estimator:

$$\widehat{\pi}_{x,k}(\theta_k) = \sum_{l=1}^{N} \frac{w_{s(x),k}\{s(y_{(l)})\}}{\theta_{k,(l)} - \theta_{k,(l-1)}} \mathbf{1}(\theta_k \in (\theta_{k,(l-1)}, \theta_{k,(l)}]), \tag{2}$$

where $\theta_{k,(1)}, \ldots, \theta_{k,(N)}$ (with $\theta_{k,(0)} \equiv \theta_{k,(1)}$) are the order statistics of $\theta_k^{(1)}, \ldots, \theta_k^{(N)}$, these order statistics corresponding to weights $w_{s(x),k}\{s(y_{(1)})\}, \ldots, w_{s(x),k}\{s(y_{(N)})\}$. Alternatively, the marginal posterior density estimates $\widehat{\pi}_{x,k}(\theta_k)$ for $k = 1, \ldots, d$ can, respectfully, be obtained by smoother, univariate local polynomial Gaussian kernel density estimators (with spline interpolation to speed up computations) and automatic bandwidth selection [80], performed on the $\theta_{k,1}, \ldots, \theta_{k,N}$, using corresponding frequency weights $N \cdot w_{s(x),k}\{s(y_{(1)})\} \ldots, N \cdot w_{s(x),k}\{s(y_{(N)})\}$, and with unbounded or bounded support [81,82] depending on the spaces (resp.) of the univariate parameters or the support of their priors as appropriate, e.g., by using the `kdeld` R software package [83].

Extensive simulation studies ([31] §4.1) showed that `drf` performs well and outperforms other machine learning methods in terms of prediction accuracy for a wide range of sample sizes and problem dimensionalities, especially in problems where p (number of covariates) is large and d (dimensionality of the dependent variable) is small to moderately large, without the need for further tuning or involved numerical optimization. The `drf` training and predictions can compute quickly using the `drf` R package [84], with minimal tuning parameters. By default, a `drf` is estimated from the given training dataset using an ensemble of $B = 2000$ CARTs, with every tree constructed (randomized) from a random subset being 50% of the size of the training set, and with a target of five for the minimum number of observations in each CART tree leaf ([31] p. 33). The estimated (induced) weight-

ing function of `drf` for each of the given model parameters $\theta_k^{(j)}$ (for $k = 1, \ldots, d$) is used to estimate a parametric meta-t multivariate copula-based distribution for these dependent variable observations, after having nonparametrically adjusted for the covariates $s(y^{(j)})$ (for $j = 1, \ldots, N$) in the spirit of [85]. All `drf` computations mentioned can be undertaken using the `drf` R package [84].

Step 4 of Algorithm 1, with the d `drfs` estimated in Step 3, uses the weights $w_{s(x)}$ and marginal posterior cdf estimates (1) to calculate, for each model parameter θ_k for $k = 1, \ldots, d$, estimates of marginal posterior expectations, including estimates of the posterior mean ($\widehat{\mathbb{E}}$), variance ($\widehat{\mathbb{V}}$), and quantiles ($\widehat{Q}$):

$$\begin{aligned}
\widehat{\mathbb{E}}_{x,k}(\theta_k) &= \textstyle\sum_{j=1}^{N} \theta_k^{(j)} w_{s(x)}\{s(y^{(j)})\}, \\
\widehat{\mathbb{V}}_{x,k}(\theta_k) &= \textstyle\sum_{j=1}^{N} \{\theta_k^{(j)} - \widehat{\mathbb{E}}_{x,k}(\theta_k)\}^2 w_{s(x)}\{s(y^{(j)})\}, \\
\widehat{Q}_{x,k}(u) &= \widehat{\Pi}_{x,k}^{-1}(u), \text{ for } k = 1, \ldots, d \text{ and } u \in (0,1).
\end{aligned}$$

Step 4 also calculates estimates of the multivariate posterior mode, $\widehat{\text{Mode}}_x(\theta)$, and MLE $\widehat{\theta}$ of the parameters θ of the (selected) model. The MLE is the multivariate mode of the ratio of the meta-t posterior pdf estimate to the prior pdf, or the posterior mode under a proper uniform prior supporting the MLE. For either the task of posterior mode or MLE estimation, a novel semiparametric mode estimator is being proposed here, which is applicable to high dimensions. Other multivariate mode estimators (see [65] for a review) seem only applicable to lower dimensions. The new multivariate mode estimator maximizes a (posterior) density estimate and finds the MLE by maximizing a ratio of the posterior density estimate to the prior density over a sample of points. The sample of points is the st of parameter samples $\theta^{(j)}$ from the reference table, or, alternatively (if necessary), can be extra parameter samples $\theta_+^{(j)}$ drawn from the meta-t posterior density estimate. Such a sampling approach avoids a possibly costly grid search involved in multivariate mode estimation by global maximization. Similarly, other 'indirect' multivariate mode estimation methods [60–64] use the data sample points to estimate the mode from a nonparametric kernel, k-nearest-neighbor, or parametric density estimate, while, respectively, relying on choice of bandwidth, or k, or on a successful normal transformation of the data. The new `copulaABCdrf` framework (Algorithm 1) estimates the MLE (mode) while performing automatic summary statistics (variable) selection via `drfs` for the meta-t marginals. Therefore, this framework advances on previous simulation-based MLE estimation methods, which do not provide this automatic selection of summaries [23–30].

In particular, the original `rejectionABC` approach to MLE and posterior mode estimation [23] is based on a multivariate kernel density estimate of the posterior distribution estimated from the subset of prior parameter samples from the reference table, corresponding to simulated summary statistics being a small distance ϵ to the observed data summary statistics (and recall that ϵ can be chosen as a small q% (e.g., 1%) quantile of the simulated distances). For example, for a one- to six-dimensional parameter, a multivariate kernel density estimate can be obtained using the `kde()` function of the `ks` R package [86] based on the automatic selection of the bandwidth matrix. For a higher-dimensional parameter (of dimension d), a multivariate kernel density estimate can be obtained by a product of d univariate local polynomial Gaussian kernel density estimators ([87] eq. 6.47) based on automatic bandwidth selection [80] performed for each of these d densities, while the accuracy of the estimator deteriorates quickly as dimension d increases (e.g., [87] p. 138). Then, given this posterior kernel density estimate, an estimate of the posterior mode is obtained by maximizing this density over the q% subsamples of the reference table, giving an estimate of the posterior mode. Also, the MLE is obtained by maximizing the ratio of this posterior density to the prior density over these subsample, while, of course, the MLE coincides with the posterior mode under a uniform prior distribution. Later, we sometimes refer to the rejection-based ABC algorithm for estimating the posterior mode or MLE, based on kernel density estimation, as `rejectionABCkern`, `rejectionABCkern.select`, or `rejectionABCprodkern`,

depending on whether the rejection-based ABC algorithm used multivariate kernel density estimation (`kern`), perhaps using a pre-selection of summary statistics (kern.select), or used multivariate product kernel density estimation (`prodkern`).

3. Results of Numerical Examples

Section 3 presents results of several simulation studies that evaluate and compare the `copulaABCdrf` (Algorithm 1) and methods based on the standard (small) q% nearest-neighbor `rejectionABC` algorithm [5] (based on Euclidean distance), in terms of the estimation accuracy of various features of the posterior distribution, for three models for n multivariate independently and identically distributed (i.i.d.) observations in Section 3.1, and for seven models for n-node networks, namely, three ERGMs in Sections 3.3 and 3.7 and four mechanistic network models in Sections 3.4 and 3.5, after Section 3.2 provides a contextualizing overview of network science and modeling. To provide further illustrations of the `copulaABCdrf` and `rejectionABC` methods, these network models are applied to analyze real-life massive network datasets in Sections 3.6–3.8. Every application of the `copulaABCdrf` and `rejectionABC` method is based on a reference table of size $N = 10{,}000$ samples and the same given set of summary statistics, after considering that this size is an automatic default choice in `ABC` applications and that some summary statistics used in this study are computationally costly, especially for large datasets.

The simulation studies will in general summarize posterior estimation accuracy results of each `ABC` method by the average and standard deviation (error) of point estimates of parameters (marginal means and medians of the posterior distribution, and the posterior mode and MLE) and various fit statistics over 10 simulated data replicas from the given model generated from specified true data-generating model parameters, mentioned in the later subsections. These fit statistics, computed relative to these true data-generating model parameters, include the mean absolute error (MAE) and (less outlier-robust) mean squared error (MSE) of the above point estimates; coverage (indicator) of marginal 95% and 50% posterior credible interval estimates of individual model parameters, to be compared to their nominal values; and Kolmogorov–Smirnov (KS) distance and test statistics of the null hypotheses that the estimated univariate marginal posterior distributions (respectively) match the corresponding true exact univariate marginal posterior distributions, specifically, the weighted one-sample (two-tailed) KS distance and corresponding significance test statistic ([88] p. 358) based on the estimated `drf` weights for `copulaABCdrf` or on equal sample weights (unweighted) for `rejectionABC`. The KS test statistic has a 95th percentile critical value of 1.358 and a 99th percentile critical value of 1.628, relative to the null hypothesis that the marginal posterior distribution of the given parameter of `copulaABCdrf` (or `rejectionABC`) equals the exact marginal posterior distribution. The KS statistics can only be computed for the three models for multivariate i.i.d. observations (in Section 3.1), since they are defined by tractable (Poisson and Gaussian) likelihoods, and thus allow for computations of the exact univariate marginal posteriors, either using direct numerical computation or by using MCMC or other suitable Monte Carlo methods. In contrast, for each network model considered in this paper, the likelihood function is intractable because it is either inexplicit or not computable when the network is of a sufficiently large size n (e.g., with more than a handful of nodes, n). This makes it impossible to compute the exact marginal posterior distributions, or at least compute them in a reasonable time using Monte Carlo methods, and thus it is impossible to compute the KS statistics for these models.

Specifically, Section 3.1 reports the results of simulation studies for three models for n multivariate i.i.d. observations, namely, a joint Poisson and Gaussian mixture model defined by two parameters, having a unimodal joint posterior and exact univariate marginal posteriors, which can be directly calculated analytically; a bivariate Gaussian model, with five parameters (two location and three covariance matrix parameters) having a multimodal joint posterior distribution; and a high-dimensional multivariate Gaussian model with three hundred location parameters, with the first two parameters having a

skewed posterior. Since each of the latter two Gaussian models is defined by a Gaussian likelihood pdf, the exact univariate marginal posterior distributions of the model parameters (respectively) can be calculated using standard MCMC methods. For each of these two models, this study estimates the exact univariate marginal posterior distributions using 10,000 samples generated from a componentwise MCMC Gibbs- and slice-sampling algorithm, which routinely displayed adequate mixing and convergence to the exact univariate marginal posterior distributions according to univariate trace plots of the MCMC parameter samples. While these MCMC samples are correlated, a more efficient estimation of posterior point estimates is obtained from all 10,000 stationary MCMC samples instead of from a thinned subsample [89].

More specifically, each sampling iteration of this MCMC algorithm performed a Gibbs sampling update of the subset of mean location parameters by drawing a sample from its explicit multivariate Gaussian full conditional posterior distribution, given the data and the remaining model parameters, that is easily derivable from the standard posterior calculus of Bayesian Gaussian models under conjugate or uniform priors (e.g., [4] Appendix A) (specifically, for the bivariate Gaussian model, the Gibbs sampling update involved simple rejection sampling from a uniform prior-truncated bivariate Gaussian for both mean parameters; and, for the 300-variate Gaussian model, the sampling update remaining mean model parameters involved a direct draw from a 297-variate Gaussian with a fixed diagonal covariance matrix); and, for the remaining model parameters having no convenient known form for the full conditional posterior distribution(s), the MCMC algorithm employed a simple version of the slice sampler [90] to perform a sampling update by repeatedly sampling from a wide uniform distribution that surely supported the entire full conditional posterior density of the parameter until a sample was obtained that yielded a full conditional posterior density value that exceeded the corresponding value of the slice variable updated in this MCMC iteration (for the bivariate Gaussian model, a trivariate slice sampler was used for all the covariance matrix parameters; and, for the 300-variate Gaussian model, a univariate slice sampling update was performed for each of the first two location parameters already known to have a skewed joint posterior, while, in each update, the shrinkage procedure [90] was used to speed the search for the slice). For these two Gaussian models in particular, this componentwise MCMC Gibbs- and slice-sampling algorithm seems to provide the simplest posterior sampling algorithm while making direct use of the fundamental theorem of simulation ([91] §2.3.1) and using no tuning parameters. This is unlike alternative viable Monte Carlo methods, such as Metropolis–Hastings, Hamiltonian [92], and affine-invariant ensemble [93] sampling algorithms, which require the use of proposal covariance matrices, gradients, or other tuning parameters.

In general, for the simulation studies throughout Section 3, all the evaluations and comparisons of `copulaABCdrf` and `rejectionABC` will be based on varying conditions of the number of data simulations from the likelihood, n_{sim}, relative to the total sample size n, with $n_{\text{sim}} < n$, where, for the network models, n is the number of nodes of a network. When constructing a reference table for ABC, the strategy of simulating summary statistics based on simulating datasets of size $n_{\text{sim}} < n$ can potentially be useful in situations where a large size n dataset is being analyzed and/or simulating from the given model is computationally costly. In such large (n) network scenarios, it is prohibitively costly and practically infeasible to simulate many networks of the same size (number of nodes) as the original network dataset over iterations in an ABC or synthetic likelihood algorithm, especially considering that it can already be prohibitively costly to simulate a single network from an ERGM or mechanistic network models (e.g., an ERGM given parameters simulates a network using MCMC) for a network of sufficiently large size (number of nodes), let alone to compute point-estimates (e.g., MLE) of the parameters of the given network model. These issues, in the context of Algorithm 1, are addressed by a strategy that simulates network datasets of a smaller size compared to the size of the network dataset under statistical analysis while using network summary statistics (calculated on the observed dataset and each simulated dataset) that take on values that have the same meaning and are

invariant to the size of the network dataset(s) being analyzed, including maximum pseudo likelihood estimates (MPLEs) of ERGM parameters based on network size offsets [94]. These network size invariant summary statistics are reviewed in Section 3.2. The R packages `ergm` [95], `ergm.count` [96], and `igraph` [97] were used to compute network summary statistics MPLEs, and Monte Carlo MLEs (MCMLEs) [98] of ERGMs, with the MCMLE being a standard commonly used MCMC approximation of the ERGM MLE. These MLEs will be compared with the MLEs of `copulaABCdrf` and `rejectionABC` by MAE and MSE. Note that some values of the ERGM parameters can concentrate probability mass on degenerate or near-degenerate networks that concentrate probability mass on a small subset of all possible network graphs with almost or exactly zero edges or almost or all possible edges among the n nodes [99,100], which can lead to infinite values of MPLEs and MLEs for one or more ERGM parameters. Therefore, for the network modeling applications of Sections 3.3–3.8, where MPLE summary statistics are used, the results of the `copulaABCdrf` and `rejectionABC` methods will only be based on the subset of the rows of the reference table corresponding to finite-valued MPLE summary statistics. This provided a natural way to address this degeneracy, at least for the models and summaries used in these sections.

For each of the multivariate i.i.d. models, ERGMs, and one of the mechanistic (DMC) network models investigated in the simulation studies of Sections 3.1, 3.3, 3.5 and 3.7, the number of summary statistics (in s) equals the number of model parameters (i.e., their dimensionality is equal, $\dim(s) = \dim(\theta)$). For each of the three other (Price, NLPA, and DMR) mechanistic network models investigated in the simulation studies of Sections 3.4 and 3.5, the initial number of summary statistics exceeds the number of model parameters (i.e., $\dim(s) \geq \dim(\theta)$). The asymptotic efficiency of posterior expectations in `ABC` relies on $\dim(s) = \dim(\theta)$ [6], as mentioned in Section 1. Therefore, in the simulation scenarios of Sections 3.4 and 3.5, where the number of summary statistics exceeds the number of model parameters, we will also consider a semi-automatic approach to `rejectionABC` based on pre-selecting the $\dim(\theta)$ most important summary statistics (predictor variables), determined by training a `drf` regression of the prior samples of $\{\theta\}_{i=1}^{N}$ on the corresponding samples of the summary statistics $\{s\}_{i=1}^{N}$ on the reference table, with the importance of each summary statistic (variable) efficiently calculated by a simple weighted sum of how many times each summary statistic was split on at each depth in the forest. This semiautomatic ABC method based on the pre-selection of summaries using `drf`, called `rejectionABCselect`, will be compared with the results obtained by `rejectionABC` using all the available summary statistics (i.e., without any pre-selection) and the results obtained by `copulaABCdrf`.

Preliminary results of simulation studies for the q% nearest-neighbor `rejectionABC` algorithm, for all models studied in Sections 3.1, 3.3–3.5 and 3.7, showed that the 1% nearest-neighbor `rejectionABC` algorithm tended to perform best in terms of the MAE, MSE, and KS statistics compared to 2% and 3% nearest neighbors. Therefore, for space considerations, only the results for 1% nearest-neighbor `rejectionABC` will be shown throughout Section 3.

For the `copulaABCdrf` method, it was found through preliminary simulation studies of all eight models considered in Sections 3.1 and 3.3–3.5 that, compared to the smooth univariate kernel density estimators (mentioned in Section 2), the empirical histogram density estimator (2) had a slightly higher tendency for producing superior posterior mode and MLE estimates based on the meta-t posterior, for most models considered in the simulation study, according to the MAE and MSE of each individual model parameter, when the dimension $\dim(\theta)$ of the given model parameters θ was less than six. The empirical density estimator also has the advantage of not using a bandwidth parameter, while the choice of bandwidth is an important and non-trivial aspect for the accuracy of smooth kernel density estimation. However, it was found that, when the given model parameter θ has a sufficiently high dimension, the product term $\widehat{\pi}_{x,k}(\theta_k^{(j)})$ (for $j = 1, \ldots, N$) of the meta-t density estimate (see Step 4 of Algorithm 1), based on the empirical density estimator, can become zero for a large majority of the N samples of $\{\theta^{(j)}\}_{i=1}^{N}$ in the reference table, thereby

requiring an extremely large sample N (e.g., well above, say, 10,000) to overcome this issue. Therefore, in such high-dimensional parameter settings, the kernel density estimator has the advantage of reducing the frequencies of zero from the reference table because it provides a smoother density estimator and therefore does not require an extremely large sample for the reference table. Also, in a related issue, according to some preliminary simulation studies carried out as a pre-cursor to the simulation studies reported later in Sections 3.1 and 3.3–3.5, it turned out that the computation of the MLE $(\widehat{\nu}, \widehat{\rho})$ of the t-copula density parameters (in Step 4 of Algorithm 1) based on multiplying (re-scaling) the $\widehat{u}_k^{(j)}$ by $N/(N+1)$, carried out to avoid evaluating the copula density at the edges of the unit hypercube (as advocated by [75,101] and others), was not obviously advantageous in terms of the MAE and MSE accuracy of posterior mean, median, mode, and MLE estimation. Therefore, Section 3 throughout presents the results of the simulation studies based on Step 3 of the `copulaABCdrf` Algorithm 1, as described in Section 2.

Further, recall that both of `copulaABCdrf`'s alternative approaches to posterior mode and MLE estimation are presented in Step 4 of Algorithm 1. It was found through additional preliminary simulation studies `copulaABCdrf` over all eight models considered in Sections 3.1 and 3.3–3.5 that, in terms of the posterior mode and MLE estimation accuracy of `copulaABCdrf`, measured by MAE and MSE, the approach of drawing an extra set of parameter samples from the meta-t posterior density estimate can be advantageous (compared to the first approach based on the original reference table), but only when the true posterior distribution is high-dimensional and symmetric or skewed (but not highly multimodal). This seems to be reflective of the geometry of the meta-t density (distribution). Therefore, for space considerations, we primarily present the results of the posterior mode and MLE estimators of `copulaABCdrf`, which is only based on the samples generated in the original reference table. We only focus on the extra-sample approach to mode and MLE estimation for a high-dimensional setting involving a skewed posterior distribution and based on the smooth univariate kernel density estimators of the univariate posterior marginals.

3.1. Simulation Study: Calculable Exact Posterior Distributions

Here, we first study the accuracy of Algorithm 1 to estimate the directly calculable true posterior distribution of a joint bivariate Poisson(θ) model and a two-component scale normal mixture model with common mean parameter μ, which has been studied in previous research on ABC methods. The Poisson(λ) distribution is defined by a tractable likelihood probability mass function (p.m.f.) $f(X_1 = x_1 \mid \lambda) = \lambda^{x_1} \exp(-\lambda)/x_1!$, assigned a gamma prior distribution $\mathcal{G}(\lambda \mid \frac{1}{2}, 0.1)$ (shape $\frac{1}{2}$ and rate 0.1), and exact gamma posterior distribution $\lambda \mid x_1 \sim \mathcal{G}(\lambda \mid \frac{1}{2} + \sum_{i=1}^{n} x_{i,1}, 0.1 + n)$, with the sample mean $\overline{x}_1 = \frac{1}{n}\sum_{i=1}^{n} x_{i,1}$ a sufficient summary statistic for λ. The scale normal (Gaussian) mixture model $f(X_2 = x_2 \mid \mu) = \frac{1}{2}\mathcal{N}(x_2 \mid \mu, 1) + \frac{1}{2}\mathcal{N}(x_2 \mid \mu, 0.01)$, and the Gaussian distribution $\mathcal{N}(x_2 \mid \mu, z + (1 - z)0.01)$ with latent covariate $z \sim$Bernoulli$(1/2)$ is assigned a uniform prior distribution $\mu \sim \mathcal{U}_{(-10,10)}$, yielding the posterior distribution $\pi(\mu \mid x_2) \propto \frac{1}{2}\mathcal{N}(\mu \mid \overline{x}_2, 1) + \frac{1}{2}\mathcal{N}(\mu \mid \overline{x}_2, 0.01)$ (effectively, with equality up to machine precision) with the summary statistic the sample mean $\overline{x}_2 = \frac{1}{n}\sum_{i=1}^{n} x_{i,2}$. The Poisson and scale normal (Gaussian) mixture distributions, together, define a bivariate distribution model, respectively, with parameters assigned the aforementioned independent prior distributions and posterior distributions as marginal distributions of $\theta = (\lambda, \mu)$. This enables Algorithm 1 to estimate a bivariate posterior mode and MLE of θ.

Ten replications of $n = 100$ bivariate observations $\{(x_{i,1}, x_{i,2})\}_{i=1}^{n=100}$ from this bivariate model were simulated from the bivariate distribution model with true data-generating parameters $\theta = (\lambda = 3, \mu = 0)$ for each of eight conditions defined by sizes of $n_{\text{sim}} = 10, 25, 33, 50, 66, 75, 90$, and 100 simulations from the exact likelihood of the model.

Table 1 presents detailed representative results for some $n_{\text{sim}} = 10, 50$ and 100 while Tables 2 and 3 present the results for all the eight conditions of n_{sim}. Overall, `copulaABCdrf` outperformed `rejectionABC` in terms of MAE, MAE, and KS statistics for each of the posterior distribution and mean, median, and mode estimation. The `rejectionABC` method

outperformed `copulaABCdrf` only for the posterior mean estimation of the mean location parameter μ of the normal mixture, especially for smaller $n_{\text{sim}} = 10$ and 25. Also, `copulaABCdrf` and `rejectionABC` performed similarly and rather close to the nominal 95% posterior credible intervals while `copulaABCdrf` performed better than `rejectionABC` in producing 50% posterior credible intervals while tending to produce interval estimates that were near the nominal values. Both `copulaABCdrf` and `rejectionABC` always produced significant KS statistics, i.e., estimates of the marginal posterior distributions of model parameters that significantly departed from the exact posterior distribution. Therefore, the univariate `drf` regressions trained on the reference table of 10,000 samples were unable to very accurately estimate these marginal posterior distributions, but they estimated marginal posteriors that were more accurate than those obtained from `rejectionABC`.

Finally, for both ABC methods, the accuracy of estimation from the posterior distribution tended to improve as n_{sim} increased toward the full dataset sample size n, more so for the rate parameter λ. This suggests that the approach of simulating datasets from the given model of size n_{sim} smaller (but not too small or zero) relative to the size n of the observed dataset being analyzed can be a reasonable approach for accurate estimation of the model's posterior distribution, using either ABC method. Intuitively, the sufficient statistic $\overline{x}_1$ for the Poisson distribution parameter λ is sufficient for a dataset of size $n_{\text{sim}} = 100$ and only approximately sufficient for smaller simulation sample sizes n_{sim}. A similar case applies for the statistic $\overline{x}_2$ summarizing the parameter μ of the scale normal mixture model.

Table 1. Poisson(λ)-NormalMixture(μ) model: mean (standard deviation) of estimators of posterior distribution, and mean 95%(50%) credible interval coverage (95%(50%)c) over 10 replicas (n = 100), under three conditions of n_{sim}.

	λ	μ	λ	μ	λ	μ
n	100	100	100	100	100	100
n_{sim}	100	100	50	50	10	10
truth	3	0	3	0	3	0
Exact mean	3.02 (0.15)	0.01 (0.10)	3.08 (0.17)	−0.02 (0.09)	2.99 (0.11)	−0.01 (0.11)
`copulaABCdrf` mean	**3.02** (0.15)	**0.01** (0.10)	3.09 (0.17)	**−0.02** (0.10)	**3.05** (0.15)	−0.03 (0.12)
`rejectionABC` mean	2.94 (0.16)	0.02 (0.10)	**3.03** (0.17)	−0.03 (0.10)	**2.95** (0.14)	**−0.02** (0.10)
Exact median	3.02 (0.15)	0.01 (0.10)	3.07 (0.17)	−0.02 (0.09)	2.99 (0.11)	−0.01 (0.11)
`copulaABCdrf` median	**3.02** (0.15)	**0.01** (0.10)	3.09 (0.16)	**−0.02** (0.09)	**3.03** (0.15)	−0.03 (0.13)
`rejectionABC` median	2.91 (0.18)	0.02 (0.11)	**3.03** (0.17)	−0.05 (0.10)	2.91 (0.12)	**−0.02** (0.09)
Exact mode	3.01 (0.15)	0.01 (0.10)	3.07 (0.17)	−0.02 (0.09)	2.98 (0.11)	−0.01 (0.11)
`copulaABCdrf` mode	**3.06** (0.24)	−0.14 (0.44)	**3.07** (0.27)	−0.16 (0.27)	3.23 (0.60)	−0.17 (0.17)
`rejectionABCkern` mode	2.81 (0.34)	**0.07** (0.33)	2.88 (0.49)	**−0.14** (0.30)	**2.90** (0.24)	**0.02** (0.22)
Exact MLE	3.02 (0.15)	0.01 (0.10)	3.08 (0.17)	−0.02 (0.09)	2.99 (0.11)	−0.01 (0.11)
`copulaABCdrf` MLE	3.06 (0.24)	−0.14 (0.44)	**3.07** (0.27)	−0.16 (0.27)	3.23 (0.60)	−0.17 (0.17)
`rejectionABCkern` MLE	**3.00** (0.29)	**0.08** (0.32)	3.20 (0.41)	**0.00** (0.35)	**3.17** (0.28)	**−0.02** (0.15)
Exact standard deviation (s.d.)	0.17 (0.00)	0.71 (0.00)	0.18 (0.00)	0.71 (0.00)	0.17 (0.00)	0.71 (0.00)
`copulaABCdrf` s.d.	**0.20** (0.01)	0.22 (0.07)	**0.26** (0.02)	0.24 (0.08)	**0.57** (0.04)	0.34 (0.06)
`rejectionABC` s.d.	0.49 (0.05)	**0.46** (0.03)	0.54 (0.03)	**0.48** (0.03)	0.73 (0.06)	**0.51** (0.03)
Exact 95%(50%)c	1.00 (0.60)	1.00 (1.00)	0.90 (0.50)	1.00 (0.90)	1.00 (0.60)	1.00 (0.80)
`copulaABCdrf`95%(50%)c	**1.00** (**0.70**)	**1.00** (**0.20**)	**1.00** (**0.70**)	**1.00** (**0.80**)	**1.00** (**1.00**)	**1.00** (**0.90**)
`rejectionABC` 95%(50%)c	**1.00** (1.00)	**1.00** (1.00)	**1.00** (1.00)	**1.00** (1.00)	**1.00** (1.00)	**1.00** (1.00)
`copulaABCdrf` KS distance	**0.09** (0.02)	**0.20** (0.01)	**0.14** (0.03)	**0.18** (0.01)	**0.31** (0.03)	**0.15** (0.02)
`rejectionABC` KS distance	0.44 (0.03)	0.22 (0.02)	0.44 (0.03)	0.23 (0.02)	0.44 (0.02)	0.22 (0.01)
`copulaABCdrf` KS test statistic	**9.40** (2.20)	**20.97** (1.78)	**15.31** (3.62)	**19.10** (1.51)	**31.83** (3.10)	**16.46** (1.84)
`rejectionABC` KS test statistic	132.62 (8.73)	65.69 (4.97)	130.74 (9.41)	68.65 (6.04)	133.49 (6.56)	66.09 (3.93)
`copulaABCdrf` copula d.f. and scale	d.f.: 7.89 (7.27)	scale: 0.19 (0.42)	d.f.: 10.74 (8.60)	scale: 0.02 (0.40)	d.f.: 9.29 (7.98)	scale: 0.10 (0.29)

Note: **Bold** indicates the more accurate ABC method for the given estimator.

Table 2. MAE and MSE for the Poisson–normal mixture model over 10 replicas (n=100) under varying conditions of n_{sim}.

		MAE: `copulaABCdrf`, `rejectionABC` (exact)		**MSE:** `copulaABCdrf`, `rejectionABC` (exact)	
n_{sim}	Posterior	λ	μ	λ	μ
10	Mean	**0.13**, **0.13** (0.09)	0.10, **0.08** (0.10)	**0.02**, **0.02** (0.01)	**0.01**, **0.01** (0.01)
	Median	**0.12**, 0.13 (0.09)	0.11, **0.06** (0.10)	**0.02**, **0.02** (0.01)	0.02, **0.01** (0.01)
	Mode	0.47, **0.21** (0.09)	0.21, **0.17** (0.10)	0.38, **0.06** (0.01)	0.06, **0.04** (0.01)
	MLE	0.47, **0.27** (0.09)	0.21, **0.12** (0.10)	0.38, 0.10 (0.01)	0.06, **0.02** (0.01)
25	Mean	**0.17**, **0.17** (0.15)	0.07, **0.04** (0.07)	**0.04**, **0.04** (0.03)	0.01, **0.00** (0.01)
	Median	**0.16**, 0.17 (0.15)	0.07, **0.05** (0.07)	**0.04**, **0.04** (0.03)	0.01, **0.00** (0.01)
	Mode	0.40, **0.27** (0.15)	**0.19**, 0.31 (0.07)	0.19, **0.11** (0.03)	**0.05**, 0.13 (0.01)
	MLE	0.40, **0.32** (0.15)	**0.19**, 0.27 (0.07)	0.19, **0.12** (0.03)	**0.05**, 0.11 (0.01)
33	Mean	**0.09**, 0.13 (0.09)	**0.06**, 0.09 (0.07)	**0.02**, **0.02** (0.02)	**0.00**, 0.01 (0.01)
	Median	**0.10**, 0.16 (0.09)	**0.06**, 0.09 (0.07)	**0.02**, 0.03 (0.02)	**0.00**, 0.01 (0.01)
	Mode	**0.21**, 0.34 (0.09)	**0.16**, 0.34 (0.07)	**0.07**, 0.15 (0.02)	**0.08**, 0.13 (0.01)
	MLE	**0.21**, 0.25 (0.09)	**0.16**, 0.33 (0.07)	**0.07**, 0.10 (0.02)	**0.08**, 0.12 (0.01)
50	Mean	**0.14**, **0.14** (0.15)	**0.07**, 0.08 (0.06)	0.04, **0.03** (0.03)	**0.01**, **0.01** (0.01)
	Median	0.14, **0.13** (0.15)	**0.07**, 0.09 (0.06)	**0.03**, **0.03** (0.03)	**0.01**, **0.01** (0.01)
	Mode	**0.17**, 0.42 (0.15)	**0.20**, 0.26 (0.06)	**0.07**, 0.23 (0.03)	**0.09**, 0.10 (0.01)
	MLE	**0.17**, 0.37 (0.15)	**0.20**, 0.28 (0.06)	**0.07**, 0.19 (0.03)	**0.09**, 0.11 (0.01)
66	Mean	**0.12**, **0.12** (0.12)	0.07, **0.06** (0.06)	**0.02**, **0.02** (0.02)	**0.01**, **0.01** (0.00)
	Median	**0.12**, **0.12** (0.12)	**0.06**, 0.08 (0.06)	**0.02**, **0.02** (0.02)	**0.01**, **0.01** (0.00)
	Mode	0.30, **0.25** (0.12)	**0.21**, 0.28 (0.06)	0.17, **0.09** (0.02)	**0.10**, **0.10** (0.00)
	MLE	0.24, **0.22** (0.12)	**0.14**, 0.32 (0.06)	0.11, **0.07** (0.02)	**0.04**, 0.13 (0.00)
75	Mean	**0.14**, 0.17 (0.15)	**0.05**, **0.05** (0.05)	**0.03**, 0.04 (0.03)	**0.00**, **0.00** (0.00)
	Median	**0.14**, 0.18 (0.15)	**0.05**, **0.05** (0.05)	**0.03**, 0.05 (0.03)	**0.00**, **0.00** (0.00)
	Mode	**0.27**, 0.29 (0.16)	**0.15**, 0.29 (0.05)	**0.09**, 0.12 (0.03)	**0.03**, 0.12 (0.00)
	MLE	**0.28**, **0.28** (0.15)	**0.14**, 0.28 (0.05)	**0.09**, 0.11 (0.03)	**0.03**, 0.12 (0.00)
90	Mean	**0.10**, **0.10** (0.10)	0.06, **0.05** (0.05)	**0.01**, **0.01** (0.01)	**0.00**, **0.00** (0.00)
	Median	**0.09**, 0.12 (0.10)	**0.06**, 0.08 (0.05)	**0.01**, 0.02 (0.01)	**0.00**, 0.01 (0.00)
	Mode	**0.21**, 0.24 (0.10)	**0.12**, 0.19 (0.05)	**0.06**, 0.08 (0.01)	**0.02**, 0.05 (0.00)
	MLE	**0.21**, 0.26 (0.10)	**0.12**, 0.26 (0.05)	**0.06**, 0.08 (0.01)	**0.02**, 0.10 (0.00)
100	Mean	**0.10**, 0.14 (0.11)	0.09, **0.07** (0.08)	**0.02**, 0.03 (0.02)	**0.01**, **0.01** (0.01)
	Median	**0.11**, 0.15 (0.11)	**0.08**, **0.08** (0.08)	**0.02**, 0.04 (0.02)	**0.01**, **0.01** (0.01)
	Mode	**0.20**, 0.30 (0.11)	0.30, **0.26** (0.08)	**0.06**, 0.14 (0.02)	0.19, **0.11** (0.01)
	MLE	**0.20**, 0.23 (0.11)	0.30, **0.26** (0.08)	**0.06**, 0.08 (0.02)	0.19, **0.10** (0.01)

Note: **Bold** indicates the more accurate ABC method for the given estimator.

Table 3. Mean (s.d.) KS distance [test statistic] and mean interval coverage for the Poisson–normal mixture model over 10 replicas ($n = 100$) under varying n_{sim}.

		Mean KS Distance [Test] (s.d.)		**Mean 95% (50%) Coverage**	
n_{sim}	**Method**	λ	μ	λ	μ
10	`copulaABCdrf`	**0.31** (0.03)	**0.15** (0.02)	**1.00** (**1.00**)	**1.00** (**1.00**)
	`rejectionABC`	0.37 (0.05)	0.20 (0.04)	**1.00** (**1.00**)	**1.00** (**1.00**)
	`copulaABCdrf`	[**31.83** (3.10)]	[**16.46** (1.84)]	exact:	exact:
	`rejectionABC`	[36.79 (4.64)]	[19.56 (4.04)]	1.00 (0.60)	1.00 (0.80)
25	`copulaABCdrf`	**0.22** (0.03)	**0.17** (0.01)	**1.00** (**0.80**)	**1.00** (**0.70**)
	`rejectionABC`	0.37 (0.04)	0.20 (0.03)	**1.00** (1.00)	**1.00** (1.00)
	`copulaABCdrf`	[**23.03** (2.45)]	[**17.85** (1.74)]	exact:	exact:
	`rejectionABC`	[36.89 (3.73)]	[20.24 (2.97)]	0.90 (0.40)	1.00 (1.00)

Table 3. *Cont.*

		Mean KS Distance [Test] (s.d.)		Mean 95% (50%) Coverage	
n_{sim}	Method	λ	μ	λ	μ
33	`copulaABCdrf`	**0.20** (0.04)	**0.17** (0.01)	**1.00** (**0.90**)	**1.00** (**0.90**)
	`rejectionABC`	0.34 (0.05)	0.20 (0.03)	**1.00** (**0.90**)	**1.00** (1.00)
	`copulaABCdrf`	[**21.06** (4.95)]	[**18.70** (1.19)]	exact:	exact:
	`rejectionABC`	[34.38 (4.83)]	[19.61 (3.34)]	1.00 (0.70)	1.00 (1.00)
50	`copulaABCdrf`	**0.14** (0.03)	**0.18** (0.01)	**1.00** (**0.70**)	**1.00** (**0.80**)
	`rejectionABC`	0.33 (0.04)	0.20 (0.04)	**1.00** (1.00)	**1.00** (1.00)
	`copulaABCdrf`	[**15.31** (3.62)]	[**19.10** (1.51)]	exact:	exact:
	`rejectionABC`	[32.83 (3.5)]	[20.35 (3.88)]	0.90 (0.50)	1.00 (0.90)
66	`copulaABCdrf`	**0.12** (0.03)	**0.19** (0.01)	**1.00** (**0.80**)	**1.00** (**0.70**)
	`rejectionABC`	0.34 (0.05)	0.20 (0.03)	**1.00** (1.00)	**1.00** (1.00)
	`copulaABCdrf`	[**13.02** (2.81)]	[**19.98** (1.32)]	exact:	exact:
	`rejectionABC`	[33.84 (5.43)]	[20.35 (2.94)]	1.00 (0.50)	1.00 (1.00)
75	`copulaABCdrf`	**0.11** (0.03)	0.19 (0.01)	**1.00** (**0.60**)	**1.00** (**0.60**)
	`rejectionABC`	0.35 (0.04)	**0.18** (0.03)	**1.00** (0.90)	**1.00** (1.00)
	`copulaABCdrf`	[**11.61** (3.71)]	[20.17 (1.47)]	exact:	exact:
	`rejectionABC`	[34.95 (4.29)]	[**17.79** (3.02)]	1.00 (0.40)	1.00 (1.00)
90	`copulaABCdrf`	**0.10** (0.03)	0.21 (0.01)	**1.00** (**0.80**)	**1.00** (**0.50**)
	`rejectionABC`	0.34 (0.04)	**0.20** (0.05)	**1.00** (1.00)	**1.00** (1.00)
	`copulaABCdrf`	[**10.64** (2.92)]	[20.93 (1.56)]	exact:	exact:
	`rejectionABC`	[34.12 (4.32)]	[**20.39** (5.06)]	1.00 (0.60)	1.00 (1.00)
100	`copulaABCdrf`	**0.09** (0.02)	0.20 (0.01)	**1.00** (**0.70**)	**1.00** (**0.20**)
	`rejectionABC`	0.36 (0.05)	**0.18** (0.04)	**1.00** (1.00)	**1.00** (1.00)
	`copulaABCdrf`	[**9.40** (2.20)]	[20.96 (1.78)]	exact:	exact:
	`rejectionABC`	[35.63 (5.41)]	[**18.35** (3.88)]	1.00 (0.60)	1.00 (1.00)

Note: **Bold** indicates the more accurate ABC method for the given estimator.

Next, for the simulation study, we consider a simple five-dimensional Gaussian model for $n = 4$ bivariate observations. While this model is simple, it nevertheless gives rise to a calculable multimodal posterior distribution that is challenging to estimate using traditional (e.g., rejection) ABC methods [32,33]. In particular, for the given set of four data observations $\{(x_{i1}, x_{i2})\}_{i=1}^{n=4}$, the model is defined by likelihood $\prod_{i=1}^{n=4} \mathcal{N}_2(x_i \mid \mu_\theta, \Sigma_\theta)$, with pdf $\mathcal{N}_2(x \mid \mu, \Sigma) = (2\pi)^{-1}\det(\Sigma)^{-1/2}\exp[-\frac{1}{2}(x-\mu)^\intercal\Sigma^{-1}(x-\mu)]$ based on means $\mu = (\mu_1, \mu_2)^\intercal$ and 2×2 covariance matrix Σ, and with mean parameters $\mu_\theta = (\theta_1, \theta_2)$ and covariance matrix parameters $\Sigma_\theta = \begin{pmatrix} s_1^2 & \rho s_1 s_2 \\ \rho s_1 s_2 & s_2^2 \end{pmatrix}$, with $s_1 = \theta_3^2$, $s_2 = \theta_4^2$, and $\rho = \tanh(\theta_5)$, assigned uniform prior distributions $\theta_1 \sim \mathcal{U}_{(-3,3)}$, $\theta_2 \sim \mathcal{U}_{(-4,4)}$, $\theta_3 \sim \mathcal{U}_{(-3,3)}$, $\theta_4 \sim \mathcal{U}_{(-3,3)}$, and $\theta_5 \sim \mathcal{U}_{(-3,3)}$ as in [33]. Approximating the posterior distribution of this model can be challenging since this posterior is complex and non-trivial because the signs of the parameters θ_3 and θ_4 are non-identifiable and thus yield four symmetric modes (due to squaring), and because the uniform prior distributions induce cutoffs in this posterior distribution. For the simulation study for this Gaussian model, ten replicas of $n = 4$ bivariate observations $x = (x_1, x_2)^\intercal$ were simulated using true parameters $\theta_0 = (-0.7, -2.9, -1.0, -0.9, 0.6)^\intercal$ as in [33]. For analyzing data simulated from this Gaussian model using `copulaABCdrf` and `rejectionABC`, five summary statistics were used, being the MLEs of the mean and variance of the two variables and of the covariance matrix parameter. The exact posterior distribution of this Gaussian model was estimated by 10,000 samples generated from the simple componentwise MCMC Gibbs- and slice-sampling algorithm mentioned earlier.

Table 4 presents the results of the simulation study for this Gaussian model with multimodal posterior distribution. Figure 1 shows trace plots of samples of the univariate marginals of 10,000 samples of the exact multimodal posterior distribution, generated by the MCMC Gibbs and slice sampler, conditionally on the first of the 10 data replications, for illustration. This trace plot shows the typical result of this MCMC sampling algorithm that was obtained for each Gaussian model studied in this subsection and given each data replica, namely, that this MCMC sampler produced samples that quickly mixed and converged to the target posterior distribution. Table 4 shows that, on average, and according to MAE and MSE, the marginal posterior mean and median estimation of the true data-generating values of the parameters θ_1, θ_2, and θ_3, `copulaABCdrf` produced marginal univariate means and medians that were on average nearer and closer to their respective (MCMC-estimated) exact values compared to 1% nearest-neighbor rejection ABC. However, for the estimation of the marginal posterior mean and median of the true data-generating values of the parameters θ_4, θ_5, the marginal posterior mean and median estimates of rejection ABC were closer than those of `copulaABCdrf` while these estimates of both methods noticeably departed from the true data-generating values. Also, the posterior mode (and MLE) estimates tended to be more accurate on average for rejection ABC based on five-dimensional kernel density estimation compared to `copulaABCdrf`. The marginal standard deviations tended to be closer to their (estimated) exact values for `copulaABCdrf` compared to `rejectionABC`. Both ABC methods produced mean marginal 95% and 50% coverage of the true data-generating parameters that were similar to those of their respective (estimated) exact values. The marginal posterior distribution estimates of `copulaABCdrf` produced smaller KS distances to the respective (estimated) exact marginal posterior distributions compared to the distance of rejection ABC, but all values of the KS test statistics associated with these distances were significant, exceeding their 99% critical value 1.628 for the two-tailed (one-sample) KS test. This indicated that each of the ABC methods again produced marginal posterior distributions of the model parameter that departed with their respective exact marginal posterior distributions. Overall, `copulaABCdrf` and `rejectionABC` performed similarly with each other in terms of accuracy in estimation from the posterior distribution of this model, while their estimates were not very accurate. For `copulaABCdrf`, this is in retrospect not a very surprising result, because it is seemingly designed to handle skewed or symmetric posterior distributions but not highly multimodal posteriors.

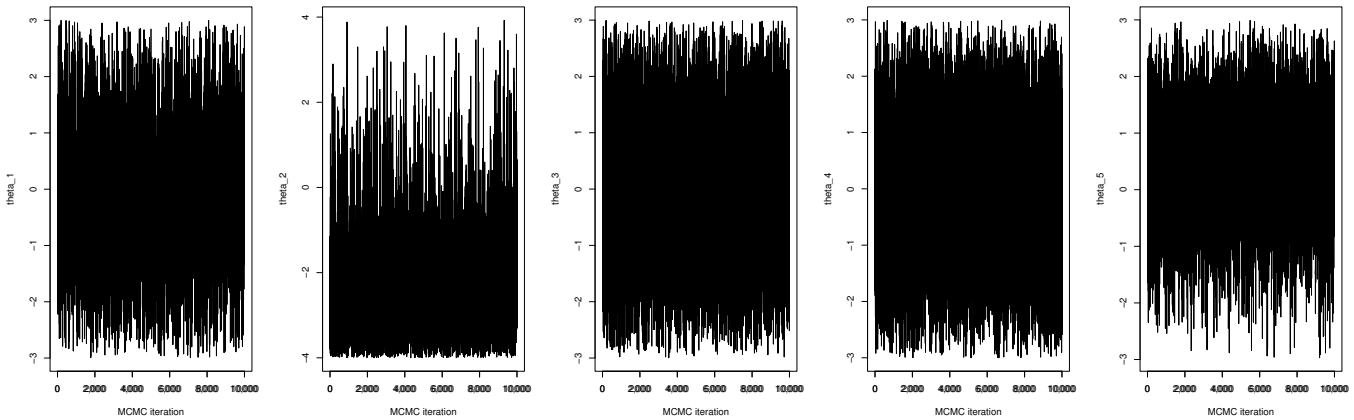


Figure 1. For each of the five model parameters (arranged in plots from left to right), trace plots of univariate marginal distributions, obtained from 10,000 MCMC Gibbs and slice samples of the exact multimodal posterior distribution, conditionally on the first data replication (for illustration). In each of the five trace plots, the x-axis refers to the sampling iteration number and the y-axis gives the realized sample value of the given model parameter.

Table 4. Simulation study results for the bivariate Gaussian model with a multimodal posterior distribution.

Mean (s.d.) of posterior estimate and 95%(50%) credible interval coverage (95%(50%)c) over 10 replicas.					
	θ_1	θ_2	θ_3	θ_4	θ_5
truth	−0.70	−2.90	−1.00	−0.90	0.60
Exact mean	−0.42 (0.44)	−2.55 (0.22)	−0.002 (0.02)	−0.01 (0.02)	0.60 (0.37)
`copulaABCdrf` mean	**−0.54** (0.44)	**−2.74** (0.27)	**0.02** (0.04)	**0.03** (0.08)	1.20 (0.29)
`rejectionABC` mean	−0.50 (0.47)	−2.73 (0.26)	0.04 (0.10)	**0.03** (0.12)	**0.21** (0.22)
Exact median	−0.48 (0.48)	−2.76 (0.26)	−0.19 (0.64)	−0.29 (0.60)	0.66 (0.42)
`copulaABCdrf` median	**−0.54** (0.47)	**−2.83** (0.29)	**0.01** (0.69)	0.33 (0.65)	1.24 (0.37)
`rejectionABC` median	−0.52 (0.49)	−2.79 (0.31)	0.15 (0.35)	**0.06** (0.18)	**0.29** (0.37)
Exact modeMLE	−0.45 (0.50)	−2.81 (0.22)	0.34 (0.84)	0.05 (0.83)	0.81 (0.68)
`copulaABCdrf` modeMLE	−0.44 (0.80)	**−3.16** (0.76)	**−0.54** (0.88)	0.33 (1.00)	1.25 (0.91)
`rejectionABCkern` modeMLE	**−0.86** (0.61)	−3.03 (0.46)	0.29 (0.66)	**−0.12** (0.60)	**0.35** (1.29)
Exact standard deviation (s.d.)	0.96 (0.13)	0.98 (0.10)	1.45 (0.19)	1.39 (0.14)	1.04 (0.14)
`copulaABCdrf` s.d.	**0.71** (0.25)	0.60 (0.16)	**1.16** (0.30)	**1.10** (0.21)	**1.05** (0.05)
`rejectionABC` s.d.	**0.71** (0.14)	**0.68** (0.10)	0.96 (0.22)	0.89 (0.14)	1.71 (0.13)
Exact 95%(50%)c	1.00 (0.40)	1.00 (0.90)	0.90 (0.10)	1.00 (0.10)	1.00 (1.00)
`copulaABCdrf` 95%(50%)c	**1.00** (**0.50**)	**1.00** (**0.50**)	**1.00** (**0.70**)	**1.00** (**0.60**)	**1.00** (**0.70**)
`rejectionABC` 95%(50%)c	**1.00** (0.70)	**1.00** (0.80)	**1.00** (0.80)	**1.00** (0.90)	**1.00** (0.80)
`copulaABCdrf` KS distance	0.22 (0.27)	0.26 (0.26)	**0.16** (0.05)	**0.23** (0.27)	**0.24** (0.06)
`rejectionABC` KS distance	**0.12** (0.03)	**0.13** (0.04)	0.27 (0.04)	0.29 (0.04)	0.25 (0.04)
`copulaABCdrf` KS test statistic	37.31 (38.78)	44.26 (47.38)	44.91 (14.13)	64.46 (72.45)	62.17 (16.70)
`rejectionABC` KS test statistic	**12.42** (3.15)	**12.77** (4.29)	**27.07** (3.98)	**29.09** (4.36)	**24.68** (3.65)
Copula	scale matrix:	θ_2	θ_3	θ_4	θ_5
	θ_1	0.15 (0.12)	−0.01 (0.04)	−0.00 (0.03)	−0.09 (0.10)
	θ_2		−0.01 (0.09)	0.00 (0.08)	−0.30 (0.13)
	θ_3			−0.00 (0.03)	0.01 (0.07)
d.f.: 8.32 (4.80)	θ_4				−0.00 (0.04)
MAE (MSE) of posterior estimate over 10 replicas.					
	θ_1	θ_2	θ_3	θ_4	θ_5
Exact mean	0.44 (0.25)	0.36 (0.17)	1.00 (1.00)	0.89 (0.79)	0.29 (0.12)
`copulaABCdrf` mean	**0.40** (**0.20**)	0.28 (**0.09**)	**1.02** (**1.04**)	**0.93** (0.88)	0.60 (0.44)
`rejectionABC` mean	0.44 (0.23)	**0.27** (**0.09**)	1.04 (1.09)	**0.93** (**0.87**)	**0.40** (**0.20**)
Exact median	0.45 (0.25)	0.24 (0.08)	0.81 (1.03)	0.61 (0.70)	0.33 (0.16)
`copulaABCdrf` median	**0.43** (**0.23**)	**0.23** (**0.08**)	**1.01** (1.45)	1.23 (1.89)	0.66 (0.53)
`rejectionABC` median	0.47 (0.25)	0.26 (0.10)	1.15 (**1.43**)	**0.96** (**0.95**)	**0.38** (**0.22**)
Exact modeMLE	0.45 (0.29)	0.20 (0.05)	1.34 (2.43)	1.01 (1.53)	0.54 (0.46)
`copulaABCdrf` modeMLE	0.56 (0.65)	0.64 (0.59)	**0.67** (**0.91**)	1.28 (2.40)	**0.94** (**1.17**)
`rejectionABCkern` modeMLE	**0.44** (**0.36**)	**0.38** (**0.21**)	1.30 (2.05)	**0.78** (**0.94**)	1.00 (1.56)

Note: **Bold** indicates the more accurate ABC method for the given estimator.

Now, we consider the twisted Gaussian model [102], a high-dimensional parametric model defined by hundreds of mean location parameters with a calculable (and known) exact but skewed posterior distribution with a strong correlation between some of the parameters. This model has received much interest in the ABC literature since [18,19]. This is because it is challenging to estimate its posterior distribution because the model has many (mean location) parameters, the model's likelihood function only provides location information, and information about the dependence among these parameters mainly comes from the prior distribution. In particular, here, we consider one 300-dimensional observation $x = (10, 0, \ldots, 0)^\intercal$ from the twisted Gaussian model defined

by likelihood $\mathcal{N}_{300}(x \mid \theta, I) = \prod_{k=1}^{d=300} \mathcal{N}(x_k \mid \theta_k, 1)$, with 300 mean (location) parameters $\theta = (\theta_1, \theta_2, \ldots, \theta_{300})^\intercal$ assigned the prior distribution (density function):

$$\pi(\theta) \propto \exp\left\{ -\frac{\theta_1^2}{200} - \frac{(\theta_2 - b\theta_1^2 + 100b)^2}{2} - \sum_{k=3}^{p=300} \theta_k^2 \right\}$$

using $b = 0.1$, as in [18,19], who considered a 250-parameter version of the model. The posterior correlation between θ_1 and θ_2 changes direction depending on whether the likelihood locates the posterior in the left or right tail of the prior, according to earlier graphical illustrations of the prior [19].

For both `rejectionABC` and `copulaABCdrf`, the summary statistic is the single 300-dimensional observation, given by $x = (10, 0, \ldots, 0)^\intercal$, with corresponding pseudo-data values $(y_{1,j}, y_{2,j}, \ldots, y_{300,j})^\intercal$ (for $j = 1, \ldots, N$) simulated from the model and used to construct the reference table for ABC, along with the prior samples of the mean locations parameters, $\theta = (\theta_1, \ldots, \theta_{300})^\intercal$. Preliminary runs of the `copulaABCdrf` framework showed that it yielded better posterior mean estimates of model parameters when univariate `drf` regression was performed on the reference table in a manner such that (for $j = 1, \ldots, N$) each of the $\theta_1^{(j)}$ and $\theta_2^{(j)}$ was regressed on $(y_1^{(j)}, y_2^{(j)})$, while the $\theta_k^{(j)}$ was regressed on the corresponding $y_k^{(j)}$ for each of $k = 3, \ldots, 300$, instead of regressing each individual model parameter on all 300 summary statistics. Therefore, the results of the former approach will only be reported. The exact posterior distribution was calculated using 10,000 samples from the MCMC Gibbs- and slice-sampling algorithm mentioned earlier.

Table 5 presents the results of the simulation study of the twisted Gaussian model based on one simulated observation and true data-generating parameter, $x = \theta_0 = (10, 0, \ldots, 0)$ as in [18,19]. These results clearly show that `copulaABCdrf` outperformed 1% nearest-neighbor rejection ABC in terms of estimation of the marginal posterior means median, mode, and MLE while producing marginal posterior distributions of the model parameters that tended to be smaller in KS distance compared to the corresponding (estimated) exact marginal posterior distributions (albeit, still with significant departures according to KS significance tests), while producing posterior means, medians, mode, and average marginal posterior interval coverages that were at least rather similar to those produced by the corresponding (estimated) exact values. In particular, for mode and MLE estimation, compared to the first mode and MLE estimation method (described earlier in Step 4 of Algorithm 1), the accuracy of the posterior mode and MLE method was improved by basing it on an extra 100,000 samples from the meta-t posterior distribution, according to the second alternative method to posterior mode and MLE estimation, as with both mode and MLE methods (described earlier in Step 4 of Algorithm 1).

In summary, all the results of the simulation studies of this subsection together show that `copulaABCdrf`, as discussed in Section 1, is equipped to estimate more unimodal, skewed, and high-dimensional posterior distributions (as in the Poisson–normal mixture model) or skewed distributions (as in the 300-dimensional twisted Gaussian model), especially compared to the rejection ABC method. However, both ABC methods produced estimates of marginal posterior distributions of model parameters that departed from the exact marginal posteriors according to Kolmogorov–Smirnov distances and their corresponding significant values of their test statistics, and could not easily handle the estimation of multimodal posterior distributions (as in the five-dimensional bivariate Gaussian model).

Table 5. Simulation results: 300-dimensional twisted Gaussian model.

	θ_1	θ_2	$\theta_3, \ldots, \theta_{300}$ (min, max)
truth	10	0	0
Exact mean	9.94	−0.04	(−0.02, 0.02)
`copulaABCdrf` mean	**9.92**	**0.02**	(−0.29, **0.21**)
`rejectionABC` mean	9.39	−0.94	(−**0.25**, 0.26)
Exact median	9.95	−0.06	(−0.02, 0.02)
`copulaABCdrf` median	**9.98**	**0.03**	(−0.36, **0.31**)
`rejectionABC` median	9.61	−0.80	(−**0.27**, **0.31**)
Exact mode	11.05	0.89	(−1.65, 1.52)
`copulaABCdrf.kern` mode	**9.84**	1.16	(−2.34, 2.96)
`copulaABCdrf.kern100k` mode	**9.80**	−**0.18**	(−**1.16**, **1.64**)
`rejectionABCprodkern` mode	10.70	3.22	(−3.23, 2.14)
Exact MLE	11.05	0.89	(−1.65, 1.52)
`copulaABCdrf.kern` MLE	9.45	−1.99	(−2.82, **2.70**)
`copulaABCdrf.kern100k` MLE	**10.36**	**1.07**	(−**2.57**, 3.51)
`rejectionABCprodkern` MLE	8.60	−1.99	(−2.60, 3.18)
Exact standard deviation (s.d.)	0.58	0.92	(0.69, 0.72)
`copulaABCdrf` s.d.	**0.63**	**1.02**	(0.56, **0.92**)
`rejectionABC` s.d.	1.58	2.79	(**0.76**, 1.23)
Exact 95%	(8.79, 11.08)	(−1.78, 1.78)	2.5%: (−1.44, −1.32); 97.5%: (1.33, 1.44)
`copulaABCdrf` 95%	(**8.62**, **11.10**)	(−**2.18**, **1.81**)	2.5%: (−**2.44**, −0.87); 97.5%: (0.78, **2.27**)
`rejectionABC` 95%	(6.02, 11.98)	(−5.20, 3.98)	2.5%: (−2.48, −**1.14**); 97.5%: (**1.21**, 2.53)
Exact 50%	(9.55, 10.34)	(−0.67, 0.58)	25%: (−0.50, −0.45); 75%: (0.45, 0.51)
`copulaABCdrf` 50%	(**9.49**, **10.34**)	(−**0.66**, **0.72**)	25%: (−**0.94**, −0.16); 75%: (**0.17**, **0.82**)
`rejectionABC` 50%	(8.31, 10.55)	(−3.00, 1.14)	25%: (−1.01, −**0.29**); 75%: (0.14, 1.00)
`copulaABCdrf` KS dist	**0.07**	**0.09**	(**0.04**, **0.22**)
`rejectionABC` KS dist	0.32	0.39	(0.06, 0.23)
`copulaABCdrf` KS test	**7.31**	**9.71**	(**2.46**, **13.49**)
`rejectionABC` KS test	31.97	39.36	(6.45, 22.77)
`copulaABCdrf` copula	scale matrix	θ_2	$\theta_3, \ldots, \theta_{300}$ (min, max)
d.f. = 117.50	θ_1	0.75	(−0.12, 0.13)
	θ_2		(−0.13, 0.13)
	$\theta_3, \ldots, \theta_{300}$		(−0.16, 0.17)

Note: **Bold** indicates the more accurate ABC method for the given estimator.

3.2. An Overview of Network Data Modeling

Network science studies systems and phenomena governing networks, which represent relational data that shape the connections of the world and are ubiquitous and focal discussion points in everyday life and the sciences. A network consists of a set of nodes (i.e., vertices, actors, individuals, sites, etc.) that can have connections (i.e., edges, ties, lines, dyadic relationships, links, bonds, etc.) with other nodes in the network; e.g., a social network consists of friendship ties among persons (nodes). A directed network treats each pair of nodes (i, j) and (j, i) as distinct, whereas an undirected network does not. Each possible edge tie connecting any given node pair in a network may be (unweighted) binary-valued; or (weighted) count-, ordinal-, categorical-, text-, ranking-, continuous-valued; and/or vector-valued, possibly as part of a multilayered (or multiplex, multi-item, or multi-relational) network representing multiple relationships among the same set of nodes. Network summary statistics include network size (number of nodes, n), edge count and density (ratio of edge count to maximum edge count in network), node degree (a node's outdegree is the number of edges starting from it and indegree is the number of edges going into it, while, in an undirected network, the outdegree and indegree coincide and are equal to the degree), the degree distribution over nodes and its geometric weighted form [103],

degree assortativity [104], counts of two-stars and triangles, and global [105] and average local [106] clustering coefficients. The presence of a dyadic network tie can depend on the presence of other dyadic ties, nodal attributes, the network's evolving dynamic structure and dyadic interactions over time, or on other factors. Also, the network's structure itself may influence or predict individuals' attributes and behaviors.

Network data are collected and analyzed in various scientific fields, including the social sciences (e.g., social or friendship networks, marriage, sexual partnerships); academic paper networks (collaboration or citation networks); neuroscience (human brain networks and interactions); political science (international relations, wars between countries, insurgencies, terrorist networks, strategic alliances and friendships); education (multilevel student relation data, item response data); economics (financial markets, economic trading, or resource distribution networks); epidemiology (disease spread dynamics, HIV, COVID-19); physics (Ising models, finding parsimonious mechanisms of network formation); biology and other natural sciences (protein–protein, molecular interaction, metabolic, cellular, genetic, ecological, and food web networks); artificial intelligence and machine learning (finding missing links in a business or a terrorist network, recommender systems, Netflix, neural networks); spatial statistics (discrete Markov random fields); traffic systems and transportation networks (roads, railways, airplanes); cooperation in an organization (advice giving in an office, identity disambiguation, business organization analysis); communication patterns and networks (detecting latent terrorist cells); telecommunications (mobile phone networks); and computer science and networks (e-mail, Internet, blogs, web, Facebook, Twitter(X), LinkedIn, information spread, viral marketing, gossiping, dating networks, blockchain network, virus propagation). Network science has seen many reviews due to its longtime scientific and societal impacts (e.g., [104,105,107–117]), while the high dimensionality of network data poses challenges to modern statistical computing methods.

Statistical and mechanistic network models are two prominent paradigms for analyzing network data. A *statistical network model* specifies an explicit likelihood function for the given network dataset, available in closed form up to a normalizing constant, making standard statistical inference tools generally available for these models, e.g., for parameter estimation and model selection. Statistical network models include the popular large family of ERGMs [118–123], which uses observable network configurations (e.g., k-stars, triangles) as sufficient statistics. An ERGM is defined by a likelihood with a normalizing constant, which is intractable for a network with more than a handful of nodes n. The ERGM for a binary undirected or directed network is as follows:

$$f(X = x \mid \theta) = \frac{\exp\{\eta(\theta)^\intercal h(x; z)\}}{\sum_{x' \in \mathcal{X}_n} \exp\{\eta(\theta)^\intercal h(x'; z)\}}, \tag{3}$$

with parameter vector $\theta \in \Theta \subseteq \mathbb{R}^q$ and its mapping to canonical parameters $\eta : \Theta \to \mathbb{R}^d$ (with (3) a linear ERGM if $\eta(\theta) \equiv \theta$, and, otherwise, a curved ERGM); sufficient statistics $h \in \mathbb{R}^d$, which are possibly dependent on any covariates z; and, for undirected binary networks X, the space $\mathcal{X}_n$ of allowable networks on n nodes is $\mathcal{X}_n = \{x \in \mathbb{R}^{n \times n}, x_{i,j} = x_{j,i} \in \{0, 1\}, x_{i,i} = 0\}$, and, for directed binary networks X, the space $\mathcal{X}_n$ is the same but without the restrictions $x_{i,j} = x_{j,i}$. The ERGM likelihood (3) is intractable for a binary network with more than a handful of nodes n, as the normalizing constant of the model's likelihood is a sum of $2^{n(n-1)/2}$ terms (or $2^{n(n-1)}$ terms for a directed network) over the sample space of allowable networks. The basic ERGM form (3) allows it to be straightforwardly extended to a general ERGM representation for multilayered weighted (valued) networks, shown later in Section 3.8, that combines certain ERGMs [123–125] and encompasses ERGMs: for binary networks; with size(n) offset adjusted (invariant) parameterizations [94,126]; with nodal random effects [127]; and for valued (weighted) [123,124], multilayered [125], and dynamic networks [128–130].

A mechanistic network model is able to incorporate domain scientific knowledge into generative mechanisms for network formation, enabling researchers to investigate complex systems using both simulation and analytical techniques. A *mechanistic network model* is an

algorithmic description of network formation defined by a few domain-specific stochastic microscopic mechanistic rules by which a network grows and evolves over time, which are hypothesized and informed by understanding and related critical elements of the given scientific problem. A typical mechanistic model generates a network by starting with a small seed network and then grows the network one node at a time according to the model's generative mechanism until some stopping condition is met, e.g., until a requisite number of nodes is reached, which can possibly number in the millions. There are easily hundreds of mechanistic network models, which originated in physics. For a long time, they were essentially the only type of network model formulated and studied using both mathematical and computer simulation methods. Mechanistic network models include the model of Price [131,132], used to study citation patterns of scientific papers, and the Barabasi and Albert [133] model, which introduce preferential attachment rules for directed and undirected networks, respectively. Also used was the Watts and Strogatz [106] model, which produces random graphs with small-world properties, including short average path lengths and high clustering. Further, the duplication–divergence class of models describes the evolution of protein–protein interaction networks (where each node is a protein in an organism and two proteins are connected if they are involved in a chemical reaction together). This class includes the duplication–mutation–complementation (DMC) model [134,135] and the duplication–mutation–random (DMR) model [136,137]. Other mechanistic models include the KM model [138,139], used to study HIV epidemics [140,141], along with other mechanistic models [142–144]. The price and nonlinear preferential attachment models and the DMC and DMR models are further described later within the simulation studies reported in Sections 3.4 and 3.5.

We will apply Algorithm 1 to the ERGM and mechanistic network models through simulation studies in Sections 3.3–3.5 and analyses and real-life large networks in Sections 3.6–3.8. Algorithm 1 can be used to analyze a network dataset x using an $M \geq 1$ ERGM and/or mechanistic network models based on K scalar and/or vector network statistics, $g_k(x)$, for $k = 1, \ldots, K$. As mentioned, for the analysis of a large network, it is computationally prohibitive to simulate one very large network from the ERGM or mechanistic network model. This issue can be addressed by a strategy that simulates network datasets of a smaller size compared to the size of the network dataset under statistical analysis while using network summary statistics (calculated on the observed dataset and each simulated dataset) that take on values that are invariant to the size (number of nodes) of the network dataset being analyzed.

In terms of Algorithm 1, we can specify computationally efficient network summary statistics s, which can be used to directly compare between two networks that may have different node sets, sizes (number of nodes), and densities, in other words, summary statistics that allow for unknown node correspondence (UNC) network comparisons ([145] p. 2), including MPLEs of size(n) offset adjusted (invariant) parameterized ERGMs [94,126], described below. Other computationally efficient UNC summary statistics include global and average local clustering coefficients (based on triangle counts) [146–149], degree distribution, including its mean and variance or geometric form, degree assortativity (propensity for similar nodes to connect), and diameter (average of shortest path lengths over all pairs of nodes in a network) [147,149]. Generally speaking, specifying the summary statistics vector s by computationally efficient summary statistics that allow for UNC (size-invariant) comparisons of networks enables Steps 1(c)–1(d) to simulate and compute these summaries of a network of a smaller size n_{sim} and to directly compare these summaries with the corresponding summaries of a larger observed network dataset x (of size n) being analyzed for each of the N_M sampling iterations of Algorithm 1. This lowers computational cost compared to computing and comparing such a summary based on simulating networks of the same size (n) as the large observed network in each sampling iteration.

As mentioned, a novel contribution is that, for either the ERGM or a mechanistic network model, the UNC network summary statistics vector $s(\cdot)$ may also include MPLEs of K ERGMs defined by sufficient statistics $g_k(x)$, respectively, for $k = 1, \ldots, K$, with these

MPLEs adjusted by network size by specifying in each ERGM an offset term x (edge count of network x) with fixed coefficient $\log(\frac{1}{n})$ [94]. The MPLE (e.g., [150]) was introduced for lattice models [151] and developed to estimate ERGM parameters [118,152] because the exact MLE of the ERGM is intractable for a binary network with more than a handful of nodes n, as mentioned above. Specifically, for example, if the observed network dataset x is an undirected binary network, $x = (x_{i,j} \in \{0,1\})_{n \times n}$ on n nodes, described by scalar or vector valued statistics $g_k(x)$ for $k = 1, \ldots, K$ (where each g_k may depend on covariates z_k), then the summary statistics s of x (or y) can be specified as $s(x) = (\widehat{\beta}^{(x)}_{\text{MPLE},1}, \ldots, \widehat{\beta}^{(x)}_{\text{MPLE},K})^\intercal$, where $\widehat{\beta}^{(x)}_{\text{MPLE},k}$ is the MPLE of the ERGM (3) using sufficient statistics $g_k(x)$ and the offset term of the network edge count x with fixed coefficient $\log(1/n)$. Each MPLE summary statistic $\widehat{\beta}_{\text{MPLE},k}$ (for $k = 1, \ldots, K$) is obtained by maximizing a logistic regression likelihood as follows:

$$\widehat{\beta}^{(x)}_{\text{MPLE},k} = \underset{\beta_k \in \mathbb{R}^{q_k}}{\arg\max} \sum\sum_{1 \le i < j \le n} \begin{pmatrix} x_{i,j}\{(\eta_k(\beta_k)^\intercal, \log(\frac{1}{n}))^\intercal \Delta_{i,j,k}\} \\ -\log[1 + \exp\{(\eta_k(\beta_k)^\intercal, \log(\frac{1}{n}))^\intercal \Delta_{i,j,k}\}] \end{pmatrix},$$

with $\beta_k \in \mathbb{R}^{q_k}$, $\eta_k : \Theta \to \mathbb{R}^{d_k}$, $g_k \in \mathbb{R}^{d_k}$, and network change statistics $\Delta_{i,j,k} = ((g_k(x^+_{i,j}) - g_k(x^-_{i,j}))^\intercal, x^+_{i,j} - x^-_{i,j})^\intercal$, where $x^+_{i,j}$ and $x^-_{i,j}$ are network x with $x_{i,j} = x_{j,i} = 1$ and $x_{i,j} = x_{j,i} = 0$, respectively. Likewise, MPLEs can be obtained from a directed network or from a weighted (undirected or directed) network based on a binary representation of a polytomous network ([123] §4.3). Such a logistic regression likelihood represents a special form of composite likelihood [153,154] that assumes (dyadic) independence of the $x_{i,j}$ observations, $\Pr_\theta(x_{i,j} = 1 \mid X^c_{i,j} = x^c_{i,j}) = \Pr_\theta(x_{i,j} = 1)$, where $x^c_{i,j}$ is the network x excluding $x_{i,j}$. Both the MPLE and MLE of the ERGM are consistent for a growing number of networks observed from the same set of fixed nodes [155]. The MPLE can be rapidly computed from large networks [156] using divide-and-conquer [157–159] or streaming methods [160], if necessary.

3.3. Simulation Study: Exponential Random Graph Models (ERGMs)

Now, we consider a simulation study to evaluate and compare the ability of `copulaABCdrf` (Algorithm 1) and `rejectionABC` to estimate the posterior distribution of ERGM parameters based on 10 replications of $n = 50$ node undirected networks from the ERGM and of $n = 300$ node undirected networks simulated from this model, for each of five conditions of n_{sim}, being at least roughly 10%, 33%, 50%, 75%, and 100% of n network nodes. Specifically, the conditions $n_{\text{sim}} = 5, 16, 25, 37$, and 50 for $n = 50$ nodes and $n_{\text{sim}} = 30, 100, 150, 225$, and 300 for $n = 300$ nodes.

Under the $n = 50$ simulation conditions, each of the 10 network datasets were simulated from the ERGM (3) using true data-generating parameters $\theta_0^\intercal = (-0.20, 0.50)$ and defined by network sufficient statistics being the number of two-stars and triangles. For each simulated network dataset x analyzed by Algorithm 1 using the ERGM, this model was assigned a g prior $\theta \sim \mathcal{N}_2(0, g(-\mathcal{H}(\widehat{\theta}_{\text{MPLE}}))^{-1})$ with $g = 100$, where $\mathcal{H}(\widehat{\theta}_{\text{MPLE}})$ is the Hessian matrix of the MPLE of θ, and the summary statistics were specified as $s(x) = (\widehat{\beta}^{(x)}_{\text{MPLE},1}, \widehat{\beta}^{(x)}_{\text{MPLE},2})^\intercal$, where $\widehat{\beta}^{(x)}_{\text{MPLE},k}$ for $k = 1, 2$ are the MPLEs for these two network sufficient statistics, respectively, based on the edge count x offset with fixed coefficient $\log(1/n)$. Likewise, $s(y) = (\widehat{\beta}^{(y)}_{\text{MPLE},1}, \widehat{\beta}^{(y)}_{\text{MPLE},2})^\intercal$ based on the edge count y offset with fixed coefficient $\log(1/n_{\text{sim}})$, and on the size $n_{\text{sim}} = 25$ of the network dataset y simulated in each iteration of Algorithm 1.

Under the $n = 300$ simulation conditions, each of 10 network datasets was simulated from the ERGM (3) using true data-generating parameters $\theta_0^\intercal = (-0.20, 0.50, 0.80)$ and defined by network sufficient statistics being the number of two-stars (`kstar(2)`), the number of triangles (`triangles`), and the geometric weighted degree distribution (`degree1.5`), with the decay parameter fixed at $\alpha \equiv \log(1.5)$ ([103] from Equations (11) and (12), pp. 112, 126). For each simulated network dataset x analyzed by Algorithm 1 using the ERGM, this

model was assigned a g prior, given by $\theta \sim \mathcal{N}_3(0, g(-\mathcal{H}(\widehat{\theta}_{\text{MPLE}}))^{-1})$ with $g = 100,000$, and the summary statistics were specified as $s(x) = (\widehat{\beta}^{(x)}_{\text{MPLE},1}, \widehat{\beta}^{(x)}_{\text{MPLE},2}, \widehat{\beta}^{(x)}_{\text{MPLE},3})^\intercal$, where $\widehat{\beta}^{(x)}_{\text{MPLE},k}$ for $k = 1, 2, 3$, are the MPLEs for these three network sufficient statistics, respectively, based on the edge count x offset with fixed coefficient $\log(1/n)$. The same applies to $s(y) = (\widehat{\beta}^{(y)}_{\text{MPLE},1}, \widehat{\beta}^{(y)}_{\text{MPLE},2}, \widehat{\beta}^{(y)}_{\text{MPLE},3})^\intercal$ based on the edge count y offset with fixed coefficient $\log(1/n_{\text{sim}})$, and on the size $n_{\text{sim}} = 100$ of the network dataset y simulated in each iteration of Algorithm 1.

Table 6 presents some detailed representative results for $n_{\text{sim}} = 25$ of $n = 50$ node networks, and for $n_{\text{sim}} = 100$ of $n = 300$ node networks. Table 7 presents the results for each of all of the five conditions of n_{sim} for $n = 50$ nodes and for $n = 300$ nodes. For $n = 50$ nodes, in terms of MAE and MSE, `rejectionABC` tended to outperform `copulaABCdrf` in the estimation accuracy of the posterior mean, median, and mode. Also, `copulaABCdrf` outperformed `rejectionABC` in terms of MLE accuracy. Further, `rejectionABC` performed similarly with `copulaABCdrf` in terms of accuracy in the estimation of 95% and 50% posterior credible intervals while being typically near the 95% nominal values. In addition, `rejectionABC` was slightly superior in the estimation of the 50% posterior credible interval, but both methods often produced 50% interval estimates that were far from this nominal value. Also, while n_{sim} increased toward the full network dataset sample size of $n = 50$ nodes, the MAEs and MSEs tended to decrease for each of the point estimates (posterior mean, median, mode, and MLE) for both ABC methods. For $n_{\text{sim}} = 37$, `copulaABCdrf` produced MLEs that were competitive with MCMLEs.

For $n = 300$ nodes, `rejectionABC` tended to outperform `copulaABCdrf` in terms of the estimation accuracy of the posterior mean, median, and mode, especially for parameters θ_1 and θ_3. Also, `copulaABCdrf` outperformed `rejectionABC` in terms of MLE accuracy for all model parameters, and in terms of accuracy in the estimation of 95% posterior credible intervals, while being typically near the nominal values. Further, `rejectionABC` tended to outperform `copulaABCdrf` in terms of accuracy in the estimation of 50% posterior credible intervals. Again, both ABC methods often produced 50% interval estimates far away from this nominal value. Finally, as n_{sim} increased toward the full network dataset sample size of $n = 300$ nodes, for each of the two ABC methods, there was no clear pattern of the MAEs and MSEs decreasing for each of the point estimates and for the posterior credible intervals to approach their nominal rates.

Table 6. Results of simulation study for the ERGM. Mean (standard deviation) of posterior estimates over 10 replicas, under some conditions of n and n_{sim}.

	kstar(2)	**triangles**	**kstar(2)**	**triangles**	**degree1.5**
n	50	50	300	300	300
n_{sim}	25	25	100	100	100
true θ	−0.20	0.50	−0.20	0.50	0.80
`copulaABCdrf` mean	−0.22 (0.03)	0.95 (0.35)	−0.03 (0.02)	**0.49** (0.10)	−0.15 (0.12)
`rejectionABC` mean	**−0.21** (0.04)	**0.79** (0.39)	**−0.05** (0.01)	−0.2 (0.07)	**0.33** (0.06)
`copulaABCdrf` median	−0.22 (0.03)	0.86 (0.33)	0.01 (0.03)	**0.28** (0.07)	−0.85 (0.28)
`rejectionABC` median	**−0.20** (0.04)	**0.67** (0.32)	**−0.04** (0.01)	−0.30 (0.03)	**0.30** (0.07)
`copulaABCdrf` mode	−0.18 (0.06)	**0.57** (0.51)	0.06 (0.07)	**1.18** (0.64)	−1.14 (0.07)
`rejectionABCkern` mode	**−0.19** (0.04)	0.67 (0.37)	**−0.03** (0.05)	−0.27 (0.11)	**0.29** (0.22)
`copulaABCdrf` MLE	**−0.22** (0.06)	**0.79** (0.48)	0.01 (0.07)	1.61 (0.58)	−1.09 (0.09)
`rejectionABCkern` MLE	−0.33 (0.06)	1.98 (1.69)	**−0.31** (0.21)	**0.89** (1.47)	**0.82** (0.89)
MCMLE	−0.20 (0.02)	0.46 (0.16)	−0.03 (0.01)	0.09 (0.01)	0.08 (0.03)
MPLE	0.15 (0.00)	1.18 (0.08)	0.03 (0.00)	0.26 (0.00)	0.18 (0.00)

Table 6. *Cont.*

	kstar(2)	**triangles**	**kstar(2)**	**triangles**	**degree1.5**
`copulaABCdrf` s.d.	0.07 (0.01)	0.66 (0.12)	0.27 (0.02)	0.79 (0.10)	1.51 (0.21)
`rejectionABC` s.d.	0.05 (0.01)	0.62 (0.19)	0.10 (0.01)	0.46 (0.14)	0.38 (0.04)
SE(MCMLE)	0.02 (0.00)	0.19 (0.01)	0.00 (0.00)	0.01 (0.00)	0.02 (0.00)
`copulaABCdrf`	d.f.	scale	d.f.:	scale matrix:	
Copula d.f. and scale	14.55 (13.33)	−0.48 (0.12)	20.48 (24.78)	triangles	degree1.5
			kstar(2)	−0.07 (0.08)	−0.97 (0.01)
			triangles		−0.09 (0.09)

Note: **Bold** indicates the more accurate ABC method for the given estimator.

Table 7. MAE (MSE), mean percent interval coverage (95%(50%)c) for ERGM over 10 replicas, varying n and n_{sim}.

		n = 50			n = 300		
n_{sim}	**Posterior**	**kstar(2)**	**triangles**	n_{sim}	**kstar(2)**	**triangles**	**degree1.5**
5	`copulaABCdrf` mean	**0.22** (**0.05**)	**0.80** (**0.64**)	30	0.29 (0.08)	**1.43** (**2.07**)	1.27 (1.60)
	`rejectionABC` mean	**0.22** (**0.05**)	0.84 (0.71)		**0.08** (**0.01**)	1.58 (2.50)	**0.56** (**0.36**)
	`copulaABCdrf` median	**0.22** (**0.05**)	**0.73** (**0.54**)		0.30 (0.09)	**1.47** (**2.20**)	1.45 (2.09)
	`rejectionABC` median	**0.22** (**0.05**)	0.77 (0.60)		**0.04** (**0.00**)	1.58 (2.51)	**0.83** (**0.70**)
	`copulaABCdrf` mode	**0.17** (**0.04**)	**0.58** (0.54)		0.19 (0.04)	**0.61** (**0.51**)	1.23 (1.63)
	`rejectionABCkern` mode	0.20 (**0.04**)	**0.58** (**0.37**)		**0.04** (**0.00**)	1.54 (2.40)	**0.76** (**0.62**)
	`copulaABCdrf` MLE	**0.16** (**0.04**)	**0.62** (**0.55**)		**0.15** (**0.03**)	**0.50** (**0.39**)	**1.23** (**1.60**)
	`rejectionABCkern` MLE	0.36 (0.17)	2.47 (7.31)		0.18 (0.04)	2.48 (6.76)	2.20 (5.67)
	MCMLE	0.01 (0.00)	0.17 (0.04)		0.16 (0.03)	0.40 (0.16)	0.72 (0.52)
	MPLE	0.35 (0.12)	0.67 (0.45)		0.23 (0.05)	0.24 (0.06)	0.62 (0.38)
	`copulaABCdrf` 95%(50%)c	0.00 (**0.00**)	**1.00** (**0.00**)		**0.90** (0.00)	**0.90** (**0.00**)	**1.00** (0.00)
	`rejectionABC` 95%(50%)c	0.10 (**0.00**)	**1.00** (**0.00**)		**1.00** (**1.00**)	0.50 (**0.00**)	**1.00** (**0.40**)
	MCMLE 95%c	1.00	1.00		0.00	0.00	0.00
16	`copulaABCdrf` mean	0.07 (0.01)	0.49 (0.25)	100	0.17 (0.03)	**0.07** (**0.01**)	0.95 (0.92)
	`rejectionABC` mean	**0.06** (**0.00**)	**0.37** (**0.14**)		**0.15** (**0.02**)	0.70 (0.49)	**0.47** (**0.22**)
	`copulaABCdrf` median	0.07 (0.01)	0.51 (0.26)		0.21 (0.04)	**0.22** (**0.05**)	1.65 (2.80)
	`rejectionABC` median	**0.06** (**0.00**)	**0.33** (**0.11**)		**0.16** (**0.02**)	0.80 (0.64)	**0.50** (**0.25**)
	`copulaABCdrf` mode	0.09 (0.01)	0.46 (0.29)		0.26 (0.07)	**0.72** (0.83)	1.94 (3.77)
	`rejectionABCkern` mode	**0.05** (**0.00**)	**0.19** (**0.05**)		**0.17** (**0.03**)	0.77 (**0.60**)	**0.51** (**0.30**)
	`copulaABCdrf` MLE	**0.07** (**0.01**)	0.53 (0.43)		0.21 (**0.05**)	**1.11** (**1.53**)	1.89 (3.60)
	`rejectionABCkern` MLE	0.09 (0.01)	**0.35** (**0.16**)		**0.20** (**0.05**)	1.26 (2.10)	**0.72** (**0.71**)
	MCMLE	0.01 (0.00)	0.14 (0.03)		**0.17** (**0.03**)	0.41 (0.16)	0.72 (0.52)
	MPLE	0.35 (0.12)	0.66 (0.44)		0.23 (0.05)	0.24 (0.06)	0.62 (0.38)
	`copulaABCdrf` 95%(50%)c	**1.00** (**0.20**)	**1.00** (0.00)		**1.00** (**0.00**)	**1.00** (**0.90**)	**1.00** (**0.20**)
	`rejectionABC` 95%(50%)c	**1.00** (0.10)	**1.00** (**0.10**)		**0.90** (**0.00**)	0.60 (0.00)	**1.00** (0.00)
	MCMLE 95%c	0.90	0.90		0.00	0.00	0.00
25	`copulaABCdrf` mean	**0.03** (**0.00**)	0.47 (0.31)	150	**0.11** (**0.01**)	**0.04** (**0.00**)	0.66 (0.46)
	`rejectionABC` mean	**0.03** (**0.00**)	**0.41** (**0.22**)		**0.11** (**0.01**)	0.68 (0.47)	**0.29** (**0.09**)
	`copulaABCdrf` median	**0.03** (**0.00**)	0.42 (0.23)		0.14 (0.02)	**0.03** (**0.00**)	0.67 (0.57)
	`rejectionABC` median	**0.03** (**0.00**)	**0.31** (**0.12**)		**0.11** (**0.01**)	0.82 (0.69)	**0.32** (**0.11**)
	`copulaABCdrf` mode	0.05 (**0.00**)	0.41 (0.24)		0.29 (0.09)	**0.60** (**0.51**)	1.98 (3.94)
	`rejectionABCkern` mode	**0.03** (**0.00**)	**0.35** (**0.15**)		**0.19** (**0.04**)	0.94 (0.90)	**0.54** (**0.38**)
	`copulaABCdrf` MLE	**0.05** (**0.00**)	**0.43** (**0.29**)		0.25 (**0.06**)	**0.93** (**1.04**)	2.01 (4.06)
	`rejectionABCkern` MLE	0.13 (0.02)	1.66 (4.77)		**0.24** (0.07)	0.94 (1.33)	**0.86** (**1.28**)
	MCMLE	0.02 (0.00)	0.12 (0.02)		0.17 (0.03)	0.40 (0.16)	0.72 (0.52)
	MPLE	0.35 (0.12)	0.68 (0.47)		0.23 (0.05)	0.24 (0.06)	0.62 (0.38)
	`copulaABCdrf` 95%(50%)c	**1.00** (**0.60**)	**1.00** (**0.50**)		**1.00** (1.00)	**1.00** (**1.00**)	**1.00** (0.90)
	`rejectionABC` 95%(50%)c	**1.00** (**0.60**)	**1.00** (**0.50**)		**1.00** (**0.20**)	**1.00** (**0.00**)	**1.00** (**0.60**)
	MCMLE 95%c	0.90	1.00		0.00	0.00	0.00

Table 7. *Cont.*

		n = 50			n = 300		
n_{sim}	**Posterior**	**kstar(2)**	**triangles**	n_{sim}	**kstar(2)**	**triangles**	**degree1.5**
37	`copulaABCdrf` mean	**0.01** (**0.00**)	0.16 (0.03)	225	0.44 (0.20)	0.29 (0.09)	2.12 (4.59)
	`rejectionABC` mean	**0.01** (**0.00**)	**0.08** (**0.01**)		**0.20** (**0.04**)	**0.16** (**0.04**)	**0.97** (**0.96**)
	`copulaABCdrf` median	**0.01** (**0.00**)	0.16 (0.04)		0.45 (0.22)	**0.25** (**0.07**)	2.12 (4.64)
	`rejectionABC` median	**0.01** (**0.00**)	**0.09** (**0.01**)		**0.18** (**0.04**)	0.32 (0.11)	**0.79** (**0.64**)
	`copulaABCdrf` mode	0.04 (**0.00**)	0.15 (0.03)		0.43 (0.20)	**0.34** (**0.15**)	2.44 (6.26)
	`rejectionABCkern` mode	**0.02** (**0.00**)	**0.07** (**0.01**)		**0.11** (**0.02**)	0.71 (0.54)	**0.55** (**0.51**)
	`copulaABCdrf` MLE	**0.02** (**0.00**)	**0.17** (**0.05**)		**0.59** (**0.47**)	**0.42** (**0.28**)	**3.21** (**12.45**)
	`rejectionABCkern` MLE	0.10 (0.01)	0.32 (0.22)		0.75 (0.60)	0.87 (1.17)	3.89 (15.68)
	MCMLE	0.01 (0.00)	0.11 (0.02)		0.17 (0.03)	0.40 (0.16)	0.73 (0.53)
	MPLE	0.35 (0.12)	0.73 (0.53)		0.23 (0.05)	0.24 (0.06)	0.62 (0.38)
	`copulaABCdrf` 95%(50%)c	**1.00** (**1.00**)	**1.00** (**0.70**)		**1.00** (1.00)	**1.00** (1.00)	**1.00** (1.00)
	`rejectionABC` 95%(50%)c	**1.00** (**1.00**)	**1.00** (0.90)		**1.00** (**0.40**)	**1.00** (**0.90**)	**1.00** (**0.10**)
	MCMLE 95%c	1.00	1.00		0.00	0.00	0.00
50	`copulaABCdrf` mean	**0.01** (**0.00**)	0.09 (0.02)	300	**0.52** (**0.28**)	**0.34** (**0.15**)	**2.91** (**8.78**)
	`rejectionABC` mean	**0.01** (**0.00**)	**0.07** (**0.01**)		0.75 (0.57)	0.44 (0.20)	4.37 (19.42)
	`copulaABCdrf` median	**0.01** (**0.00**)	0.10 (**0.01**)		**0.43** (**0.19**)	**0.44** (**0.23**)	**2.33** (**5.76**)
	`rejectionABC` median	0.02 (**0.00**)	**0.07** (**0.01**)		0.70 (0.51)	0.49 (0.26)	4.09 (17.20)
	`copulaABCdrf` mode	0.04 (**0.00**)	0.32 (0.14)		**0.17** (**0.06**)	**0.33** (**0.15**)	**0.92** (**1.92**)
	`rejectionABCkern` mode	**0.02** (**0.00**)	**0.12** (**0.02**)		0.60 (0.42)	0.58 (0.43)	3.60 (14.50)
	`copulaABCdrf` MLE	**0.04** (**0.00**)	**0.33** (**0.17**)		**0.19** (**0.07**)	**0.39** (**0.23**)	**0.96** (**2.04**)
	`rejectionABCkern` MLE	0.11 (0.01)	0.44 (0.45)		2.05 (4.46)	1.47 (2.61)	11.67 (141.65)
	MCMLE	0.01 (0.00)	0.14 (0.03)		0.17 (0.03)	0.40 (0.16)	0.74 (0.55)
	MPLE	0.35 (0.12)	0.68 (0.46)		0.23 (0.05)	0.24 (0.06)	0.62 (0.38)
	`copulaABCdrf` 95%(50%)c	**1.00** (0.80)	**1.00** (**0.90**)		**1.00** (**0.10**)	**1.00** (**0.80**)	**1.00** (**0.00**)
	`rejectionABC` 95%(50%)c	**1.00** (**0.70**)	**1.00** (1.00)		0.70 (0.00)	**1.00** (1.00)	0.30 (**0.00**)
	MCMLE 95%c	1.00	1.00		0.00	0.00	0.00

Note: **Bold** indicates the more accurate ABC method for the given estimator.

3.4. Simulation Study: Preferential Attachment (PA) Models

Mechanistic network models defined by preferential attachment rules, including the price model and the nonlinear preferential attachment (NLPA) model, grow a directed or undirected network by introducing a new node at each stage and linking the new node to any given existing node i with probability $\Pr(\text{new node attaches to old node } i) \propto k_0 + k_i^\alpha$, with some constant parameter $k_0 \in \mathbb{R}$, where k_i is the degree of node i (indegree for a directed network), and power parameter $\alpha \in (0, \infty)$. In particular, the price model grows a directed binary citation network one article (node) at a time (each edge being a paper citing another paper), with $\Pr(\text{new article cites existing article } i) \propto k_0 + k_i$, with $\alpha \equiv 1$ and constant parameter k_0 centered around 1, and k_i is the indegree k_i of existing article i, such that each new article cites m existing articles on average, and the number of articles that a new article cites follows a binomial (n_0, p) distribution, with success probability parameter p, and with n_0 set as the maximum outdegree of x [161]. The unknown price model parameters to be estimated from the observed directed network dataset x are (k_0, p). For an undirected (binary) network x, the Barabasi and Albert [133] (BA) model has the same form, but without the assumptions $k_0 = 0$ and $\alpha \equiv 1$, and based on k_i defined as the degree of existing node i.

The NLPA model [162] generalizes the BA model using $\Pr(\text{new node attaches to existing node } i) = \dfrac{k_i^\alpha}{\sum_{\text{Existing nodes } j} k_j^\alpha}$, $\alpha > 0$, with k_i the degree of existing node i. Since possibly $\alpha \neq 1$, the number of nodes that a new node attaches was assumed to follow a truncated binomial (n_0, p) distribution that supports only positive counts. For the NLPA model, the unknown parameters to be estimated from the given undirected network dataset are (α, p). Thus, from an observed network dataset x, the unknown model parameters to be estimated are (α, p), and n_0 is chosen as the maximum degree of x.

Now, we consider simulation studies that evaluate and compare the ability of `copulaABC-drf` (Algorithm 1) and `rejectionABC` to estimate the posterior distribution of the parameters of the price model, and the parameters of the posterior distribution of the NLPA model. The simulation studies are based on 10 replications of $n = 50$ node-directed networks from the price model (undirected networks from the NLPA model, respectively), and of $n = 300$ node-undirected networks simulated from the price model (undirected networks from the NLPA model, respectively). These simulations were performed for each of five conditions of n_{sim}, being at least roughly 10%, 33%, 50%, 75%, and 100% of n network nodes; specifically, the conditions $n_{\text{sim}} = 5, 16, 25, 37$, and 50 for $n = 50$ nodes, and $n_{\text{sim}} = 30, 100, 150, 225$, and 300 for $n = 300$ nodes. The true data-generating parameters of the price model were specified as $\theta_0^{\mathsf{T}} = (k_0, p) = (1, 0.02)$, while the true data-generating parameters of the NLPA model were $\theta_0^{\mathsf{T}} = (\alpha, p) = (1.2, 0.02)$.

For each network dataset x simulated from the price model and analyzed by Algorithm 1 using this model, this model was assigned a uniform prior $(k_0, p) \sim \mathcal{U}_{(0.9,1.1)}\mathcal{U}_{(0,b)}$ (with $b = 0.10$ under $n = 50$ and $b = 0.20$ for $n = 300$) and used a vector of summaries $s(\cdot)$ specified by network size invariant offset MPLEs of geometrically weighted degree, its decay estimate, and triangle count. For each network dataset x simulated from the NLPA model and analyzed by Algorithm 1 using this model, the NLPA model was assigned a uniform prior $(\alpha, p) \sim \mathcal{U}_{(0,3)}\mathcal{U}_{(0,0.20)}$ using a vector of summaries $s(\cdot)$ specified by the network size invariant offset MPLEs of density, average clustering coefficient, and diameter. The MPLE summaries $s(x)$ of each network x used offsets based on n nodes ($n = 50$ or 300). The MPLE summaries $s(y)$ of each network y of smaller size $n_{\text{sim}} < n$ (being $n_{\text{sim}} = 25$ for $n = 50$, and $n_{\text{sim}} = 100$ for $n = 300$) were simulated in each iteration of the ABC algorithm using offsets based on n_{sim} nodes and based on a binomial $(n_{\text{sim}} - 1, p)$ distribution (or the truncated binomial for NLPA).

Table 8 presents some representative detailed results of the simulation study of `copulaABCdrf` and `rejectionABC` for the price and NLPA models, for pairs of values of n_{sim} and n with $n_{\text{sim}} < n$, namely, $n_{\text{sim}} = 25$ for $n = 50$, and $n_{\text{sim}} = 100$ for $n = 300$. Tables 9 and 10 present the results of MAE, MSE, and accuracy of 95% and 50% posterior interval credible interval estimates of `copulaABCdrf` and `rejectionABC` for the price and NLPA models, for each of the five n_{sim} conditions within each of $n = 50$ and $n = 300$.

In particular, since, for each of the price and NLPA models, there were three summary statistics relative to the two models parameters, we not only considered for each model implementations of `rejectionABC` using all three summary statistics but also considered, for comparison purposes, each implementation of `rejectionABC` based on pre-selecting two of the most important of the three summary statistics, according to the results obtained from a variable importance analysis based on training a `drf` multivariate regression on a reference table of simulated model parameters on all three summary statistics (covariate variables), as described earlier in Section 3. For the price model, `drf` variable importance analyses found that, over all ten replicas of network datasets under the $n_{\text{sim}} = 5$ and $n = 50$ condition, the network size invariant offset MPLEs of geometrically weighted degree, and their decay estimate, were always the most important summary statistics of the three total summaries. Also, for each of the other simulation conditions of n_{sim} and n, the `drf` variable importance analyses found that, over all ten replicas of network datasets, the most important summaries were always the network size invariant offset MPLEs of geometrically weighted degree and triangle counts. For the NLPA model, the `drf` variable importance analyses always found that network density and average clustering coefficient were the two most important network summaries of the three total summaries compared to the network diameter summary, for each of the 10 network replicas, and within each of the conditions $n_{\text{sim}} = 5, 16, 25, 37$, and 50 for $n = 50$ nodes, and the conditions of $n_{\text{sim}} = 30, 100, 150, 225$, and 300 for $n = 300$ nodes.

From the results of Table 9 on the price model, it can be concluded overall that `copulaABCdrf` and `rejectionABC` performed similarly. Also, `rejectionABC`, based on using `drf` to pre-select two of the three most important summaries, produced results that sometimes slightly improved but were usually similar to the results of `rejectionABC` using all three

summaries. More specifically, for the parameter k_0 and $n = 50$ nodes, `copulaABCdrf` slightly outperformed `rejectionABC` in MSE of posterior mode estimation and MLE, and `rejectionABC` outperformed `copulaABCdrf` in MAE of posterior mode and MLE. For the parameter p and $n = 50$ nodes, `copulaABCdrf` slightly outperformed `rejectionABC` in the estimation accuracy of the 95% credible interval while being close to the nominal level, while `rejectionABC` slightly outperformed `copulaABCdrf` in terms of MSE of the estimation of the posterior mode and MLE. For the parameter k_0 and $n = 300$ nodes, `copulaABCdrf` slightly outperformed `rejectionABC` in MSE of the estimation of the posterior mode and MLE, and `rejectionABC` outperformed `copulaABCdrf` in MAE of the estimation of the posterior mode and MLE. For the parameter p and $n = 300$ nodes, `copulaABCdrf` outperformed `rejectionABC` in accuracy in the estimation of the 95% posterior credible interval and slightly outperformed `rejectionABC` in terms of accuracy in the estimation of the 50% posterior credible interval, while not being near the nominal rate a few times.

Table 8. Simulation study results for price and NLPA models. Mean (standard deviation) of posterior estimates over 10 replicas, under some conditions n and n_{sim}.

Price model	k_0	p	k_0	p
n	50	50	300	300
n_{sim}	25	25	100	100
true θ	1.00	0.02	1.00	0.02
`copulaABCdrf` mean	**1.00** (0.00)	**0.05** (0.01)	0.99 (0.01)	**0.05** (0.00)
`rejectionABC` mean	**1.00** (0.01)	**0.05** (0.01)	**1.00** (0.01)	**0.05** (0.00)
`rejectionABCselect` mean	**1.00** (0.01)	**0.05** (0.01)	**1.00** (0.01)	**0.05** (0.00)
`copulaABCdrf` median	**1.00** (0.01)	0.05 (0.01)	0.99 (0.01)	**0.05** (0.00)
`rejectionABC` median	0.99 (0.01)	0.05 (0.01)	**1.00** (0.01)	**0.05** (0.00)
`rejectionABCselect` median	0.99 (0.01)	**0.04** (0.02)	**1.00** (0.01)	**0.05** (0.00)
`copulaABCdrf` modeMLE	0.97 (0.05)	**0.05** (0.02)	1.02 (0.06)	**0.05** (0.00)
`rejectionABCkern` modeMLE	**1.00** (0.04)	**0.05** (0.01)	**0.99** (0.05)	**0.05** (0.00)
`rejectionABCkern.select` modeMLE	0.99 (0.04)	0.06 (0.01)	**0.99** (0.05)	**0.05** (0.00)
`copulaABCdrf` s.d.	0.06 (0.00)	0.02 (0.00)	0.06 (0.00)	0.01 (0.00)
`rejectionABC` s.d.	0.06 (0.00)	0.02 (0.00)	0.06 (0.00)	0.01 (0.00)
`rejectionABCselect` s.d.	0.06 (0.00)	0.02 (0.00)	0.06 (0.00)	0.01 (0.00)
`copulaABCdrf`	d.f.	scale	d.f.	scale
Copula d.f. and scale	633.32 (474.75)	0.06 (0.05)	568.93 (456.18)	0.04 (0.07)
NLPA model	α	p	α	p
n	50	50	300	300
n_{sim}	25	25	100	100
true θ	1.20	0.02	1.20	0.02
`copulaABCdrf` mean	**1.24** (0.30)	0.01 (0.00)	**1.12** (0.09)	**0.02** (0.00)
`rejectionABC` mean	0.94 (0.30)	0.03 (0.00)	0.89 (0.17)	0.04 (0.01)
`rejectionABCselect` mean	1.37 (0.21)	**0.02** (0.00)	**1.12** (0.07)	**0.02** (0.00)
`copulaABCdrf` median	**1.23** (0.30)	0.01 (0.00)	**1.12** (0.08)	**0.02** (0.00)
`rejectionABC` median	0.90 (0.33)	0.03 (0.00)	0.87 (0.19)	0.04 (0.01)
`rejectionABCselect` median	1.36 (0.22)	**0.02** (0.00)	1.11 (0.07)	**0.02** (0.00)
`copulaABCdrf` modeMLE	**1.35** (0.50)	**0.01** (0.00)	1.11 (0.22)	**0.02** (0.00)
`rejectionABCkern` modeMLE	0.86 (0.43)	**0.03** (0.01)	0.82 (0.19)	0.04 (0.01)
`rejectionABCkern.select` modeMLE	1.42 (0.33)	**0.01** (0.01)	**1.12** (0.09)	**0.02** (0.00)
`copulaABCdrf` standard deviation (s.d.)	0.51 (0.02)	0.01 (0.00)	0.27 (0.04)	0.00 (0.00)
`rejectionABC` s.d.	0.48 (0.05)	0.01 (0.00)	0.25 (0.03)	0.01 (0.00)
`rejectionABCselect` s.d.	0.59 (0.06)	0.01 (0.00)	0.25 (0.04)	0.01 (0.00)
`copulaABCdrf`	d.f.	scale	d.f.	scale
Copula d.f. and scale	13.52 (9.78)	0.21 (0.25)	7.39 (9.64)	0.10 (0.10)

Note: **Bold** indicates the more accurate ABC method for the given estimator.

Table 9. MAE (MSE), mean percent interval coverage (95%(50%)c) for price model over 10 replicas, under varying n and n_{sim}.

		n = 50			n = 300	
n_{sim}	**Posterior**	k_0	p	n_{sim}	k_0	p
5	`copulaABCdrf` mean	**0.00** (**0.00**)	**0.06** (**0.00**)	30	**0.01** (**0.00**)	**0.10** (**0.01**)
	`rejectionABCselect` mean	**0.01**$_{-0.02}$(**0.00**)	**0.06**$_{-0.01}$(**0.00**)		**0.01** (**0.00**)	**0.11** (**0.01**)
	`copulaABCdrf` median	**0.01** (**0.00**)	**0.06** (**0.00**)		**0.01** (**0.00**)	**0.11** (**0.01**)
	`rejectionABCselect` median	**0.01**$_{-0.02}$(**0.00**)	**0.06**$_{-0.01}$(**0.00**)		**0.01** (**0.00**)	**0.11** (**0.01**)
	`copulaABCdrf` modeMLE	0.06 (**0.00**)	**0.07** (**0.00**)		0.05 (**0.00**)	**0.11** (**0.01**)
	`rejectionABCkern.select` modeMLE	**0.04**$_{-0.03}$(**0.00**)	**0.07**(**0.00**$_{-0.01}$)		**0.04** (**0.00**)	**0.11**$_{-0.01}$(**0.01**)
	`copulaABCdrf` 95%(50%)c	**1.00** (1.00)	**0.00** (**0.00**)		**1.00** (**1.00**)	**0.10** (**0.00**)
	`rejectionABCselect` 95%(50%)c	**0.90**$_{+0.10}$(**0.90**$_{+0.60}$)	**0.00** (**0.00**)		**1.00** (**1.00**)	0.00 (**0.00**)
16	`copulaABCdrf` mean	**0.00** (**0.00**)	**0.05** (**0.00**)	100	**0.01** (**0.00**)	**0.03** (**0.00**)
	`rejectionABCselect` mean	**0.01** (**0.00**)	**0.05** (**0.00**)		**0.01** (**0.00**)	**0.03** (**0.00**)
	`copulaABCdrf` median	**0.01** (**0.00**)	**0.05** (**0.00**)		**0.01** (**0.00**)	**0.03** (**0.00**)
	`rejectionABCselect` median	**0.01** (**0.00**)	**0.05** (**0.00**)		**0.01** (**0.00**)	**0.03** (**0.00**)
	`copulaABCdrf` modeMLE	0.06 (**0.00**)	**0.05** (**0.00**)		**0.05** (**0.00**)	**0.03** (**0.00**)
	`rejectionABCkern.select` modeMLE	**0.04**$_{+0.01}$(**0.00**)	**0.05** (**0.00**)		**0.05** (**0.00**)	**0.03** (**0.00**)
	`copulaABCdrf` 95%(50%)c	**1.00** (**1.00**)	**0.00** (**0.00**)		**1.00** (**1.00**)	**0.20** (**0.00**)
	`rejectionABCselect` 95%(50%)c	**1.00** (**1.00**)	**0.00** (**0.00**)		**1.00** (**1.00**)	0.00 (**0.00**)
25	`copulaABCdrf` mean	**0.00** (**0.00**)	**0.03** (**0.00**)	150	**0.01** (**0.00**)	**0.01** (**0.00**)
	`rejectionABCselect` mean	**0.00**$_{-0.01}$(**0.00**)	**0.03** (**0.00**)		**0.01** (**0.00**)	0.02 (**0.00**)
	`copulaABCdrf` median	**0.00** (**0.00**)	**0.03** (**0.00**)		**0.01** (**0.00**)	**0.01** (**0.00**)
	`rejectionABCselect` median	**0.01** (**0.00**)	**0.03** (**0.00**)		**0.01** (**0.00**)	0.02 (**0.00**)
	`copulaABCdrf` modeMLE	0.05 (**0.00**)	**0.03** (**0.00**)		0.05 (**0.00**)	**0.01** (**0.00**)
	`rejectionABCkern.select` modeMLE	**0.04** (**0.00**)	0.04$_{+0.01}$(**0.00**)		**0.03**$_{-0.01}$(**0.00**)	**0.01** (**0.00**)
	`copulaABCdrf` 95%(50%)c	**1.00** (**1.00**)	**0.40** (**0.00**)		**1.00** (**1.00**)	0.00 (**0.00**)
	`rejectionABCselect` 95%(50%)c	**1.00** (**1.00**)	0.20 (**0.00**)		**1.00** (**1.00**)	**0.30**$_{+0.20}$(**0.00**)
37	`copulaABCdrf` mean	**0.00** (**0.00**)	**0.02** (**0.00**)	225	**0.01** (**0.00**)	**0.01** (**0.00**)
	`rejectionABC` mean	**0.00** (**0.00**)	**0.02** (**0.00**)		**0.01** (**0.00**)	**0.01** (**0.00**)
	`copulaABCdrf` median	**0.00** (**0.00**)	**0.02** (**0.00**)		**0.01** (**0.00**)	**0.01** (**0.00**)
	`rejectionABC` median	**0.01** (**0.00**)	**0.02** (**0.00**)		**0.01** (**0.00**)	**0.01** (**0.00**)
	`copulaABCdrf` modeMLE	0.05 (**0.00**)	**0.02** (**0.00**)		**0.05** (**0.00**)	**0.01** (**0.00**)
	`rejectionABCkern.select` modeMLE	**0.04** (**0.00**)	**0.02** (**0.00**)		**0.05**$_{+0.02}$(0.00)	**0.01** (**0.00**)
	`copulaABCdrf` 95%(50%)c	**1.00** (**1.00**)	**0.80** (**0.10**)		**1.00** (**1.00**)	**0.60** (**0.00**)
	`rejectionABCselect` 95%(50%)c	**1.00** (**1.00**)	**0.80**$_{+0.10}$(**0.10**)		**1.00** (**1.00**)	0.40$_{-0.20}$(**0.00**)
50	`copulaABCdrf` mean	**0.01** (**0.00**)	**0.02** (**0.00**)	300	**0.01** (**0.00**)	**0.00** (**0.00**)
	`rejectionABCselect` mean	**0.01** (**0.00**)	**0.02** (**0.00**)		**0.01** (**0.00**)	**0.00** (**0.00**)
	`copulaABCdrf` median	**0.01** (**0.00**)	**0.02** (**0.00**)		**0.01** (**0.00**)	**0.00** (**0.00**)
	`rejectionABCselect` median	**0.01** (**0.00**)	**0.02** (**0.00**)		**0.01** (**0.00**)	**0.00** (**0.00**)
	`copulaABCdrf` modeMLE	0.05 (**0.00**)	0.12 (0.10)		0.05 (**0.00**)	**0.00** (**0.00**)
	`rejectionABCkern.select` modeMLE	**0.04**$_{+0.01}$(0.00)	**0.02** (**0.00**)		**0.03** (**0.00**)	**0.00** (**0.00**)
	`copulaABCdrf` 95%(50%)c	**1.00** (**1.00**)	**1.00** (**0.50**)		**1.00** (**1.00**)	**1.00** (**0.80**)
	`rejectionABCselect` 95%(50%)c	**1.00** (**1.00**)	0.80$_{-0.10}$(**0.50**$_{+0.10}$)		**1.00** (**1.00**)	**1.00** (0.90)

Note: Subscript is change in MAE (MSE) based on `drf` selecting the 2 best of the 3 summaries. **Bold** indicates the more accurate ABC method for the given estimator.

From the results of Table 10 on the NLPA model, it can be concluded that, overall, `copulaABCdrf` and `rejectionABC` performed similarly, while `rejectionABC`, based on using `drf` to pre-select two of the three most important summaries, produced results that often improved on results of `rejectionABC` using all three summaries. More specifically, for the parameter α and $n = 50$ nodes, `copulaABCdrf` slightly outperformed `rejectionABC` in MAE and MSE of posterior mean estimation and MAE of posterior median estimation. Also, `rejectionABC` slightly outperformed `copulaABCdrf` in MSE of posterior median estimation and in MAE and MSE in the estimation of the posterior mode and MLE. For the parameter p and $n = 50$ nodes, `copulaABCdrf` slightly outperformed `rejectionABC` in accuracy of the 50% credible interval estimation while not often being close to the nominal rate. In addition,

`rejectionABC` outperformed `copulaABCdrf` in MAE of posterior mean and median estimation and in MAE and MSE of the estimation of the posterior mode and the MLE, and in the estimation of the 95% posterior credible interval, while typically being close to the nominal rate, while, for the parameter α and $n = 300$ nodes, `copulaABCdrf` slightly outperformed `rejectionABC` in accuracy of the 50% credible interval estimation, while `rejectionABC` outperformed `copulaABCdrf` in terms of MAE and MSE in the estimation of the posterior mean, median, mode, and the MLE, and in the accuracy of estimation of the 50% posterior credible interval, while often being far from the nominal level. For the parameter p and $n = 300$ nodes, `copulaABCdrf` slightly outperformed `rejectionABC` in MAE in posterior mean estimation, MSE in the estimation of the posterior median, the posterior mode, and the MLE, and in accuracy in the estimation of the 50% posterior credible interval, while not always close to the nominal level. Also, `rejectionABC` outperformed `copulaABCdrf` in MAE of the estimation of the posterior mean, median, and 95% posterior credible interval coverage, while usually being close to the nominal level.

Table 10. MAE (MSE), mean credible interval coverage (95%(50%)c) for NLPA model over 10 replicas, varying n and n_{sim}.

		n = 50			n = 300	
n_{sim}	**Posterior**	α	p	n_{sim}	α	p
5	`copulaABCdrf` mean	**0.20** (**0.04**)	**0.07** (**0.00**)	30	0.67 (0.48)	0.02 (**0.00**)
	`rejectionABCselect` mean	$0.29_{+0.01}(0.09_{+0.01})$	$\mathbf{0.07}_{+0.01}(\mathbf{0.00})$		$\mathbf{0.07}_{-0.03}(\mathbf{0.01})$	**0.01** (**0.00**)
	`copulaABCdrf` median	0.37 (0.14)	**0.06** (**0.00**)		0.65 (0.46)	0.02 (**0.00**)
	`rejectionABCselect` median	$\mathbf{0.29}_{-0.16}(\mathbf{0.08}_{-0.12})$	$\mathbf{0.06}_{+0.01}(\mathbf{0.00})$		$\mathbf{0.07}_{-0.05}(\mathbf{0.01}_{-0.01})$	$\mathbf{0.01}_{-0.01}(\mathbf{0.00})$
	`copulaABCdrf` modeMLE	0.76 (0.68)	0.10 (0.01)		0.55 (0.37)	0.02 (**0.00**)
	`rejectionABCkern.select` modeMLE	$\mathbf{0.61}_{-0.27}(\mathbf{0.52}_{-0.25})$	$\mathbf{0.02}_{+0.01}(\mathbf{0.00})$		$\mathbf{0.21}_{-0.15}(\mathbf{0.06}_{-0.11})$	$\mathbf{0.01}_{-0.01}(\mathbf{0.00})$
	`copulaABCdrf`95%(50%)c	**1.00** (**1.00**)	**1.00** (**0.00**)		**1.00** (**0.00**)	0.00 (0.00)
	`rejectionABCselect` 95%(50%)c	**1.00** (**1.00**)	$\mathbf{0.90}_{-0.10}(\mathbf{0.00})$		**1.00** (**1.00**)	**1.00** ($\mathbf{0.30}_{-0.70}$)
16	`copulaABCdrf` mean	0.35 (0.16)	0.01 (**0.00**)	100	0.09 (**0.01**)	**0.00** (**0.00**)
	`rejectionABCselect` mean	$\mathbf{0.25}_{-0.14}(\mathbf{0.07}_{-0.14})$	$\mathbf{0.00}_{-0.01}(\mathbf{0.00})$		$\mathbf{0.08}_{-0.23}(\mathbf{0.01}_{-0.11})$	$\mathbf{0.00}_{-0.02}(\mathbf{0.00})$
	`copulaABCdrf` median	0.40 (0.21)	0.01 (**0.00**)		**0.08** (**0.01**)	**0.00** (**0.00**)
	`rejectionABCselect` median	$\mathbf{0.26}_{-0.18}(\mathbf{0.08}_{-0.19})$	$\mathbf{0.00}_{-0.01}(\mathbf{0.00})$		$\mathbf{0.09}_{-0.24}(\mathbf{0.01}_{-0.13})$	$\mathbf{0.00}_{-0.02}(\mathbf{0.00})$
	`copulaABCdrf` modeMLE	0.61 (0.53)	0.02 (**0.00**)		0.16 (0.05)	**0.00** (**0.00**)
	`rejectionABCkern.select` modeMLE	$\mathbf{0.34}_{-0.24}(\mathbf{0.14}_{-0.27})$	$\mathbf{0.01}_{+0.01}(\mathbf{0.00})$		$\mathbf{0.09}_{-0.29}(\mathbf{0.01}_{-0.17})$	$\mathbf{0.00}_{-0.02}(\mathbf{0.00})$
	`copulaABCdrf`95%(50%)c	**1.00** (**0.30**)	**1.00** (**0.30**)		**1.00** (**0.80**)	**1.00** (**0.00**)
	`rejectionABCselect` 95%(50%)c	**1.00** ($1.00_{+0.70}$)	$\mathbf{1.00}(1.00_{+0.10})$		$\mathbf{1.00}_{+0.10}(\mathbf{0.80}_{+0.80})$	$\mathbf{1.00}_{+0.20}(\mathbf{1.00}_{+1.00})$
25	`copulaABCdrf` mean	0.23 (0.08)	**0.01** (**0.00**)	150	0.12 (0.02)	**0.00** (**0.00**)
	`rejectionABCselect` mean	$\mathbf{0.22}_{-0.13}(\mathbf{0.07}_{-0.08})$	$\mathbf{0.01}_{-0.01}(\mathbf{0.00})$		$\mathbf{0.10}_{-0.14}(\mathbf{0.01}_{-0.06})$	$\mathbf{0.00}_{-0.01}(\mathbf{0.00})$
	`copulaABCdrf` median	**0.23** (0.08)	0.02 (**0.00**)		0.11 (0.02)	**0.00** (**0.00**)
	`rejectionABCselect` median	$0.24_{-0.14}(\mathbf{0.07}_{-0.12})$	$\mathbf{0.00}_{-0.01}(\mathbf{0.00})$		$\mathbf{0.10}_{-0.13}(\mathbf{0.01}_{-0.06})$	$\mathbf{0.00}_{-0.01}(\mathbf{0.00})$
	`copulaABCdrf` modeMLE	0.41 (0.25)	**0.01** (**0.00**)		**0.12** (**0.02**)	**0.00** (**0.00**)
	`rejectionABCkern.select` modeMLE	$\mathbf{0.34}_{-0.10}(0.15_{-0.13})$	**0.01** (**0.00**)		$\mathbf{0.12}_{-0.10}(\mathbf{0.02}_{-0.05})$	0.01 (**0.00**)
	`copulaABCdrf`95%(50%)c	**1.00** (**0.70**)	0.60 (0.00)		**1.00** (**0.50**)	**1.00** (**0.40**)
	`rejectionABCselect` 95%(50%)c	**1.00** ($1.00_{+0.70}$)	**1.00** ($\mathbf{0.90}_{+0.40}$)		$\mathbf{1.00}_{+0.30}(0.60_{+0.30})$	$\mathbf{1.00}_{+0.30}(1.00_{+0.90})$
37	`copulaABCdrf` mean	**0.21** (**0.05**)	0.01 (**0.00**)	225	**0.05** (**0.00**)	**0.00** (**0.00**)
	`rejectionABCselect` mean	$0.24_{-0.04}(0.08_{-0.04})$	$\mathbf{0.00}_{-0.01}(\mathbf{0.00})$		$0.06_{-0.09}(\mathbf{0.00}_{-0.03})$	**0.00** (**0.00**)
	`copulaABCdrf` median	**0.21** (**0.05**)	0.01 (**0.00**)		**0.05** (**0.00**)	**0.00** (**0.00**)
	`rejectionABCselect` median	$0.25_{-0.07}(0.09_{-0.07})$	$\mathbf{0.00}_{-0.01}(\mathbf{0.00})$		$\mathbf{0.06}_{-0.08}(\mathbf{0.00}_{-0.02})$	**0.00** (**0.00**)
	`copulaABCdrf` modeMLE	0.28 (0.11)	**0.01** (**0.00**)		0.11 (0.04)	**0.00** (**0.00**)
	`rejectionABCkern.select` modeMLE	$\mathbf{0.24}_{-0.24}(\mathbf{0.09}_{-0.21})$	**0.01** (**0.00**)		$\mathbf{0.06}_{-0.10}(\mathbf{0.01}_{-0.03})$	$\mathbf{0.00}_{-0.01}(\mathbf{0.00})$
	`copulaABCdrf`95%(50%)c	**1.00** (0.70)	0.50 (**0.10**)		**1.00** (**0.80**)	**1.00** (**0.90**)
	`rejectionABCselect` 95%(50%)c	**1.00** (**0.60**)	**1.00** ($0.90_{+0.10}$)		**1.00** ($\mathbf{0.80}_{+0.30}$)	$\mathbf{1.00}_{+0.20}(1.00_{+0.20})$

Table 10. *Cont.*

		n = 50			n = 300	
n_{sim}	**Posterior**	α	p	n_{sim}	α	p
50	`copulaABCdrf` mean	**0.11** (**0.02**)	**0.00** (**0.00**)	300	**0.04** (**0.00**)	**0.00** (**0.00**)
	`rejectionABCselect` mean	$0.17_{-0.05}(0.03_{-0.03})$	$\mathbf{0.00}_{-0.01}(\mathbf{0.00})$		$0.05_{-0.10}(\mathbf{0.00}_{-0.02})$	$\mathbf{0.00}_{-0.01}(\mathbf{0.00})$
	`copulaABCdrf` median	**0.10** (**0.02**)	**0.00** (**0.00**)		**0.04** (**0.00**)	**0.00** (**0.00**)
	`rejectionABCselect` median	$0.16_{-0.06}(0.03_{-0.04})$	$\mathbf{0.00}_{-0.01}(\mathbf{0.00})$		$0.05_{-0.09}(\mathbf{0.00}_{-0.02})$	$\mathbf{0.00}_{-0.01}(0.00)$
	`copulaABCdrf` modeMLE	0.26 (0.09)	0.12 (0.14)		0.07 (**0.01**)	**0.00** (**0.00**)
	`rejectionABCkern.select` modeMLE	$\mathbf{0.18}_{-0.05}(\mathbf{0.05}_{-0.03})$	**0.01** (**0.00**)		$\mathbf{0.05}_{-0.12}(\mathbf{0.01}_{-0.03})$	$\mathbf{0.00}_{-0.01}(0.00)$
	`copulaABCdrf`95%(50%)c	**1.00** (0.80)	**0.90** (**0.60**)		**1.00** (**0.80**)	**1.00** (**0.60**)
	`rejectionABCselect` 95%(50%)c	**1.00** ($\mathbf{0.70}_{+0.10}$)	**1.00** ($0.90_{+0.40}$)		**1.00** ($\mathbf{0.80}_{+0.20}$)	$\mathbf{1.00}_{+0.10}(1.00_{+0.40})$

Note: Subscript is change in MAE (MSE) based on `drf` selecting the 2 best of the 3 summaries. **Bold** indicates the more accurate ABC method for the given estimator.

Tables 9 and 10 show that, for each of the `copulaABCdrf` and `rejectionABC` methods, as n_{sim} gets smaller relative to n, roughly, and on the most part, the MAE and MSE of estimates tend to get smaller, and the 95% and 50% posterior credible intervals more closely approach their respective nominal rates. Intuitively, the summary statistics s become less sufficient as n_{sim} gets smaller relative to n.

3.5. Simulation Study: Duplication–Divergence Models

The DMC model and the DMR model each grow an undirected network of n nodes by starting with a seed network of $n' \geq 1$ nodes, and then repeats the following steps until the requisite number of nodes ($n = n'$) is reached: add new node ($n' \equiv n' + 1$), then add an edge between the new node and each neighbor of a randomly chosen existing node, $k \sim \mathcal{U}_{\{1,n'-1\}}$. Then, for the DMC model, the remaining steps are as follows: for each neighbor of the chosen node, randomly select either the edge between the chosen node and the neighbor or the edge between the new node and the neighbor, and remove that edge with probability q_{mod}; then, add an edge between the chosen node and the new node with probability q_{con}. Alternatively, for the DMR model, the remaining steps are as follows: each edge connected to the new node is removed independently with probability q_{del}, and an edge between any existing node and the new node is added with probability $q_{\text{new}}/(n'-1)$.

Now, we consider simulation studies that evaluate and compare the ability of `copulaABCdrf` (Algorithm 1) and `rejectionABC` to estimate the posterior distribution of the parameters of the DMC model, and the parameters of the posterior distribution of the DMR model. As before, the simulation studies are based on 10 replications of $n = 50$ node-undirected networks from the DMC model (undirected networks from the DMR model, respectively) and of $n = 300$ node-undirected networks simulated from the DMC model (undirected networks from the DMR model, respectively). These simulations are performed for each of five conditions of n_{sim}, being at least roughly 10%, 33%, 50%, 75%,, and 100% of n network nodes; specifically, the conditions $n_{\text{sim}} = 5, 16, 25, 37$, and 50 for $n = 50$ nodes, and $n_{\text{sim}} = 30, 100, 150, 225$, and 300 for $n = 300$ nodes. The true data-generating parameters for the DMC model were set as $\theta_0^{\mathsf{T}} = (q_{\text{mod}}, q_{\text{con}}) = (0.20, 0.10)$, while the true data-generating parameters of the DMR model were $\theta_0^{\mathsf{T}} = (q_{\text{del}}, q_{\text{new}}) = (0.20, 0.10)$.

For each network dataset x simulated from the DMC model (DMR model, respectively) and analyzed by `copulaABCdrf` and `rejectionABCselect`, the DMC model (DMR model, resp.) was assigned a uniform prior $(q_{\text{mod}}, q_{\text{con}}) \sim \mathcal{U}_{(0.15,0.35)}\mathcal{U}_{(0,1)}$ (uniform prior $(q_{\text{del}}, q_{\text{new}}) \sim \mathcal{U}_{(0.15,0.35)}\mathcal{U}_{(0,1)}$, resp.) using vector of summaries $s(\cdot)$ specified by network size invariant offset MPLEs of undirected network summary statistics of mean degree and triangles (of local and global average clustering coefficients, and degree assortativity, resp.). The MPLE summaries $s(x)$ of each network x used offsets based on n nodes ($n = 50$ or 300), while the MPLE summaries $s(y)$ of each network y of smaller size $n_{\text{sim}} < n$ (being $n_{\text{sim}} = 25$ for $n = 50$, and $n_{\text{sim}} = 100$ for $n = 300$) were simulated in each iteration of the ABC algorithm using offsets based on n_{sim} nodes.

Table 11 presents some representative detailed results of the simulation study of `copulaABCdrf` and `rejectionABC`, for the DMC and DMR models, for pairs of values of n_{sim} and n. Specifically, for the DMC model, $n_{\text{sim}} = 50$ for $n = 50$ and $n_{\text{sim}} = 300$ for $n = 300$; and, for the DMR model, $n_{\text{sim}} = 50$ for $n = 150$ and $n_{\text{sim}} = 300$ for $n = 300$. Tables 12 and 13 present the results of the MAE, MSE, and accuracy of 95% and 50% posterior interval credible interval estimates of `copulaABCdrf` and `rejectionABC` for the DMC and DMR models for each of the five n_{sim} conditions within each of $n = 50$ and $n = 300$.

In particular, since, for the DMR models, there were three summary statistics relative to the two models parameters, we not only considered for each model implementations of `rejectionABC` using all three summary statistics but also considered, for comparison purposes, each implementation of `rejectionABC` based on pre-selecting two of the most important of the three summary statistics according to the results obtained from a variable importance analysis based on training a `drf` multivariate regression on a reference table of simulated model parameters on all three summary statistics (covariate variables). For this network model, `drf` variable importance analyses found that, over all ten replicas of network datasets under the $n_{\text{sim}} = 16$ and $n = 50$ condition, the network size invariant offset MPLEs of global average clustering coefficients (importance measure of 1) and degree assortativity (importance measure of 0.60) were the two most important summary statistics. Meanwhile, for each of the other simulation conditions of n_{sim} and n, `drf` variable importance analyses found that, over all ten replicas of network datasets, the network size invariant offset MPLEs of undirected network summary statistics of local and global average clustering coefficients were always the most important two summaries of the three total summary statistics.

From the results of Table 12 on the DMC model, it can be concluded that, overall, `copulaABCdrf` and `rejectionABC` performed fairly similarly. Specifically, for the parameter q_{mod} and $n = 50$ nodes, `copulaABCdrf` slightly outperformed `rejectionABC` in MAE of posterior mean and median estimation and in accuracy in the estimation of the 50% credible interval, while usually being around the nominal rate. Also, `rejectionABC` slightly outperformed `copulaABCdrf` in MSE of posterior mean estimation and in MAE and MSE in the estimation of the posterior mode and MLE. For the parameter q_{con} and $n = 50$ nodes, `copulaABCdrf` outperformed `rejectionABC` in MAE and MSE of the estimation of the posterior mean, median, mode, and MLE, as well as accuracy in the estimation of the 95% credible interval and the 50% credible interval, while usually being around the nominal rate in each case, especially for $n_{\text{sim}} = 37$ and 50. For the parameter q_{mod} and $n = 300$ nodes, `copulaABCdrf` slightly outperformed `rejectionABC` in accuracy in the estimation of the 50% credible interval, while often not being around the nominal rate. Further, `rejectionABC` slightly outperformed `copulaABCdrf` in MSE of the estimation of the posterior mode and MLE. For the parameter q_{con} and $n = 300$ nodes, `copulaABCdrf` outperformed `rejectionABC` in MAE and MSE of the estimation of the posterior median, median, mode, and the MLE.

From the results of Table 13, on the DMR model, it can be concluded that, overall, `copulaABCdrf` and `rejectionABC` performed rather similarly. Meanwhile, `rejectionABC`, based on using `drf` to pre-select two of the three most important summaries, produced results that often noticeably improved on the results of `rejectionABC` using all three summaries. More specifically, for the parameter q_{del} and $n = 50$ nodes, `copulaABCdrf` slightly outperformed `rejectionABC` in MSE of posterior mean estimation, MAE and MSE of posterior mode estimation, and in the accuracy of the 50% credible interval estimation, sometimes being close to the nominal rate. For the parameter q_{new} and $n = 50$ nodes, `rejectionABC` outperformed `copulaABCdrf` in MAE and MSE in the estimation of the posterior mean, median, and mode, and MSE in posterior median estimation, and slightly outperformed in the accuracy of the 95% credible interval estimation while usually being close to the nominal rate. Also, `copulaABCdrf` outperformed `rejectionABC` in the accuracy of the 50% credible interval estimation while usually being not close to the nominal rate. For the parameter q_{del}

and $n = 300$ nodes, `copulaABCdrf` slightly outperformed `rejectionABC` in MSE of posterior mean estimation, MAE in the estimation of the posterior median, mode, and the MLE, and in the accuracy of the 50% credible interval estimation, while sometimes being rather near the nominal rate. Meanwhile, `rejectionABC` slightly outperformed `copulaABCdrf` in terms of MSE in the estimation of the posterior mode and MLE. Finally, for the parameter q_{new} and $n = 300$ nodes, `rejectionABC` slightly outperformed `copulaABCdrf` in MAE and MSE in the estimation of the posterior mode and the MLE, and in the accuracy of 95% credible interval estimation, while usually being close to the nominal rate.

Tables 12 and 13 show that, for each of the `copulaABCdrf` and `rejectionABC` methods, as n_{sim} gets smaller relative to n, at least roughly and on the most part, the MAE and MSE of estimates tend to get smaller, and the 95% and 50% posterior credible intervals more closely approach their respective nominal rates. Intuitively, the summary statistics s become less sufficient as n_{sim} gets smaller relative to n.

Table 11. Simulation study results of `copulaABCdrf` for the DMC and DMR models. Mean (standard deviation) of posterior estimates over 10 replicas, under some varying n and n_{sim}.

DMC model	q_{mod}	q_{con}	q_{mod}	q_{con}
n	50	50	300	300
n_{sim}	50	50	300	300
true θ	0.20	0.10	0.20	0.10
`copulaABCdrf` mean	**0.24** (0.02)	**0.16** (0.05)	**0.22** (0.03)	**0.16** (0.05)
`rejectionABC` mean	0.26 (0.01)	0.32 (0.17)	0.26 (0.01)	0.34 (0.21)
`copulaABCdrf` median	**0.24** (0.03)	**0.15** (0.05)	**0.22** (0.04)	**0.14** (0.04)
`rejectionABC` median	0.26 (0.01)	0.32 (0.18)	0.27 (0.01)	0.33 (0.22)
`copulaABCdrf` modeMLE	**0.25** (0.06)	**0.17** (0.09)	**0.26** (0.04)	**0.13** (0.09)
`rejectionABCkern` modeMLE	**0.25** (0.05)	0.32 (0.23)	0.28 (0.04)	0.38 (0.31)
`copulaABCdrf` s.d.	0.05 (0.01)	0.07 (0.02)	0.04 (0.01)	0.07 (0.04)
`rejectionABC` s.d.	0.06 (0.00)	0.14 (0.08)	0.06 (0.00)	0.13 (0.07)
`copulaABCdrf`	d.f.	scale	d.f.	scale
Copula d.f. and scale	339.94 (456.81)	0.09 (0.24)	131.75 (289.23)	0.21 (0.30)
DMR model	q_{del}	q_{new}	q_{del}	q_{new}
n	50	50	300	300
n_{sim}	50	50	150	150
true θ	0.20	0.10	0.20	0.10
`copulaABCdrf` mean	**0.24** (0.03)	0.24 (0.19)	0.26 (0.01)	0.15 (0.11)
`rejectionABC` mean	**0.24** (0.02)	**0.23** (0.16)	**0.25** (0.02)	**0.13** (0.04)
`rejectionABCselect` mean	0.25 (0.02)	0.27 (0.23)	**0.25** (**0.00**)	0.14 (0.11)
`copulaABCdrf` median	**0.24** (0.03)	0.23 (0.21)	0.26 (0.01)	0.14 (0.11)
`rejectionABC` median	**0.24** (0.03)	**0.21** (0.18)	**0.25** (0.02)	**0.11** (0.03)
`rejectionABCselect` median	0.25 (0.03)	0.26 (0.26)	0.26 (0.01)	0.20 (0.10)
`copulaABCdrf` modeMLE	**0.23** (0.06)	0.28 (0.20)	**0.24** (0.05)	0.19 (0.14)
`rejectionABCkern` modeMLE	0.25 (0.05)	**0.22** (0.24)	**0.24** (0.05)	0.09 (0.03)
`rejectionABCkern.select` modeMLE	0.27 (0.05)	**0.22** (0.29)	**0.26** (0.03)	**0.10** (0.13)
`copulaABCdrf` s.d.	0.05 (0.01)	0.12 (0.04)	0.06 (0.00)	0.08 (0.03)
`rejectionABC` s.d.	0.05 (0.01)	0.12 (0.05)	0.06 (0.01)	0.09 (0.02)
`rejectionABCselect` s.d.	0.06 (0.01)	0.13 (0.03)	0.06 (0.00)	0.08 (0.03)
`copulaABCdrf`	d.f.	scale	d.f.	scale
Copula d.f. and scale	426.42 (425.06)	-0.05 (0.19)	440.06 (399.70)	-0.07 (0.06)

Note: **Bold** indicates the more accurate ABC method for the given estimator.

Table 12. MAE (MSE), mean credible interval coverage (95%(50%)c) for the DMC model over 10 replicas, under varying conditions of n and n_{sim}.

		n = 50			n = 300	
n_{sim}	**Posterior**	$\mathbf{q}_{\text{mod}}$	q_{con}	n_{sim}	q_{mod}	q_{con}
5	`copulaABCdrf` mean	0.05 (**0.00**)	**0.57** (**0.33**)	30	0.04 (**0.00**)	**0.69** (**0.52**)
	`rejectionABC` mean	**0.04** (**0.00**)	0.61 (0.39)		**0.03** (**0.00**)	0.71 (0.56)
	`copulaABCdrf` median	0.05 (**0.00**)	**0.60** (**0.37**)		**0.04** (**0.00**)	**0.73** (**0.59**)
	`rejectionABC` median	**0.04** (**0.00**)	0.65 (0.43)		**0.04** (**0.00**)	0.74 (0.60)
	`copulaABCdrf` modeMLE	0.05 (0.00)	**0.61** (**0.42**)		0.06 (0.01)	**0.71** (**0.57**)
	`rejectionABCkern` modeMLE	**0.03** (**0.00**)	0.75 (0.58)		**0.04** (**0.00**)	0.76 (0.64)
	`copulaABCdrf` 95%(50%)c	**1.00** (**0.60**)	**0.00** (**0.00**)		**1.00** (**0.30**)	**0.20** (**0.10**)
	`rejectionABC` 95%(50%)c	**1.00** (0.90)	**0.00** (**0.00**)		**1.00** (0.00)	0.10 (**0.10**)
16	`copulaABCdrf` mean	**0.05** (**0.00**)	**0.24** (**0.09**)	100	0.05 (**0.00**)	0.57 (0.42)
	`rejectionABC` mean	**0.05** (**0.00**)	0.31 (0.14)		**0.04** (**0.00**)	**0.53** (**0.37**)
	`copulaABCdrf` median	**0.05** (**0.00**)	**0.23** (**0.09**)		**0.05** (**0.00**)	0.59 (0.44)
	`rejectionABC` median	**0.05** (**0.00**)	0.31 (0.14)		**0.05** (**0.00**)	**0.55** (**0.40**)
	`copulaABCdrf` modeMLE	**0.06** (**0.01**)	0.30 (**0.14**)		0.08 (0.01)	0.61 (0.46)
	`rejectionABCkern` modeMLE	0.07 (**0.01**)	**0.29** (**0.14**)		**0.06** (**0.00**)	**0.59** (**0.45**)
	`copulaABCdrf` 95%(50%)c	**1.00** (0.30)	**0.80** (**0.00**)		**0.80** (**0.30**)	**0.20** (**0.20**)
	`rejectionABC` 95%(50%)c	**1.00** (**0.40**)	0.60 (**0.00**)		0.50 (0.10)	**0.20** (0.10)
25	`copulaABCdrf` mean	**0.05** (**0.00**)	**0.10** (**0.01**)	150	**0.05** (**0.00**)	**0.30** (**0.18**)
	`rejectionABC` mean	**0.05** (**0.00**)	0.18 (0.04)		**0.05** (**0.00**)	0.31 (0.19)
	`copulaABCdrf` median	**0.05** (**0.00**)	**0.09** (**0.01**)		**0.05** (0.00)	**0.30** (**0.19**)
	`rejectionABC` median	0.06 (**0.00**)	0.16 (0.04)		0.06 (**0.00**)	0.32 (0.20)
	`copulaABCdrf` modeMLE	0.10 (**0.01**)	**0.13** (**0.02**)		**0.06** (**0.01**)	**0.32** (**0.22**)
	`rejectionABCkern` modeMLE	**0.07** (**0.01**)	0.18 (0.08)		0.07 (**0.01**)	0.34 (0.23)
	`copulaABCdrf` 95%(50%)c	**1.00** (**0.50**)	**1.00** (**0.30**)		**0.80** (**0.30**)	0.40 (0.10)
	`rejectionABC` 95%(50%)c	**1.00** (0.20)	**0.90** (0.10)		**0.80** (0.00)	**0.60** (**0.20**)
37	`copulaABCdrf` mean	**0.04** (**0.00**)	**0.11** (**0.03**)	225	**0.06** (**0.00**)	**0.05** (**0.00**)
	`rejectionABC` mean	0.05 (**0.00**)	0.36 (0.16)		0.07 (**0.00**)	0.10 (0.03)
	`copulaABCdrf` median	**0.04** (**0.00**)	**0.10** (**0.03**)		**0.07** (**0.00**)	**0.05** (**0.00**)
	`rejectionABC` median	0.05 (**0.00**)	0.36 (0.16)		**0.07** (0.01)	0.10 (0.03)
	`copulaABCdrf` modeMLE	0.07 (0.01)	**0.12** (**0.02**)		**0.08** (**0.01**)	**0.04** (**0.00**)
	`rejectionABCkern` modeMLE	**0.05** (**0.00**)	0.38 (0.20)		**0.08** (**0.01**)	0.10 (0.03)
	`copulaABCdrf` 95%(50%)c	**1.00** (**0.50**)	**0.90** (**0.50**)		**1.00** (**0.10**)	0.80 (**0.40**)
	`rejectionABC` 95%(50%)c	**1.00** (0.60)	0.20 (0.10)		**1.00** (0.00)	**0.90** (**0.40**)
50	`copulaABCdrf` mean	**0.04** (**0.00**)	**0.06** (**0.01**)	300	**0.03** (**0.00**)	**0.06** (**0.01**)
	`rejectionABC` mean	0.06 (**0.00**)	0.22 (0.08)		0.06 (**0.00**)	0.24 (0.10)
	`copulaABCdrf` median	**0.05** (**0.00**)	**0.05** (**0.00**)		**0.03** (**0.00**)	**0.05** (**0.00**)
	`rejectionABC` median	0.06 (**0.00**)	0.22 (0.08)		0.07 (**0.00**)	0.24 (0.10)
	`copulaABCdrf` modeMLE	0.06 (0.01)	**0.08** (**0.01**)		**0.06** (**0.00**)	**0.06** (**0.01**)
	`rejectionABCkern` modeMLE	**0.05** (**0.00**)	0.22 (0.10)		0.09 (0.01)	0.30 (0.17)
	`copulaABCdrf` 95%(50%)c	**1.00** (**0.60**)	**1.00** (**0.50**)		**1.00** (**0.80**)	**1.00** (**0.30**)
	`rejectionABC` 95%(50%)c	**1.00** (0.20)	**0.90** (0.10)		**1.00** (0.10)	0.60 (**0.30**)

Note: **Bold** indicates the more accurate ABC method for the given estimator.

3.6. Real Citation Network Analysis

With Algorithm 1 validated by simulation studies, we apply this algorithm to analyze real-life datasets. Specifically, we now analyze a large binary directed citation `HepPh` network dataset, an arXiv High Energy Physics paper citation network of n = 28,093 papers (nodes) and 3,148,447 citations (directed edges) among them [163] (this dataset was obtained from https://networkrepository.com/cit.php and accessed on 20 January 2024). This dataset is analyzed by the ERGM model and by the price model, each model using a vector of summary statistics $s(\cdot)$ that are the network size invariant offset MPLEs (resp.) of geometrically weighted indegree distribution (`gwidegree`), its decay (degree weighting) parameter estimate (`gwidegree.decay`), and `triangle` count. The ERGM, based on network sufficient statistics `gwidegree`, `gwidegree.decay`, and `triangle`, was assigned a trivariate

normal prior $\theta \sim \mathcal{N}(0, 10I_3)$. The price model was assigned priors $k_0 \sim \mathcal{U}_{(0.9,1.1)}$ and $p \sim \mathcal{U}_{(0,0.20)}$.

Table 13. MAE (MSE), mean credible interval coverage (95%(50%)c) for DMR model over 10 replicas, varying n and n_{sim}.

		n = 50			n = 300	
n_{sim}	**Posterior**	q_{del}	q_{new}	n_{sim}	q_{del}	q_{new}
5	`copulaABCdrf` mean	0.06 (**0.00**)	0.49 (0.25)	30	0.06 (**0.00**)	0.21 (0.05)
	`rejectionABCselect` mean	**0.05**$_{-0.01}$(**0.00**)	**0.43**$_{-0.06}$(**0.18**$_{-0.06}$)		**0.05**$_{-0.01}$(**0.00**)	**0.17**$_{-0.01}$(**0.03**)
	`copulaABCdrf` median	0.06 (**0.00**)	0.50 (0.26)		0.06 (**0.00**)	0.19 (0.04)
	`rejectionABCselect` median	**0.05**$_{-0.02}$(**0.00**)	**0.42**$_{-0.07}$(**0.18**$_{-0.07}$)		**0.05**$_{-0.01}$(**0.00**)	**0.14**$_{-0.02}$(**0.03**)
	`copulaABCdrf` modeMLE	**0.07** (**0.01**)	0.53 (0.32)		0.06 (0.01)	0.23 (0.07)
	`rejectionABCkern.select` modeMLE	0.08$_{-0.02}$(**0.01**)	**0.34**$_{-0.20}$(**0.12**$_{-0.18}$)		**0.04**$_{-0.02}$(**0.00**)	**0.09**$_{-0.04}$(**0.02**)
	`copulaABCdrf`95%(50%)c	**1.00** (0.00)	0.00 (**0.00**)		**1.00** (0.00)	0.70 (0.00)
	`rejectionABCselect` 95%(50%)c	**1.00**(**0.10**$_{+0.10}$)	**0.60**$_{+0.60}$(**0.00**)		**1.00** (**0.20**$_{+0.10}$)	**0.90** (**0.30**$_{+0.30}$)
16	`copulaABCdrf` mean	**0.05** (**0.00**)	0.29 (0.10)	100	**0.05** (**0.00**)	**0.12** (**0.02**)
	`rejectionABCselect` mean	0.06$_{+0.01}$(0.00)	**0.25**$_{-0.03}$(**0.07**$_{-0.02}$)		**0.05** (**0.00**)	**0.12**$_{+0.02}$(**0.02**$_{+0.01}$)
	`copulaABCdrf` median	**0.06** (**0.00**)	0.27 (0.09)		**0.05** (**0.00**)	**0.10** (**0.02**)
	`rejectionABCselect` median	**0.06** (**0.00**)	**0.23**$_{-0.03}$(**0.06**$_{-0.02}$)		0.06$_{+0.01}$(**0.00**)	**0.10**$_{+0.02}$(**0.02**$_{+0.01}$)
	`copulaABCdrf` modeMLE	**0.05** (**0.00**)	0.32 (0.17)		**0.04** (**0.00**)	**0.10** (**0.02**)
	`rejectionABCkern.select` modeMLE	0.06$_{-0.02}$(0.01)	**0.17**$_{-0.06}$(**0.06**$_{-0.02}$)		0.05$_{-0.01}$(**0.00**$_{-0.01}$)	0.11$_{+0.05}$(**0.02**$_{+0.01}$)
	`copulaABCdrf`95%(50%)c	**1.00** (**0.30**)	0.60 (**0.00**)		**1.00** (0.30)	**0.80** (**0.40**)
	`rejectionABCselect` 95%(50%)c	**1.00** (0.00$_{-0.20}$)	**0.80** (**0.00**)		**1.00** (**0.40**$_{-0.10}$)	**0.80** (**0.40**)
25	`copulaABCdrf` mean	**0.05** (**0.00**)	0.17 (0.04)	150	0.06 (**0.00**)	0.06 (**0.01**)
	`rejectionABCselect` mean	**0.05** (**0.00**)	**0.16**$_{-0.02}$(**0.03**$_{-0.01}$)		**0.05**$_{-0.01}$(**0.00**)	**0.05**$_{-0.02}$(**0.01**)
	`copulaABCdrf` median	**0.05** (**0.00**)	**0.14** (**0.03**)		**0.06** (**0.00**)	0.06 (**0.01**)
	`rejectionABCselect` median	0.06 (**0.00**)	**0.14**$_{-0.02}$(**0.03**$_{-0.01}$)		**0.06** (**0.00**)	**0.06**$_{+0.01}$(**0.01**)
	`copulaABCdrf` modeMLE	**0.07** (**0.01**)	0.17 (0.08)		**0.05** (**0.00**)	0.11 (0.03)
	`rejectionABCkern.select` modeMLE	**0.07**$_{+0.01}$(**0.01**)	**0.11**$_{-0.02}$(**0.04**)		0.06$_{-0.01}$(**0.00**$_{-0.01}$)	**0.09**$_{+0.03}$(**0.01**)
	`copulaABCdrf`95%(50%)c	**1.00** (**0.30**)	**0.90** (**0.10**)		**1.00** (**0.30**)	**0.90** (**0.70**)
	`rejectionABCselect` 95%(50%)c	**1.00** (0.00$_{-0.40}$)	**1.00** (0.00$_{-0.10}$)		**1.00** (0.00)	**0.90** (0.80$_{-0.10}$)
37	`copulaABCdrf` mean	**0.05** (**0.00**)	**0.10** (**0.01**)	225	**0.05** (**0.00**)	**0.04** (**0.00**)
	`rejectionABCselect` mean	**0.05** (**0.00**)	0.11$_{+0.01}$(0.02$_{+0.01}$)		**0.05** (**0.00**)	0.06$_{+0.02}$(**0.00**)
	`copulaABCdrf` median	**0.05** (**0.00**)	**0.08** (**0.01**)		**0.05** (**0.00**)	**0.04** (**0.00**)
	`rejectionABCselect` median	0.06$_{+0.01}$(**0.00**)	0.09$_{+0.02}$(**0.01**)		**0.05** (**0.00**)	0.06$_{+0.03}$(**0.00**)
	`copulaABCdrf` modeMLE	**0.06** (**0.00**)	0.10 (0.02)		0.08 (0.01)	0.11 (0.02)
	`rejectionABCkern.select` modeMLE	0.07$_{+0.04}$(0.01$_{+0.01}$)	**0.06**$_{+0.01}$(**0.01**$_{+0.01}$)		**0.06** (**0.00**$_{-0.01}$)	**0.06**$_{+0.01}$(**0.00**)
	`copulaABCdrf`95%(50%)c	**1.00** (**0.60**)	**1.00** (**0.40**)		**1.00** (**0.40**)	**1.00** (0.80)
	`rejectionABCselect` 95%(50%)c	**1.00** (0.00$_{-0.40}$)	**0.90**$_{-0.10}$(0.8$_{+0.50}$)		**1.00** (0.00$_{-0.40}$)	**1.00** (**0.50**$_{-0.40}$)
50	`copulaABCdrf` mean	**0.05** (**0.00**)	**0.14** (**0.05**)	300	**0.05** (**0.00**)	**0.04** (**0.00**)
	`rejectionABCselect` mean	**0.05** (**0.00**)	0.17$_{+0.04}$(0.07$_{+0.03}$)		0.06$_{+0.01}$(0.00)	0.07$_{+0.03}$(0.01$_{+0.01}$)
	`copulaABCdrf` median	**0.05** (**0.00**)	**0.13** (**0.06**)		**0.05** (**0.00**)	**0.05** (**0.00**)
	`rejectionABCselect` median	**0.05** (**0.00**)	0.16$_{+0.04}$(0.08$_{+0.04}$)		0.06$_{+0.01}$(**0.00**)	0.07$_{+0.04}$(0.01$_{+0.01}$)
	`copulaABCdrf` modeMLE	**0.05** (**0.00**)	0.19 (**0.07**)		**0.06** (**0.01**)	0.08 (**0.01**)
	`rejectionABCkern.select` modeMLE	0.08$_{+0.02}$(0.01$_{+0.01}$)	**0.14**$_{+0.01}$(0.09$_{+0.02}$)		0.07$_{+0.02}$(**0.01**$_{+0.01}$)	**0.07**$_{+0.04}$(**0.01**$_{+0.01}$)
	`copulaABCdrf`95%(50%)c	**1.00** (**0.50**)	**0.90** (**0.50**)		**1.00** (**0.40**)	**1.00** (**0.60**)
	`rejectionABCselect` 95%(50%)c	**1.00** (0.20$_{-0.20}$)	0.80$_{-0.10}$(0.70$_{+0.20}$)		**1.00** (0.10$_{-0.50}$)	**1.00** (0.30$_{-0.70}$)

Note: Subscript is change in MAE (MSE) based on `drf` selecting the 2 best of the 3 summaries. **Bold** indicates the more accurate ABC method for the given estimator.

The MPLEs for the three network summaries (resp.) could not be computed directly on this large network; these computations were prohibitively long. Therefore, these three MPLEs were estimated by their geometric median over 100,000 subsamples of the 28,093 papers (nodes), each subsample inducing a subgraph of size $\lfloor\sqrt{28,093}\rfloor = 167$, providing an outlier-robust divide-and-conquer (DAC) subsampling MPLE estimator [159], resulting in the following MPLEs (resp.): for `gwidegree`, −0.8418461; for `gwidegree.decay`, 0.2365466; for `triangle`: 1.6904300. Recall from Section 3.2 that the MPLE is the MLE for logistic

regression, and that they are both consistent over a growing number of networks, observed from the same set of fixed nodes (while the MPLE assumes dyadic independence of the edge observations). These three DAC MPLEs then specified the vector of summary statistics $s(x)$ for the real observed citation `HepPh` network dataset, analyzed by each of the ERGM and price model using Algorithm 1, with each model simulating a network of size $n_{\text{sim}} = 167$ in each algorithm iteration.

Table 14 presents the posterior estimates and MLEs obtained from `copulaABCdrf`, the ABC algorithm. Running Step 2 of this algorithm for model selection output a posterior probability estimate of 0.99 for the price model compared to the ERGM. This table also presents the results of `rejectionABC` for the ERGM and for the price model based on all three network summary statistics. In addition, results are presented for the price model after pre-selecting the most important summary statistics (variable) of the three total summaries based on a `drf` regression fit to the reference table, regressing the prior parameter samples on the samples of all three summary statistics. It was found that, among all three network statistics, those being the network size (n) invariant MPLEs of `gwidegree` (importance measure of 0.14), `gwidegree.decay` (importance 0.07), and `triangle` count (importance 0.79), the first and third of these statistics were most important and thus pre-selected. Table 14 shows that, for the price model, the posterior-based parameter estimates were similar across all `copulaABCdrf` and `rejectionABC` methods, while, for the ERGM, they were less similar.

Table 14. ABC posterior estimates for the price model and for the ERGM, obtained from the arXiv High Energy Physics paper citation network dataset.

	Price Model		**ERGM**		
	k_0	p	θ **Gwidegree**	θ **Decay**	θ **Triangles**
`copulaABCdrf` mean	1.01	0.01	−4.46	3.34	−1.07
`rejectionABC` mean	1.01	0.01	−1.49	−0.72	−0.58
`rejectionABCselect` mean	1.01	0.01			
`copulaABCdrf` median	1.02	0.01	−4.96	3.66	−0.58
`rejectionABC` median	1.02	0.01	−1.86	−1.45	−0.53
`rejectionABCselect` median	1.02	0.01			
`copulaABCdrf` mode	1.08	0.01	−1.76	−0.80	−4.59
`rejectionABC` mode	1.02	0.01	−2.18	4.49	−0.51
`rejectionABCkern.select` mode	1.02	0.01			
`copulaABCdrf` MLE	1.08	0.01	−4.82	6.00	−0.58
`rejectionABC` MLE	1.02	0.01	5.39	−5.86	−7.30
`rejectionABCkern.select` MLE	1.02	0.01			
`copulaABCdrf` s.d.	0.05	0.00	2.41	2.62	2.06
`rejectionABC` s.d.	0.06	0.003	2.81	3.29	2.85
`rejectionABCselect` s.d.	0.06	0.003			
`copulaABCdrf` 50%	(0.97, 1.06)	(0.01, 0.01)	(−6.09, −2.81)	(2.23, 4.96)	(−1.17, −0.23)
`rejectionABC` 50%	(0.97, 1.05)	(0.01, 0.01)	(−2.62, 0.51)	(−3.21, 1.92)	(−1.99, 1.07)
`rejectionABCselect` 50%	(0.97, 1.05)	(0.01, 0.01)			
`copulaABCdrf` 95%	(0.91, 1.09)	(0.00, 0.02)	(−9.12, 0.73)	(−2.08, 6.80)	(−6.09, 3.22)
`rejectionABC` 95%	(0.90, 1.09)	(0.003, 0.01)	(−6.60, 4.00)	(−5.44, 5.86)	(−6.12, 4.22)
`rejectionABCselect` 95%	(0.90, 1.09)	(0.003, 0.01)			
`copulaABCdrf`	d.f.	scale	d.f.	scale matrix	
Copula d.f. and scale	50.42	−0.20	6.31	θ decay	θ triangles
			θ gwidegree	−0.60	0.30
			θ decay		−0.28

3.7. Real Multilayer Network Analysis

The multilayer, weighted, and directed `BostonBomb2013` Twitter network dataset consists of n = 4,377,184 persons (nodes) and 9,480,331 count-weighted edges across $L = 3$ layers: retweets, mentions, and replies occurring between dates 15 April 2013 and 22 April 2013 of the Boston Bombing Attacks of 2013 [164] (the dataset was obtained from https://manliodedomenico.com/data.php and accessed on 26 January 2024). Using Algorithm 1, the `BostonBomb2013` network dataset was analyzed by a three-layer ERGM with Poisson reference measure b,

$$f(X = x \mid \theta) = \prod_{l=1}^{L=3} f(X_l = x_l \mid \theta) = \frac{b(x_l) \exp\left\{\sum_{l=1}^{L=3} \theta_l^\intercal h_l(x_l)\right\}}{\prod_{l=1}^{L=3} \sum_{x' \in \mathcal{X}} b(x') \exp\left\{\sum_{l=1}^{L=3} \theta_l^\intercal h_l(x')\right\}},$$

which was assigned a multivariate normal prior $\theta \sim \mathcal{N}(0, 10I_{18})$ for 18 parameters θ of six network sufficient statistics (vector h_l) specified for each of the three layers, namely (1) indicator of edge weight equal to 1 in the given layer (`equalto(1)`); (2) indicator of edge weight greater than 1 in the given layer; (3) minimum mutuality (`mutual.min`), which is analogous to mutuality for binary networks; (4) triad closure represented by transitive weights with twopath = min, combine = sum, and affect = min(`transitiveweights.min.sum.min`), which is analogous to triangle count from a binary network; (5) triad closure represented by transitive weights with twopath = min, combine = max, and affect = min(`transitiveweights.min.max.min`), which is analogous to a count of transitive ties from a binary network; and, finally, (6) the term `CMP`, which specifies the reference measure b by the Conway–Maxwell–Poisson (CMP) distribution for the count-weighted edges, with corresponding coefficient parameter θ_{CMP}, which controls the degree of dispersion relative to Poisson distribution. In particular, $\theta_{\text{CMP}} = 0$ defines the Poisson distribution; $\theta_{\text{CMP}} < 0$ defines a more underdispersed distribution; the value $\theta_{\text{CMP}} \to -\infty$ leads to the Bernoulli distribution; $\theta_{\text{CMP}} > 0$ defines a more overdispersed distribution; and $\theta_{\text{CMP}} = 1$ corresponds to the geometric distribution, being the most overdispersed. More details about these network statistics of valued networks and the CMP distribution for the ERGM are provided elsewhere [124].

These 18 ERGM sufficient statistics could not be easily computed on the `BostonBomb2013` network dataset due to its massive size. Therefore, these sufficient statistics were approximated by computing 21 (mostly) UMP summary statistics $s(x)$, each of which was computable on the full network for each of the three network layers. The values of the summaries $s(x)$ of the `BostonBomb2013` network dataset are shown in Table 15.

Table 15. Summary statistics $s(x)$ of the `BostonBomb2013` network dataset.

Summary Statistic	Layer 1	Layer 2	Layer 3
`meanIndegree`	1.35253692785	0.6141674647	0.1590769317
`varIndegree`	8474.74052782986	1804.0675198631	31.5665679809
`meanOutdegree`	1.35253692785	0.6141674647	0.1590769317
`varOutdegree`	7.61861428910	4.5875615155	0.8667280838
`wClusteringCoef`	0.00008622621	0.0004401573	0.0001867074
`assortativityDegree`	−0.06229008973	−0.0233535797	−0.0165706996
`reciprocity`	0.00379811641	0.0379874337	0.0625188215

The `copulaABCdrf` (Algorithm 1) was run on this model, based on these network data summaries $s(x)$ and based on summaries $s(y)$ of each network y of $n_{\text{sim}} = 500$ nodes simulated from the ERGM, conditionally on given proposed model parameters, in each iteration of this algorithm. The posterior distribution and MLE estimates delivered by the algorithm are shown in Table 16. This table also shows the posterior parameter estimates of `rejectionABC` using the same summary statistics and n_{sim}. The posterior estimates noticeably differed between these two ABC methods.

Table 16. ABC posterior estimates, 3-layer-valued ERGM parameters, from `BostonBomb2013` network dataset.

Layer 1	Method	Mean	Median	Mode	MLE	S.D.	2.5%	25%	75%	97.5%
`equalto.1.pm.0`	`copulaABCdrf`	−5.65	−5.42	−4.07	−4.69	1.38	−9.07	−6.30	−4.69	−3.59
	`rejectionABC`	−3.40	−3.73	−3.65	−5.55	2.59	−7.45	−4.56	−2.96	3.16
`greaterthan.1`	`copulaABCdrf`	−2.85	−3.18	−0.67	1.33	2.74	−7.72	−4.77	−1.06	3.06
	`rejectionABC`	−1.47	−1.56	−1.95	−13.22	3.37	−7.51	−3.33	0.63	4.54
`mutual.min`	`copulaABCdrf`	−3.85	−3.63	−2.43	0.12	2.20	−8.72	−5.27	−2.42	0.14
	`rejectionABC`	−0.17	−0.13	0.13	4.18	2.71	−5.56	−1.86	1.65	5.07
`tw.min.sum.min`	`copulaABCdrf`	0.05	0.18	−2.04	1.24	3.25	−6.71	−2.30	2.38	5.74
	`rejectionABC`	3.06	2.84	0.12	1.80	2.31	−1.59	1.76	4.67	7.54
`tw.min.max.min`	`copulaABCdrf`	−1.88	−1.90	−0.33	−1.30	2.89	−7.29	−3.84	0.06	3.78
	`rejectionABC`	2.98	3.01	2.32	5.96	2.36	−1.05	1.48	4.31	7.24
`CMP`	`copulaABCdrf`	−2.15	−2.24	3.29	5.45	2.66	−7.33	−3.89	−0.38	3.04
	`rejectionABC`	−0.11	−0.19	1.79	5.17	3.65	−7.69	−2.08	2.36	7.23
Layer 2	Method	Mean	Median	Mode	MLE	S.D.	2.5%	25%	75%	97.5%
`equalto.1.pm.0`	`copulaABCdrf`	−6.39	−6.33	−3.81	−5.26	1.44	−9.76	−7.26	−5.39	−3.91
	`rejectionABC`	−0.08	0.06	0.19	−0.06	3.17	−5.75	−2.44	1.72	6.60
`greaterthan.1`	`copulaABCdrf`	−2.76	−2.66	1.68	3.49	2.80	−7.83	−4.71	−1.02	3.13
	`rejectionABC`	−0.44	−0.32	−4.59	−2.16	3.38	−6.38	−2.88	1.97	6.20
`mutual.min`	`copulaABCdrf`	1.70	1.59	0.57	0.44	1.25	−0.32	0.78	2.57	3.83
	`rejectionABC`	0.16	0.50	0.64	−2.14	2.99	−5.79	−1.67	1.70	5.91
`tw.min.sum.min`	`copulaABCdrf`	0.23	0.27	−2.83	3.39	2.98	−5.81	−1.64	2.33	5.87
	`rejectionABC`	1.98	1.42	0.60	8.29	3.14	−3.33	−0.10	4.11	8.10
`tw.min.max.min`	`copulaABCdrf`	−2.48	−2.45	1.65	1.59	2.79	−8.02	−4.36	−0.64	3.30
	`rejectionABC`	1.37	1.41	2.10	5.36	2.79	−3.48	−0.60	3.13	6.96
`CMP`	`copulaABCdrf`	−2.04	−2.10	−2.82	4.43	2.93	−7.55	−3.95	−0.01	3.84
	`rejectionABC`	0.18	0.32	0.83	−2.18	3.27	−5.74	−2.00	2.54	5.99
Layer 3	Method	Mean	Median	Mode	MLE	S.D.	2.5%	25%	75%	97.5%
`equalto.1.pm.0`	`copulaABCdrf`	−5.67	−5.62	−4.63	−4.28	1.25	−8.44	−6.52	−4.78	−3.71
	`rejectionABC`	0.41	0.31	0.24	4.40	3.22	−5.31	−1.63	2.47	7.00
`greaterthan.1`	`copulaABCdrf`	−3.22	−3.17	0.82	−0.24	2.76	−8.22	−5.20	−1.45	2.37
	`rejectionABC`	0.21	−0.21	−1.93	3.16	2.73	−4.26	−1.80	1.79	6.05
`mutual.min`	`copulaABCdrf`	2.62	2.41	4.05	0.78	1.21	0.65	1.86	3.33	5.16
	`rejectionABC`	−0.07	0.03	0.77	−2.10	3.23	−5.58	−2.57	2.22	5.60
`tw.min.sum.min`	`copulaABCdrf`	−1.14	−1.23	−0.15	−6.25	2.82	−6.86	−2.79	0.68	4.23
	`rejectionABC`	0.22	0.35	0.18	2.70	3.26	−5.67	−1.81	1.98	5.74
`tw.min.max.min`	`copulaABCdrf`	−2.15	−2.28	−3.66	−0.16	2.85	−7.68	−4.03	−0.33	3.92
	`rejectionABC`	0.31	0.46	1.22	−0.54	2.67	−5.07	−1.24	1.84	6.21
`CMP`	`copulaABCdrf`	−1.34	−1.20	−1.99	2.29	2.72	−6.82	−3.20	0.47	4.11
	`rejectionABC`	−0.32	−0.47	−1.09	−1.36	2.98	−5.98	−2.00	1.72	5.59

Note: `tw` refers to `transitiveweights`.

To provide a further investigation, a simulation study of `copulaABCdrf` and `rejectionABC` was conducted using the posterior mean estimates shown in Table 16 as the true data-generating parameters of the same multilayer ERGM, and based on $n_{\text{sim}} = 100$ nodes and $n_{\text{sim}} = 10$ and 20, while noting that even these analyses were very computationally expensive. Tables 17–19 present the results of this simulation study. Table 17 shows the results of the MSEs and MAEs for the posterior mean, median, and mode estimates. Table 18 presents the MLE estimates compared to MCMLEs. Table 19 presents the mean 95% and 50% interval coverage. According to these three tables, both ABC methods were competitive, while `copulaABCdrf` tended to produce superior results.

Table 17. ABC posterior mean, median, and mode estimates of the 3-layer-valued ERGM parameters from network datasets simulated to reflect the `BostonBomb2013` network dataset, over 10 replications, for $n = 100$ nodes and for n_{sim} = 10 and 20.

	$n_{\text{sim}} =$	10	20	10	20	10	20
Layer 1	Method	Mean	Mean	Median	Median	Mode	Mode
`equalto.1.pm.0`	`copulaABCdrf`	**3.15** (**9.96**)	**2.28** (**5.19**)	**3.34** (**11.20**)	**2.43** (**5.93**)	4.04 (17.43)	**3.38** (**11.95**)
	`rejectionABC`	4.02 (16.28)	3.91 (15.37)	4.08 (16.97)	4.00 (16.3)	**3.82** (**16.49**)	3.76 (16.02)
`greaterthan.1`	`copulaABCdrf`	1.92 (3.71)	**1.00** (**1.10**)	1.98 (4.00)	**1.03** (**1.17**)	**1.81** (4.17)	**1.95** (**6.06**)
	`rejectionABC`	**1.72** (**3.59**)	1.52 (2.67)	**1.48** (**2.79**)	1.18 (1.87)	1.91 (5.31)	2.40 (6.89)
`mutual.min`	`copulaABCdrf`	**1.06** (**1.23**)	**1.14** (**1.36**)	**1.25** (**1.64**)	**1.31** (**1.80**)	**2.15** (**6.03**)	3.20 (13.16)
	`rejectionABC`	1.89 (3.98)	2.19 (5.37)	1.83 (4.04)	2.11 (4.98)	2.82 (8.81)	**2.91** (**11.09**)
`tw.min.sum.min`	`copulaABCdrf`	2.25 (**5.08**)	2.27 (5.19)	**2.17** (**4.71**)	2.15 (4.69)	**0.87** (**1.13**)	**1.22** (**2.26**)
	`rejectionABC`	**2.22** (5.29)	**1.97** (**4.04**)	2.32 (5.71)	**1.81** (**3.60**)	2.69 (10.21)	1.80 (4.43)
`tw.min.max.min`	`copulaABCdrf`	**0.09** (**0.01**)	**0.30** (**0.14**)	**0.07** (**0.01**)	**0.25** (**0.10**)	1.64 (3.62)	1.96 (6.51)
	`rejectionABC`	0.48 (0.31)	0.76 (0.90)	0.45 (0.36)	0.82 (1.00)	**1.38** (**3.29**)	**1.39** (**2.09**)
`CMP`	`copulaABCdrf`	**0.13** (**0.03**)	**0.26** (**0.08**)	**0.27** (**0.12**)	**0.15** (**0.02**)	1.96 (4.60)	2.11 (5.07)
	`rejectionABC`	0.41 (0.33)	0.34 (0.18)	0.66 (0.70)	0.43 (0.23)	**1.31** (**3.26**)	**1.03** (**1.50**)
Layer 2	Method	Mean	Mean	Median	Median	Mode	Mode
`equalto.1.pm.0`	`copulaABCdrf`	**3.96** (**15.73**)	**3.11** (**9.69**)	**4.13** (**17.11**)	**3.34** (**11.16**)	**4.15** (**18.40**)	**4.67** (**22.17**)
	`rejectionABC`	4.77 (23.18)	4.35 (19.03)	4.92 (24.50)	4.34 (19.00)	5.41 (33.17)	4.96 (26.54)
`greaterthan.1`	`copulaABCdrf`	**1.78** (**3.21**)	**0.94** (**0.92**)	**1.79** (**3.30**)	**1.02** (**1.09**)	3.71 (17.44)	2.34 (7.59)
	`rejectionABC`	2.03 (4.84)	1.31 (1.88)	2.06 (5.32)	1.11 (1.35)	**2.06** (**6.73**)	**1.78** (**4.80**)
`mutual.min`	`copulaABCdrf`	**3.04** (**9.83**)	**2.36** (**6.27**)	**2.91** (**9.06**)	**2.28** (**5.83**)	**1.92** (**4.17**)	2.91 (9.06)
	`rejectionABC`	3.72 (14.03)	3.65 (13.78)	3.80 (14.86)	3.39 (12.10)	3.65 (15.88)	**1.86** (**4.75**)
`tw.min.sum.min`	`copulaABCdrf`	2.36 (5.60)	2.56 (6.57)	2.26 (**5.15**)	2.57 (6.62)	**1.43** (**3.07**)	1.68 (4.81)
	`rejectionABC`	**2.33** (**5.80**)	**2.18** (**5.16**)	**2.22** (5.60)	**2.18** (**5.30**)	3.00 (15.47)	**1.49** (**3.51**)
`tw.min.max.min`	`copulaABCdrf`	**0.56** (**0.33**)	**0.41** (**0.20**)	**0.61** (**0.41**)	**0.47** (**0.27**)	1.45 (2.88)	1.83 (7.65)
	`rejectionABC`	0.61 (0.43)	0.81 (0.81)	0.67 (0.59)	1.10 (1.38)	**1.33** (**2.62**)	**1.20** (**1.89**)
`CMP`	`copulaABCdrf`	**0.12** (**0.02**)	**0.32** (**0.14**)	**0.17** (**0.04**)	**0.18** (**0.06**)	1.39 (2.79)	2.21 (6.37)
	`rejectionABC`	0.59 (0.61)	0.75 (0.92)	0.73 (0.85)	0.58 (0.78)	**1.19** (**2.24**)	**1.68** (**3.68**)
Layer 3	Method	Mean	Mean	Median	Median	Mode	Mode
`equalto.1.pm.0`	`copulaABCdrf`	**3.94** (**15.57**)	**2.61** (**6.91**)	**4.20** (**17.72**)	**2.88** (**8.37**)	4.92 (25.87)	2.91 (9.88)
	`rejectionABC`	4.56 (21.06)	3.73 (14.11)	4.59 (21.38)	3.75 (14.41)	**4.68** (**24.47**)	**2.84** (**9.46**)
`greaterthan.1`	`copulaABCdrf`	2.45 (6.03)	**1.45** (**2.19**)	2.52 (6.40)	**1.48** (**2.25**)	3.32 (12.91)	**2.54** (**9.59**)
	`rejectionABC`	**2.04** (**4.34**)	2.38 (6.05)	**2.05** (**4.67**)	2.26 (5.35)	**1.68** (**3.75**)	3.66 (16.5)
`mutual.min`	`copulaABCdrf`	**2.79** (**7.81**)	**2.12** (**4.64**)	**2.70** (**7.30**)	**2.07** (**4.37**)	**3.57** (**13.85**)	**2.40** (**6.56**)
	`rejectionABC`	4.81 (23.87)	4.74 (22.61)	4.77 (23.84)	4.57 (21.18)	4.05 (21.73)	4.55 (23.11)
`tw.min.sum.min`	`copulaABCdrf`	0.97 (**0.96**)	1.08 (**1.19**)	**0.88** (**0.80**)	1.01 (1.05)	**0.93** (**1.15**)	**0.86** (**1.07**)
	`rejectionABC`	**0.96** (1.02)	**0.98** (1.22)	0.99 (1.16)	**0.74** (**0.85**)	1.46 (2.61)	2.64 (9.04)
`tw.min.max.min`	`copulaABCdrf`	**0.21** (**0.05**)	**0.19** (**0.05**)	**0.28** (**0.10**)	0.30 (0.12)	2.21 (**6.86**)	1.98 (6.30)
	`rejectionABC`	0.61 (0.51)	0.26 (0.10)	0.67 (0.59)	**0.24** (**0.07**)	**2.16** (6.89)	**1.49** (**3.71**)
`CMP`	`copulaABCdrf`	0.89 (0.83)	1.06 (1.22)	0.65 (**0.48**)	0.80 (0.76)	**1.12** (**2.14**)	1.53 (3.00)
	`rejectionABC`	**0.50** (**0.33**)	**0.67** (**0.67**)	**0.64** (0.56)	**0.69** (**0.61**)	1.77 (5.75)	**0.67** (**0.68**)

Note: **Bold** indicates the more accurate ABC method for the given estimator. `tw` refers to `transitiveweights`.

Table 18. ABC MLEs and MCMLEs of the 3-layer-valued ERGM parameters from network datasets simulated to reflect the BostonBomb2013 network dataset, over 10 replications, for $n = 100$ nodes and for n_{sim} = 10 and 20.

	$n_{\text{sim}} =$	10	20	10	20
Layer 1	Method	MLE	MLE	MCMLE	MCMLE
`equalto.1.pm.0`	`copulaABCdrf`	**3.63** (**15.07**)	**3.27** (**11.11**)	0.19 (0.05)	0.21 (0.06)
	`rejectionABCkern`	4.20 (18.15)	4.34 (19.74)		
`greaterthan.1`	`copulaABCdrf`	3.15 (13.25)	2.90 (**8.75**)	1.31 (5.35)	0.42 (0.24)
	`rejectionABCkern`	**1.79** (**6.89**)	**2.61** (8.85)		
`mutual.min`	`copulaABCdrf`	**1.72** (**4.33**)	2.19 (6.15)	4.96 (43.20)	6.79 (138.42)
	`rejectionABCkern`	2.30 (6.95)	**1.68** (**3.74**)		
`tw.min.sum.min`	`copulaABCdrf`	**2.16** (**5.91**)	3.10 (16.61)	∞ (∞)	∞ (∞)
	`rejectionABCkern`	3.05 (15.01)	**1.30** (**1.96**)		
`tw.min.max.min`	`copulaABCdrf`	**2.23** (**9.67**)	2.82 (8.40)	∞ (∞)	∞ (∞)
	`rejectionABCkern`	3.00 (16.48)	**2.27** (**6.40**)		
`CMP`	`copulaABCdrf`	**1.48** (**2.83**)	**1.78** (**4.38**)	1.71 (9.56)	0.36 (0.18)
	`rejectionABCkern`	2.00 (5.89)	3.32 (20.56)		
Layer 2	Method	MLE	MLE	MCMLE	MCMLE
`equalto.1.pm.0`	`copulaABCdrf`	**4.36** (**20.49**)	**3.62** (**13.42**)	0.39 (0.20)	0.35 (0.15)
	`rejectionABCkern`	5.00 (31.16)	4.79 (26.93)		
`greaterthan.1`	`copulaABCdrf`	2.88 (8.89)	2.93 (11.84)	0.33 (0.16)	0.50 (0.39)
	`rejectionABCkern`	**2.48** (**8.05**)	**2.17** (**5.73**)		
`mutual.min`	`copulaABCdrf`	4.24 (26.38)	**1.24** (**1.61**)	0.38 (0.22)	0.48 (0.35)
	`rejectionABCkern`	**3.29** (**18.69**)	4.63 (39.55)		
`tw.min.sum.min`	`copulaABCdrf`	**2.25** (**7.67**)	5.12 (28.20)	∞ (∞)	∞ (∞)
	`rejectionABCkern`	3.36 (17.19)	**2.24** (**8.13**)		
`tw.min.max.min`	`copulaABCdrf`	2.35 (8.18)	**3.11** (14.38)	∞ (∞)	∞ (∞)
	`rejectionABCkern`	**1.62** (**3.58**)	3.31 (**14.09**)		
`CMP`	`copulaABCdrf`	**1.61** (**3.46**)	1.38 (2.83)	0.30 (0.12)	0.49 (0.40)
	`rejectionABCkern`	2.05 (5.21)	**1.04** (**1.76**)		
Layer 3	Method	MLE	MLE	MCMLE	MCMLE
`equalto.1.pm.0`	`copulaABCdrf`	4.39 (**21.34**)	3.56 (**13.41**)	0.24 (0.07)	0.23 (0.06)
	`rejectionABCkern`	**4.21** (21.71)	**3.29** (17.39)		
`greaterthan.1`	`copulaABCdrf`	4.67 (30.13)	3.36 (15.32)	0.24 (0.09)	0.23 (0.08)
	`rejectionABCkern`	**2.13** (**5.91**)	**3.14** (**10.99**)		
`mutual.min`	`copulaABCdrf`	**2.79** (**9.52**)	**2.85** (**8.74**)	0.58 (0.36)	0.68 (0.47)
	`rejectionABCkern`	5.45 (35.26)	6.51 (46.49)		
`tw.min.sum.min`	`copulaABCdrf`	2.14 (5.98)	3.23 (15.80)	∞ (∞)	∞ (∞)
	`rejectionABCkern`	**2.03** (**5.16**)	**2.97** (**11.92**)		
`tw.min.max.min`	`copulaABCdrf`	**1.85** (**4.07**)	**1.39** (**2.41**)	∞ (∞)	∞ (∞)
	`rejectionABCkern`	2.43 (11.27)	1.58 (3.61)		
`CMP`	`copulaABCdrf`	**1.32** (**2.44**)	**0.93** (**1.29**)	0.20 (0.08)	0.31 (0.12)
	`rejectionABCkern`	1.62 (4.22)	3.34 (16.36)		

Note: **Bold** indicates the more accurate ABC method for the given estimator. `tw` refers to `transitiveweights`.

Table 19. ABC and MCMLE mean 95% and 50% coverage of the 3-layer-valued ERGM parameters from network datasets simulated to reflect the BostonBomb2013 network dataset, over 10 replications, for $n = 100$ nodes and for n_{sim} = 10 and 20.

	n_{sim} =	10	20	10	20
Layer 1	Method	95%(50%)c	95%(50%)c	MCMLE95%c	MCMLE95%c
`equalto.1.pm.0`	`copulaABCdrf`	**0.90** (**0.00**)	**1.00** (**0.00**)	0.90	0.80
	`rejectionABC`	0.30 (**0.00**)	0.20 (**0.00**)		
`greaterthan.1`	`copulaABCdrf`	**1.00** (0.30)	**1.00** (1.00)	1.00	1.00
	`rejectionABC`	**1.00** (**0.40**)	**1.00** (**0.60**)		
`mutual.min`	`copulaABCdrf`	**1.00** (**0.70**)	**1.00** (**0.60**)	1.00	1.00
	`rejectionABC`	**1.00** (0.20)	**1.00** (0.20)		
`tw.min.sum.min`	`copulaABCdrf`	**1.00** (0.00)	**1.00** (0.00)	0.10	0.00
	`rejectionABC`	**1.00** (**0.10**)	**1.00** (**0.20**)		
`tw.min.max.min`	`copulaABCdrf`	**1.00** (**1.00**)	**1.00** (**1.00**)	0.10	0.00
	`rejectionABC`	**1.00** (**1.00**)	**1.00** (**1.00**)		
`CMP`	`copulaABCdrf`	**1.00** (**1.00**)	**1.00** (**1.00**)	1.00	1.00
	`rejectionABC`	**1.00** (**1.00**)	**1.00** (**1.00**)		
Layer 2	Method	95%(50%)c	95%(50%)c	MCMLE95%c	MCMLE95%c
`equalto.1.pm.0`	`copulaABCdrf`	**0.50** (**0.00**)	**0.80** (**0.00**)	0.70	0.80
	`rejectionABC`	0.30 (**0.00**)	0.00 (**0.00**)		
`greaterthan.1`	`copulaABCdrf`	**1.00** (**0.60**)	**1.00** (1.00)	1.00	1.00
	`rejectionABC`	**1.00** (**0.40**)	**1.00** (**0.80**)		
`mutual.min`	`copulaABCdrf`	0.80 (**0.00**)	**1.00** (**0.00**)	0.60	0.40
	`rejectionABC`	**1.00** (**0.00**)	0.60 (**0.00**)		
`tw.min.sum.min`	`copulaABCdrf`	**1.00** (0.00)	**1.00** (**0.00**)	0.00	0.00
	`rejectionABC`	**1.00** (**0.10**)	**1.00** (**0.00**)		
`tw.min.max.min`	`copulaABCdrf`	**1.00** (1.00)	**1.00** (1.00)	0.00	0.00
	`rejectionABC`	**1.00** (**1.00**)	**1.00** (**0.60**)		
`CMP`	`copulaABCdrf`	**1.00** (**1.00**)	**1.00** (1.00)	1.00	1.00
	`rejectionABC`	**1.00** (**0.90**)	**1.00** (**0.80**)		
Layer 3	Method	95%(50%)c	95%(50%)c	MCMLE95%c	MCMLE95%c
`equalto.1.pm.0`	`copulaABCdrf`	**0.90** (**0.00**)	**1.00** (**0.00**)	0.60	0.80
	`rejectionABC`	0.00 (**0.00**)	0.60 (**0.00**)		
`greaterthan.1`	`copulaABCdrf`	**1.00** (0.00)	**1.00** (**0.60**)	0.90	1.00
	`rejectionABC`	**1.00** (**0.10**)	**1.00** (0.00)		
`mutual.min`	`copulaABCdrf`	**0.50** (**0.00**)	**0.60** (**0.00**)	0.00	0.00
	`rejectionABC`	0.30 (**0.00**)	0.20 (**0.00**)		
`tw.min.sum.min`	`copulaABCdrf`	**1.00** (1.00)	**1.00** (1.00)	0.00	0.00
	`rejectionABC`	**1.00** (**0.90**)	**1.00** (**0.80**)		
`tw.min.max.min`	`copulaABCdrf`	**1.00** (**1.00**)	**1.00** (**1.00**)	0.00	0.00
	`rejectionABC`	**1.00** (**1.00**)	**1.00** (**1.00**)		
`CMP`	`copulaABCdrf`	**1.00** (**1.00**)	**1.00** (1.00)	0.90	0.80
	`rejectionABC`	**1.00** (**1.00**)	**1.00** (**0.80**)		

Note: **Bold** indicates the more accurate ABC method for the given estimator.

3.8. Real Social Network Analysis

We now apply Algorithm 1 to analyze the massive `friendster` social network of n = 65,608,366 persons (nodes) and 1,806,067,135 undirected edges [165] (this dataset was obtained from https://snap.stanford.edu/data/com-Friendster.html and accessed on 7 February 2024) using the NLPA model, assigned uniform priors $(\alpha, p) \sim \mathcal{U}_{(0,3)}\mathcal{U}_{(0,0.20)}$, and using summary statistics $s(x)$ = (average clustering coefficient, diameter, density) = $(0.1623,\ 32,\ 8.39161 \times 10^{-7})$. Also, these network summaries were also computed from a network of size n_{sim} = 1000 simulated from the model (given a set of proposed parameters) in each of the N = 10,000 iterations of Algorithm 1. The algorithm routinely delivered posterior distribution and MLE estimates, shown in Table 20, despite the sheer size of the dataset. Also, this table reports the results of `rejectionABC` using all the three summary statistics, as well as the results of `rejectionABC`. For the `rejectionABC` analysis, the first two summary statistics were pre-selected as the most important among the three total summary statistics, namely, average clustering coefficient (with an importance measure of 0.66), diameter (importance 0.22), and density (importance of 0.12). The importance of each summary was measured from `drf` regression performed on the reference table as before, regressing the prior parameter samples on all three summaries. The pre-selection of summaries led `rejectionABC` to produce posterior mean and median estimates that were similar to those of `copulaABCdrf`.

Table 20. Posterior estimates of the NLPA model from the `friendster` network dataset.

	α	p
`copulaABCdrf` mean	1.21	0.001
`rejectionABC` mean	0.69	0.002
`rejectionABCselect` mean	1.21	0.02
`copulaABCdrf` median	1.25	0.001
`rejectionABC` median	0.64	0.002
`rejectionABCselect` median	1.20	0.02
`copulaABCdrf` mode	0.24	0.0002
`rejectionABC` mode	0.42	0.001
`rejectionABCkern.select` mode	1.25	0.01
`copulaABCdrf` MLE	0.24	0.0002
`rejectionABC` MLE	0.42	0.001
`rejectionABCkern.select` MLE	1.25	0.01
`copulaABCdrf` s.d.	0.28	0.002
`rejectionABC` s.d.	0.44	0.002
`rejectionABCselect` s.d.	0.12	0.01
`copulaABCdrf` 50%	(1.18, 1.31)	(0.0003, 0.001)
`rejectionABC` 50%	(0.33, 1.01)	(0.001, 0.004)
`rejectionABCselect` 50%	(1.14, 1.26)	(0.01, 0.02)
`copulaABCdrf` 95%	(0.37, 2.05)	(0.0001, 0.005)
`rejectionABC` 95%	(0.04, 1.44)	(0.0002, 0.005)
`rejectionABCselect` 95%	(1.04, 1.37)	(0.002, 0.03)
`copulaABCdrf` copula	d.f. 28.91	scale 0.32

4. Conclusions and Discussion

This article introduced `copulaABCdrf` as a framework that unifies previous methods of `copulaABC` and `abcrf`, which were introduced separately in previous articles. This unified method aimed to provide, in a single unified framework, a wide range of inferences from the approximate posterior distribution (including posterior means, medians, mode, and the MLE) for models that may be defined by intractable likelihoods and many parameters, as well as model selection. All of this was carried out while being able to automatically select the subset of the most relevant summary statistics, from a potentially high number of summary statistics, without requiring the user to identify all the important summary statistics by hand before data analysis. Further, `copulaABCdrf` avoids the need to use tolerance and distance measure tuning parameters of `rejectionABC` methods.

This paper also proposed a new solution to the simulation problem in ABC, which is based on simulating datasets from the exact model likelihood of size smaller than the potentially very large size of the datasets being analyzed by the given model. This strategy could potentially be useful for models from which it is computationally costly to simulate datasets, including models for large networks such as the ERGM and mechanistic network models, using UNC summary statistics calculated on simulated network datasets that were smaller than the potentially large network dataset being analyzed.

Based on the results of many simulation studies performed in this paper, evaluating and comparing `copulaABCdrf`, `rejectionABC`, and semiautomatic ABC `rejectionABCselect` methods based on `drf`, the following general conclusions can be made:

- For the low-dimensional parametric models ($d \leq 5$) considered in this paper, `copulaABCdrf` and `rejectionABC` methods (including `rejectionABCselect`, which uses pre-selection of summary statistics using `drf`) were competitive in terms of MAE and MSE of the estimation of marginal posterior mean and median, and of the posterior mode and MLE of the model parameters. Also, `rejectionABCselect` often outperformed `rejectionABC` without pre-selection of the subset of summaries and often produced posterior estimates that were similar to those of `copulaABCdrf`.
- The `copulaABCdrf` framework tended to outperform `rejectionABC` methods in the estimation accuracy of univariate marginal posterior distributions according to the KS distance to the exact univariate marginal posterior distributions. However, all KS tests of all ABC methods indicated statistically significant departures.

- For high-dimensional parametric models ($d \geq 5$) with a posterior that was not highly multimodal, `copulaABCdrf` far outperformed `rejectionABC` methods.
- For all ABC methods, posterior estimation accuracy tended to be best when $n_{sim}/n = 1$, and accuracy decreased as n_{sim}/n decreased below 1, which is not surprising because, then, the summaries become less informative about the data.
- For all ABC methods, estimation from the posterior distribution can be challenging when the true posterior is highly multimodal.

Further, it was found in the simulations that the network size offset MPLEs were useful, but too computationally costly, compared to the other UNC network summary statistics, which were far more rapidly computable because they did not require optimization algorithms. For future research, computational speed can be drastically improved by using parallel computing methods to compute each of the summary statistics while constructing the reference table.

The results suggest that `drf` is the main driving force of `copulaABCdrf`, as suggested, for example, by the fact that `copulaABCdrf` often produced results similar to those of `rejectionABCselect`. However, according to the KS tests, `drf` always produced univariate marginal posterior estimates that significantly departed from the corresponding exact posterior densities. On a related note, after the initial writing of this manuscript in early January 2024 (see Acknowledgements), `drf` was proposed as a tool for recursively estimating weights for sequential Monte Carlo ABC algorithms [166] for models defined by intractable likelihoods with fewer than five parameters. The `copulaABCdrf` framework, seemingly because it is based on the meta-t posterior distribution approximation, was limited (like the `rejectionABC` methods) to the accurate estimation of more symmetric or skewed unimodal posterior distributions. The framework is partially motivated by the fact that the posterior given n i.i.d. observations converges to a multivariate normal distribution asymptotically as $n \to \infty$ of i.i.d. observations, according to the Bernstein von Mises theorem, under mild regularity conditions [34].

In other words, `copulaABCdrf` based on the meta-t is limited in estimating highly multimodal posteriors, at least according to one simulation study. For future research, this issue can potentially be addressed by using more flexible copula density functions that can handle nonlinear relationships among parameters, perhaps while taking advantage of recent developments in the continually active field of copula-based multivariate density estimation. However, such more flexible nonlinear copulas can be very computationally costly to estimate, especially for high-dimensional parameters. Still, finding a method to efficiently compute estimates of such a more flexible copula density in high-dimensional settings seems worth pursuing in future research.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable. This is because the results of this study is not based on Human Subjects Research. In particular, this study only analyzed simulated datasets, and analyzed publicly available secondary datasets, as mentioned in Section 3 of this article.

Informed Consent Statement: Not applicable, because this study did not involve human subjects research.

Data Availability Statement: The real datasets can be obtained through sources cited within the paper. The R software code files used to simulate the data and to analyze the real and simulated datasets can be obtained from the author via https://github.com/GeorgeKarabatsos/copulaABCdrf which is currently accessible.

Acknowledgments: The author thanks anonymous reviewers for providing helpful comments on a previous version of this manuscript, which helped to improve its presentation. This manuscript was presented at a Biostatistics Seminar at the Medical College of Wisconsin during 19 March 2024, and at the conferences COMPSTAT 2024 (University of Giessen, 27–30 August 2024) and CFE-CMStatistics 2024 (King's College London, 14–16 December 2024). The initial version of this manuscript, including the results of simulation studies and analyses of real networks, obtained using

Algorithm 1 implementing distribution random forests and copula modeling, was presented as a report in a grant proposal submitted to the U.S. National Science Foundation on 14 February 2024.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Tavaré, S.; Balding, D.; Griffiths, R.; Donnelly, P. Inferring coalescence times from DNA sequence data. *Genetics* **1997**, *145*, 505–518. [CrossRef] [PubMed]
2. Pritchard, J.; Seielstad, M.; Perez-Lezaun, A.; Feldman, M. Population growth of human Y chromosomes: A study of Y chromosome microsatellites. *Mol. Biol. Evol.* **1999**, *16*, 1791–1798. [CrossRef] [PubMed]
3. Marin, J.M.; Pudlo, P.; Robert, C.; Ryder, R. Approximate Bayesian Computational methods. *Stat. Comput.* **2012**, *22*, 1167–1180. [CrossRef]
4. Bernardo, J.; Smith, A. *Bayesian Theory*; Wiley: Chichester, UK, 1994.
5. Biau, G.; Cérou, F.; Guyader, A. New insights into Approximate Bayesian Computation. *Ann. L'Institut Henri Poincaré Probab. Stat.* **2015**, *51*, 376–403. [CrossRef]
6. Li, W.; Fearnhead, P. On the asymptotic efficiency of approximate Bayesian computation estimators. *Biometrika* **2018**, *105*, 285–299. [CrossRef]
7. Fearnhead, P.; Prangle, D. Constructing summary statistics for Approximate Bayesian Computation: Semi-automatic Approximate Bayesian Computation. *J. R. Stat. Soc. Ser. B* **2012**, *74*, 419–474. [CrossRef]
8. Blum, M.; Nunes, M.; Prangle, D.; Sisson, S. A comparative review of dimension reduction methods in Approximate Bayesian Computation. *Stat. Sci.* **2013**, *28*, 189–208. [CrossRef]
9. Sunnåker, M.; Busetto, A.; Numminen, E.; Corander, J.; Foll, M.; Dessimoz, C. Approximate Bayesian Computation. *PLoS Comput. Biol.* **2013**, *9*, 1–10. [CrossRef]
10. Karabatsos, G.; Leisen, F. An approximate likelihood perspective on ABC methods. *Stat. Surv.* **2018**, *12*, 66–104. [CrossRef]
11. Sisson, S.; Fan, Y.; Beaumont, M. *Handbook of Approximate Bayesian Computation*; Chapman and Hall/CRC Press: Boca Raton, FL, USA, 2018.
12. Grazian, C.; Fan, Y. A review of approximate Bayesian computation methods via density estimation: Inference for simulator-models. *WIREs Comput. Stat.* **2020**, *12*, e1486. [CrossRef]
13. Cranmer, K.; Brehmer, J.; Louppe, G. The frontier of simulation-based inference. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 30055–30062. [CrossRef] [PubMed]
14. Craiu, R.; Levi, E. Approximate methods for Bayesian computation. *Annu. Rev. Stat. Its Appl.* **2022**, *10*, 379–399. [CrossRef]
15. Karabatsos, G. Approximate Bayesian computation using asymptotically normal point estimates. *Comput. Stat.* **2023**, *38*, 531–568. [CrossRef]
16. Pesonen, H.; Simola, U.; Köhn-Luque, A.; Vuollekoski, H.; Lai, X.; Frigessi, A.; Kaski, S.; Frazier, D.; Maneesoonthorn, W.; Martin, G.; et al. ABC of the future. *Int. Stat. Rev.* **2023**, *91*, 243–268. [CrossRef]
17. Martin, G.; Frazier, D.; Robert, C. Approximating Bayes in the 21st Century. *Stat. Sci.* **2023**, *39*, 20–45. [CrossRef]
18. Li, J.; Nott, D.; Fan, Y.; Sisson, S. Extending Approximate Bayesian Computation methods to high dimensions via a Gaussian copula model. *Comput. Stat. Data Anal.* **2017**, *106*, 77–89. [CrossRef]
19. Nott, D.; Ong, V.; Fan, Y.; Sisson, S. High-dimensional ABC. In *Handbook of Approximate Bayesian Computation*; Sisson, S., Fan, Y., Beaumont, M., Eds.; Chapman and Hall/CRC: Boca Raton, FL, USA, 2018; pp. 106–117.
20. Chen, Y.; Gutmann, M. Adaptive Gaussian Copula ABC. In Proceedings of the Twenty-Second International Conference on Artificial Intelligence and Statistics, Okinawa, Japan, 16–18 April 2019; Chaudhuri, K., Sugiyama, M., Eds.; PMLR: Cambridge, MA, USA, 2019 ; Volume 89, pp. 1584–1592.
21. Klein, N.; Stanley Smith, M.; Nott, D.; Chrisholm, R. Regression copulas for multivariate responses. *arXiv* **2024**, arXiv:2401.11804.
22. Raynal, L.; Marin, J.; Pudlo, P.; Ribatet, M.; Robert, C.; Estoup, A. ABC random forests for Bayesian parameter inference. *Bioinformatics* **2018**, *35*, 1720–1728. [CrossRef]
23. Rubio, F.; Johansen, A. A simple approach to maximum intractable likelihood estimation. *Electron. J. Stat.* **2013**, *7*, 1632–1654. [CrossRef]
24. Kajihara, T.; Kanagawa, M.; Yamazaki, K.; Fukumizu, K. Kernel recursive ABC: Point estimation with intractable likelihood. In Proceedings of the 35th International Conference on Machine Learning, Stockholm, Sweden, 10–15 July 2018; Dy, J., Krause, A., Eds.; PMLR: Cambridge, MA, USA, 2018 ; Volume 80, pp. 2400–2409.
25. Picchini, U.; Anderson, R. Approximate maximum likelihood estimation using data-cloning ABC. *Comput. Stat. Data Anal.* **2017**, *105*, 166–183. [CrossRef]
26. Gutmann, M.; Corander, J. Bayesian optimization for likelihood-free inference of simulator-based statistical models. *J. Mach. Learn. Res.* **2016**, *17*, 1–47.
27. Yildirim, S.; Singh, S.; Dean, T.; Jasra, A. Parameter estimation in hidden Markov Models with intractable likelihoods using sequential Monte Carlo. *J. Comput. Graph. Stat.* **2015**, *24*, 846–865. [CrossRef]
28. Dean, T.; Singh, S.; Jasra, A.; Peters, G. Parameter estimation for hidden Markov models with intractable likelihoods. *Scand. J. Stat.* **2014**, *41*, 970–987. [CrossRef]

29. Gourieroux, C.; Monfort, A.; Renault, E. Indirect inference. *J. Appl. Econ.* **1993**, *8*, S85–S118. [CrossRef]
30. McFadden, D. A method of simulated moments for estimation of discrete response models without numerical integration. *Econometrica* **1989**, *57*, 995–1026. [CrossRef]
31. Ćevid, D.; Michel, L.; Näf, J.; Bühlmann, P.; Meinshausen, N. Distributional random forests: Heterogeneity adjustment and multivariate distributional regression. *J. Mach. Learn. Res.* **2022**, *23*, 14987–15065.
32. Papamakarios, G.; Sterratt, D.; Murray, I. Sequential neural likelihood: Fast likelihood-free inference with autoregressive flows. In Proceedings of the Twenty-Second International Conference on Artificial Intelligence and Statistics, Cambridge, MA, USA, 16–18 April 2019; Chaudhuri, K., Sugiyama, M., Eds.; PMLR: Cambridge, MA, USA, 2019 ; Volume 89, pp. 837–848.
33. Wang, Y.; Rocková, V. Adversarial Bayesian simulation. *arXiv* **2023**, arXiv:2208.12113.
34. Kleijn, B.; van der Vaart, A. The Bernstein-Von-Mises theorem under misspecification. *Electron. J. Stat.* **2012**, *6*, 354–381. [CrossRef]
35. Sklar, M. Fonctions de repartition an dimensions et leurs marges. *Ann. l'ISUP* **1959**, *8*, 229–231.
36. Denuit, M.; Lambert, P. Constraints on concordance measures in bivariate discrete data. *J. Multivar. Anal.* **2005**, *93*, 40–57. [CrossRef]
37. Madsen, L.; Fang, Y. Joint regression analysis for discrete longitudinal data. *Biometrics* **2011**, *67*, 1171–1176. [CrossRef] [PubMed]
38. Nelsen, R. *An Introduction to Copulas*; Springer: New York, NY, USA, 2006.
39. Hutson, A.; Wilding, G.; Mashtare, T.; Vexler, A. Measures of biomarker dependence using a copula-based multivariate epsilon-skew-normal family of distributions. *J. Appl. Stat.* **2015**, *42*, 2734–2753. [CrossRef] [PubMed]
40. Smith, M.; Vahey, S. Asymmetric forecast densities for U.S. macroeconomic variables from a Gaussian copula model of cross-sectional and serial dependence. *J. Bus. Econ. Stat.* **2016**, *34*, 416–434. [CrossRef]
41. Baillien, J.; Gijbels, I.; Verhasselt, A. Estimation in copula models with two-piece skewed margins using the inference for margins method. *Econom. Stat.* **2022**, *in press*. [CrossRef]
42. Wei, Z.; Kim, S.; Kim, D. Multivariate Skew Normal Copula for Non-exchangeable Dependence. *Procedia Comput. Sci.* **2016**, *91*, 141–150. [CrossRef]
43. Yoshiba, T. Maximum likelihood estimation of skew-t copulas with its applications to stock returns. *J. Stat. Comput. Simul.* **2018**, *88*, 2489–2506. [CrossRef]
44. Demarta, S.; McNeil, A. The *t* copula and related copulas. *Int. Stat. Rev.* **2005**, *73*, 111–129. [CrossRef]
45. Daul, S.; De Giorgi, E.; Lindskog, F.; McNeil, A. The grouped *t*-copula with an application to credit risk. *SSRN* **2003**, *1358956*, 1–7
46. Kosmidis, I.; Karlis, D. Model-based clustering using copulas with applications. *Stat. Comput.* **2016**, *26*, 1079–1099. [CrossRef]
47. Smith, M.; Klein, N. Bayesian inference for regression copulas. *J. Bus. Econ. Stat.* **2021**, *39*, 712–728. [CrossRef]
48. Acar, E.; Craiu, R.; Yao, F. Statistical testing of covariate effects in conditional copula models. *Electron. J. Stat.* **2013**, *7*, 2822–2850. [CrossRef]
49. Hintz, E.; Hofert, M.; Lemieux, C. Computational challenges of t and related copulas. *J. Data Sci.* **2022**, *20*, 95–110. [CrossRef]
50. Dellaportas, P.; Tsionas, M. Importance sampling from posterior distributions using copula-like approximations. *J. Econ.* **2019**, *210*, 45–57. [CrossRef]
51. Qu, L.; Lu, Y. Copula density estimation by finite mixture of parametric copula densities. *Commun. Stat. Simul. Comput.* **2021**, *50*, 3315–3337. [CrossRef]
52. Fang, H.; Fang, K.; Kotz, S. The meta-elliptical distributions with given marginals. *J. Multivar. Anal.* **2002**, *82*, 1–16. [CrossRef]
53. Pitt, M.; Chan, D.; Kohn, R. Efficient Bayesian inference for Gaussian copula regression models. *Biometrika* **2006**, *93*, 537–554. [CrossRef]
54. Song, P. Multivariate dispersion models generated from Gaussian copula. *Scand. J. Stat.* **2000**, *27*, 305–320. [CrossRef]
55. Lange, K.; Little, R.; Taylor, J. Robust statistical modeling using the *t* distribution. *J. Am. Stat. Assoc.* **1989**, *84*, 881–896. [CrossRef]
56. Drovandi, C.; Nott, D.; Frazier, D. Improving the accuracy of marginal approximations in likelihood-free inference via localization. *J. Comput. Graph. Stat.* **2023**, *33*, 101–111. [CrossRef]
57. An, Z.; Nott, D.; Drovandi, C. Robust Bayesian synthetic likelihood via a semi-parametric approach. *Stat. Comput.* **2020**, *30*, 543–557. [CrossRef]
58. Pudlo, P.; Marin, J.M.; Estoup, A.; Cornuet, J.M.; Gautier, M.; Robert, C. Reliable ABC model choice via random forests. *Bioinformatics* **2016**, *32*, 859–866. [CrossRef] [PubMed]
59. Marin, J.; Pudlo, P.; Estoup, A.; Robert, C. Likelihood-free model choice. In *Handbook of Approximate Bayesian Computation*; Sisson, S., Fan, Y., Beaumont, M., Eds.; Chapman and Hall/CRC: Boca Raton, FL, USA, 2018.
60. Devroye, L. Recursive estimation of the mode of a multivariate density. *Can. J. Stat.* **1979**, *7*, 159–167. [CrossRef]
61. Abraham, C.; Biau, G.; Cadre, B. Simple estimation of the mode of a multivariate density. *Can. J. Stat.* **2003**, *31*, 23–34. [CrossRef]
62. Abraham, C.; Biau, G.; Cadre, B. On the asymptotic properties of a simple estimate of the mode. *ESAIM Probab. Stat.* **2004**, *8*, 1–11. [CrossRef]
63. Hsu, C.; Wu, T. Efficient estimation of the mode of continuous multivariate data. *Comput. Stat. Data Anal.* **2013**, *63*, 148–159. [CrossRef]
64. Dasgupta, S.; Kpotufe, S. Optimal rates for k-NN density and mode estimation. In *Advances in Neural Information Processing Systems*; Ghahramani, Z., Welling, M., Cortes, C., Lawrence, N., Weinberger, K., Eds.; Curran Associates, Inc.: Red Hook, NY, USA, 2014; Volume 27, pp. 1–9.
65. Chacón, J. The modal age of statistics. *Int. Stat. Rev.* **2020**, *88*, 122–141. [CrossRef]

66. Picchini, U.; Simola, U.; Corander, J. Sequentially guided MCMC proposals for synthetic likelihoods and correlated synthetic likelihoods. *Bayesian Anal.* **2022**, *18*, 1099–1129. [CrossRef]
67. Picchini, U.; Tamborrino, M. Guided sequential ABC schemes for intractable Bayesian models. *arXiv* **2024**, arXiv:2206.12235. [CrossRef]
68. Newton, M.; Polson, N.; Xu, J. Weighted Bayesian bootstrap for scalable posterior distributions. *Can. J. Stat.* **2021**, *49*, 421–437. [CrossRef]
69. Barrientos, A.; Peña, V. Bayesian bootstraps for massive data. *Bayesian Anal.* **2020**, *15*, 363–388. [CrossRef]
70. Lyddon, S.; Holmes, C.; Walker, S. General Bayesian updating and the loss-likelihood bootstrap. *Biometrika* **2019**, *106*, 465–478. [CrossRef]
71. Zhu, W.; Marin, J.; Leisen, F. A Bootstrap likelihood approach to Bayesian computation. *Aust. N. Z. J. Stat.* **2016**, *58*, 227–244. [CrossRef]
72. Robert, C.; Cornuet, J.; Marin, J.; Pillai, N. Lack of confidence in Approximate Bayesian Computation model choice. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 15112–15117. [CrossRef]
73. Okhrin, O.; Ristig, A.; Xu, Y. Copulae in High Dimensions: An Introduction. In *Applied Quantitative Finance*; Härdle, W., Chen, C., Overbeck, L., Eds.; Springer: Berlin/Heidelberg, Germany, 2017; pp. 247–277.
74. Geenens, G. Copula modeling for discrete random vectors. *Depend. Model.* **2020**, *8*, 417–440. [CrossRef]
75. Genest, C.; Ghoudi, K.; Rivest, L.P. A semiparametric estimation procedure of dependence parameters in multivariate families of distributions. *Biometrika* **1995**, *82*, 543–552. [CrossRef]
76. Hintz, E.; Hofert, M.; Lemieux, C. Multivariate normal variance mixtures in R: The R package nvmix. *J. Stat. Softw.* **2022**, *102*, 1–31. [CrossRef]
77. Lin, Y.; Jeon, Y. Random forests and adaptive nearest neighbors. *J. Am. Stat. Assoc.* **2006**, *101*, 578–590. [CrossRef]
78. Breiman, L.; Friedman, J.; Stone, C.; Olshen, R. *Classification and Regression Trees*; CRC Press: Boca Raton, FL, USA, 1984.
79. Gretton, A.; Borgwardt, K.; Rasch, M.; Schölkopf, B.; Smola, A. A kernel method for the two-sample-problem. In *Advances in Neural Information Processing Systems*; Platt, J., Koller, D., Singer, Y., Roweis, S., Eds.; The MIT Press: Cambridge, MA, USA, 2007; Volume 19, pp. 513–520.
80. Sheather, S.; Jones, M. A reliable data-based bandwidth selection method for kernel density estimation. *J. R. Stat. Soc. Ser. B* **1991**, *53*, 683–690. [CrossRef]
81. Geenens, G. Probit transformation for kernel density estimation on the unit interval. *J. Am. Stat. Assoc.* **2014**, *109*, 346–358. [CrossRef]
82. Geenens, G.; Wang, C. Local-likelihood transformation kernel density estimation for positive random variables. *J. Comput. Graph. Stat.* **2018**, *27*, 822–835. [CrossRef]
83. Nagler, T.; Vatter, T. kde1d: Univariate Kernel Density Estimation. R Package Version 1.0.7. 2024. Available online: https://cran.r-project.org/web/packages/kde1d/kde1d.pdf (accessed on 21 June 2024).
84. Michel, L.; Ćevid, D. drf: Distributional Random Forests. R Package Version 1.1.0. 2021. Available online: https://cran.r-project.org/web/packages/drf/drf.pdf (accessed on 3 January 2024).
85. Bickel, P.; Klaassen, C.; Ritov, Y.; Wellner, J. *Efficient and Adaptive Estimation for Semiparametric Models*; Johns Hopkins University Press: Baltimore, MD, USA, 1993.
86. Duong, T. ks: Kernel Smoothing. R Package Version 1.14.2. 2024. Available online: https://cran.r-project.org/web/packages/ks/ks.pdf (accessed on 21 June 2024).
87. Wasserman, L. *All of Nonparametric Statistics*; Springer: New York, NY, USA, 2006.
88. Monahan, J. *Numerical Methods of Statistics*; Cambridge University Press: New York, NY, USA, 2011.
89. MacEachern, S.; Berliner, L. Subsampling the Gibbs sampler. *Am. Stat.* **1994**, *48*, 188–190. [CrossRef]
90. Neal, R. Density Modeling and Clustering Using Dirichlet Diffusion Trees. *Bayesian Stat.* **2003**, *7*, 619–629.
91. Robert, C.; Casella, G. *Monte Carlo Statistical Methods*, 2nd ed.; Springer: New York, NY, USA, 2004.
92. Hoffman, M.; Gelman, A. The No-U-Turn sampler: Adaptively setting path lengths in Hamiltonian Monte Carlo. *J. Mach. Learn. Res.* **2014**, *15*, 1593–1623.
93. Goodman, J.; Weare, J. Ensemble samplers with affine invariance. *Commun. Appl. Math. Comput. Sci.* **2010**, *5*, 65–80. [CrossRef]
94. Krivitsky, P.; Handcock, M.; Morris, M. Adjusting for network size and composition effects in exponential-family random graph models. *Stat. Methodol.* **2011**, *8*, 319–339. [CrossRef]
95. Handcock, M.; Hunter, D.; Butts, C.; Goodreau, A.; Krivitsky, P.; Morris, M. ergm: Fit, Simulate and Diagnose Exponential-Family Models for Networks. The Statnet Project. R Package Version 4.5.0. 2023. Available online: https://cloud.r-project.org/web/packages/ergm/ergm.pdf (accessed on 3 January 2024).
96. Krivitsky, P. ergm.count: Fit, Simulate and Diagnose Exponential-Family Models for Networks with Count Edges. R Package Version 4.1.1. 2022. Available online: https://rdrr.io/github/statnet/ergm.count/man/ergm.count-package.html (accessed on 3 January 2024).
97. Csárdi, G.; Nepusz, T.; Traag, V.; Horvát, S.; Zanini, F.; Noom, D.; Müller, K. igraph: Network Analysis and Visualization in R. R Package Version 1.5.1. 2024. Available online: https://CRAN.R-project.org/package=igraph (accessed on 3 January 2024).
98. Snijders, T. Markov chain Monte Carlo estimation of exponential random graph models. *J. Soc. Struct.* **2002**, *3*, 1–40.
99. Strauss, D. On a general class of models for interaction. *SIAM Rev.* **1986**, *28*, 513–527. [CrossRef]

100. Handcock, M. *Assessing Degeneracy in Statistical Models of Social Networks*; Technical Report; University of Washington, Center for Statistics and the Social Sciences: Washington, DC, USA, 2003.
101. Genest, C.; Nešlehová, J. A primer on copulas for count data. *ASTIN Bull.* **2007**, *37*, 475–515. [CrossRef]
102. Haario, H.; Saksman, E.; Tamminen, J. Adaptive proposal distribution for random walk Metropolis algorithm. *Comput. Stat.* **1999**, *14*, 375–395. [CrossRef]
103. Snijders, T.; Pattison, P.; Robins, G.; Handcock, M. New specifications for exponential random graph models. *Sociol. Methodol.* **2006**, *36*, 99–153. [CrossRef]
104. Newman, M. The structure and function of complex networks. *SIAM Rev.* **2003**, *45*, 167–256. [CrossRef]
105. Wasserman, S.; Faust, K. *Social Network Analysis: Methods and Applications*; Cambridge University Press: Cambridge, UK, 1994.
106. Watts, D.; Strogatz, S. Collective dynamics of small-world networks. *Nature* **1998**, *393*, 440–442. [CrossRef]
107. Goldenberg, A.; Zheng, A.; Fienberg, S.; Airoldi, E. A survey of statistical network models. *Found. Trends Mach. Learn.* **2010**, *2*, 129–233. [CrossRef]
108. Snijders, T. Statistical models for social networks. *Annu. Rev. Sociol.* **2011**, *37*, 131–153. [CrossRef]
109. Pastor-Satorras, R.; Vespignani, A. *Evolution and Structure of the Internet: A Statistical Physics Approach*; Cambridge University Press: Cambridge, UK, 2004.
110. Raval, R.; Ray, A. *Introduction to Biological Networks*; Taylor & Francis: Andover, MA, USA; London, UK, 2013.
111. Newman, M. *Networks*, 2nd ed.; Oxford University Press: Oxford, UK, 2018.
112. Horvat, E.; Zweig, K. Multiplex Networks. In *Encyclopedia of Social Network Analysis and Mining*; Alhajj, R., Rokne, J., Eds.; Springer: New York, NY, USA, 2018; pp. 1430–1434.
113. Bródka, P.; Kazienko, P. Multilayer Social Networks. In *Encyclopedia of Social Network Analysis and Mining*; Alhajj, R., Rokne, J., Eds.; Springer: New York, NY, USA, 2018; pp. 1408–1422.
114. Ghafouri, S.; Khasteh, S. A survey on exponential random graph models: An application perspective. *PeerJ Comput. Sci.* **2020**, *6*, e269. [CrossRef]
115. Loyal, J.; Chen, Y. Statistical network analysis: A review with applications to the coronavirus disease 2019 pandemic. *Int. Stat. Rev.* **2020**, *88*, 419–440. [CrossRef]
116. Hammoud, Z.; Kramer, F. Multilayer networks: Aspects, implementations, and application in biomedicine. *Big Data Anal.* **2020**, *5*, 1–18. [CrossRef]
117. Kinsley, A.; Rossi, G.; Silk, M.; VanderWaal, K. Multilayer and multiplex networks: An introduction to their use in veterinary epidemiology. *Front. Vetinary Sci.* **2020**, *7*, 596. [CrossRef]
118. Frank, O.; Strauss, D. Markov graphs. *J. Am. Stat. Assoc.* **1986**, *81*, 832–842. [CrossRef]
119. Wasserman, S.; Pattison, P. Logit models and logistic regressions for social networks: I. An introduction to Markov graphs and *p**. *Psychometrika* **1996**, *61*, 401–425. [CrossRef]
120. Lusher, D.; Koskinen, J.; Robins, G. *Exponential Random Graph Models for Social Networks: Theory, Methods, and Applications*; Cambridge University Press: Cambridge, UK, 2013.
121. Harris, J. *An Introduction to Exponential Random Graph Modeling*; Sage: Thousand Oaks, CA, USA, 2013.
122. Schweinberger, M.; Krivitsky, P.; Butts, C.; Stewart, J. Exponential-family models of random graphs: Inference in finite, super and infinite population scenarios. *Stat. Sci.* **2020**, *35*, 627–662. [CrossRef]
123. Caimo, A.; Gollini, I. Recent advances in exponential random graph modelling. *Math. Proc. R. Ir. Acad.* **2022**, *123*, 1–12. [CrossRef]
124. Krivitsky, P. Exponential-family random graph models for valued networks. *Electron. J. Stat.* **2012**, *6*, 1100–1128. [CrossRef]
125. Krivitsky, P.; Koehly, L.; Marcum, C. Exponential-family random graph models for multi-layer networks. *Psychometrika* **2020**, *85*, 630–659. [CrossRef]
126. Stewart, J.; Schweinberger, M.; Bojanowski, M.; Morris, M. Multilevel network data facilitate statistical inference for curved ERGMs with geometrically weighted terms. *Soc. Netw.* **2019**, *59*, 98–119. [CrossRef] [PubMed]
127. Thiemichen, S.; Friel, N.; Caimo, A.; Kauermann, G. Bayesian exponential random graph models with nodal random effects. *Soc. Netw.* **2016**, *46*, 11–28. [CrossRef]
128. Hanneke, S.; Fu, W.; Xing, E. Discrete temporal models of social networks. *Electron. J. Stat.* **2010**, *4*, 585–605. [CrossRef]
129. Krivitsky, P.; Handcock, M. A separable model for dynamic networks. *J. R. Stat. Soc. Ser. B* **2014**, *76*, 29–46. [CrossRef] [PubMed]
130. Lee, J.; Li, G.; Wilson, J. Varying-coefficient models for dynamic networks. *Comput. Stat. Data Anal.* **2020**, *152*, 107052. [CrossRef]
131. Price, D. Networks of scientific papers. *Science* **1965**, *149*, 510–515. [CrossRef]
132. Price, D. A general theory of bibliometric and other cumulative advantage processes. *J. Am. Soc. Inf. Sci.* **1976**, *27*, 292–306. [CrossRef]
133. Barabasi, A.; Albert, R. Emergence of scaling in random networks. *Science* **1999**, *286*, 509–512. [CrossRef] [PubMed]
134. Vazquez, A.; Flammini, A.; Maritan, A.; Vespignani, A. Modeling of protein interaction networks. *ComPlexUs* **2003**, *1*, 38–44. [CrossRef]
135. Vazquez, A. Growing network with local rules: Preferential attachment, clustering hierarchy, and degree correlations. *Phys. Rev. E* **2003**, *67*, 056104. [CrossRef]
136. Solé, R.; Pastor-Satorras, R.; Smith, E.; Kepler, T. A model of large-scale proteome evolution. *Adv. Complex Syst.* **2002**, *5*, 43–54. [CrossRef]

137. Pastor-Satorras, R.; Smith, E.; Solé, R. Evolving protein interaction networks through gene duplication. *J. Theor. Biol.* **2003**, *222*, 199–210. [CrossRef]
138. Kretzschmar, M.; Morris, M. Measures of concurrency in networks and the spread of infectious disease. *Math. Biosci.* **1996**, *133*, 165–195. [CrossRef]
139. Morris, M.; Kretzschmar, M. Concurrent partnerships and the spread of HIV. *AIDS* **1997**, *11*, 641–648. [CrossRef]
140. Morris, M.; Goodreau, S.; Moody, J. Sexual networks, concurrency, and STD/HIV. In *Sexually Transmitted Diseases*; Holmes, K., Sparling, P., Stamm, W., Eds.; McGraw-Hill Companies: New York, NY, USA, 2007; pp. 109–126.
141. Palombi, L.; Bernava, G.; Nucita, A.; Giglio, P.; Liotta, G.; Nielsen-Saines, K.; Orlando, S.; Mancinelli, S.; Buonomo, E.; Scarcella, P.; et al. Predicting trends in HIV-1 sexual transmission in sub-Saharan Africa through the drug resource enhancement against AIDS and malnutrition model: Antiretrovirals for reduction of population infectivity, incidence and prevalence at the district level. *Clin. Infect. Dis.* **2012**, *55*, 268–275. [CrossRef]
142. Klemm, K.; Eguiluz, V. Highly clustered scale-free networks. *Phys. Rev. E* **2002**, *65*, 036123. [CrossRef] [PubMed]
143. Kumpula, J.; Onnela, J.; Saramäki, J.; Kaski, K.; Kertész, J. Emergence of communities in weighted networks. *Phys. Rev. Lett.* **2007**, *99*, 228701. [CrossRef]
144. Procopio, A.; Cesarelli, G.; Donisi, L.; Merola, A.; Amato, F.; Cosentino, C. Combined mechanistic modeling and machine-learning approaches in systems biology—A systematic literature review. *Comput. Methods Programs Biomed.* **2023**, *240*, 107681. [CrossRef] [PubMed]
145. Tantardini, M.; Ieva, F.; Tajoli, L.; Piccardi, C. Comparing methods for comparing networks. *Sci. Rep.* **2019**, *9*, 17557. [CrossRef]
146. Pržulj, N.; Corneil, D.; Jurisica, I. Modeling interactome: Scale-free or geometric? *Bioinformatics* **2004**, *20*, 3508–3515. [CrossRef]
147. Yaveroğlu, Ö.; Malod-Dognin, N.; Davis, D.; Levnajic, Z.; Janjic, V.; Karapandza, R.; Stojmirovic, A.; Pržulj, N. Revealing the hidden language of complex networks. *Sci. Rep.* **2014**, *4*, 4547. [CrossRef]
148. Yaveroğlu, Ö.; Milenković, T.; Pržulj, N. Proper evaluation of alignment-free network comparison methods. *Bioinformatics* **2015**, *31*, 2697–2704. [CrossRef]
149. Faisal, F.; Newaz, K.; Chaney, J.; Li, J.; Emrich, S.; Clark, P.; Milenković, T. GRAFENE: Graphlet-based alignment-free network approach integrates 3D structural and sequence (residue order) data to improve protein structural comparison. *Sci. Rep.* **2017**, *7*, 14890. [CrossRef]
150. Schmid, C.; Hunter, D. Computing pseudolikelihood estimators for exponential-family random graph models. *J. Data Sci.* **2023**, *21*, 295–309. [CrossRef]
151. Besag, J. Spatial interaction and the statistical analysis of lattice systems (with discussion). *J. R. Stat. Soc. Ser. B* **1974**, *36*, 192–236. [CrossRef]
152. Strauss, D.; Ikeda, M. Pseudolikelihood estimation for social networks. *J. Am. Stat. Assoc.* **1990**, *85*, 204–212. [CrossRef]
153. Lindsay, B. Composite likelihood methods. *Contemp. Math.* **1988**, *80*, 221–239.
154. Varin, C.; Reid, N.; Firth, D. An overview of composite likelihood methods. *Stat. Sin.* **2011**, *21*, 5–42.
155. Arnold, B.; Strauss, D. Pseudolikelihood estimation: Some examples. *Sankhya Ser. B* **1991**, *53*, 233–243.
156. Schmid, C.; Desmarais, B. Exponential random graph models with big networks: Maximum pseudolikelihood estimation and the parametric bootstrap. In Proceedings of the 2017 IEEE International Conference on Big Data, Boston, MA, USA, 11–14 December 2017; pp. 116–121.
157. Gao, Y.; Liu, W.; Wang, H.; Wang, X.; Yan, Y.; Zhang, R. A review of distributed statistical inference. *Stat. Theory Relat. Fields* **2022**, *6*, 89–99. [CrossRef]
158. Rosenblatt, J.; Nadler, B. On the optimality of averaging in distributed statistical learning. *Inf. Inference J. IMA* **2016**, *5*, 379–404. [CrossRef]
159. Minsker, S. Distributed statistical estimation and rates of convergence in normal approximation. *Electron. J. Stat.* **2019**, *13*, 5213–5252. [CrossRef]
160. Luo, L.; Song, P. Renewable estimation and incremental inference in generalized linear models with streaming data sets. *J. R. Stat. Soc. Ser. B* **2019**, *82*, 69–97. [CrossRef]
161. Raynal, L.; Chen, S.; Mira, A.; Onnela, J. Scalable Approximate Bayesian Computation for growing network models via extrapolated and sampled summaries. *Bayesian Anal.* **2022**, *17*, 165–192. [CrossRef]
162. Krapivsky, P.; Redner, S.; Leyvraz, F. Connectivity of Growing Random Networks. *Phys. Rev. Lett.* **2000**, *85*, 4629–4632. [CrossRef]
163. Rossi, R.; Ahmed, N. The network data repository with interactive graph analytics and visualization. In Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence, Austin, TX, USA, 25–30 January 2015.
164. De Domenico, M.; Altmann, E. Unraveling the origin of social bursts in collective attention. *Sci. Rep.* **2020**, *10*, 4629. [CrossRef]
165. Yang, J.; Leskovec, J. Defining and evaluating network communities based on ground-truth. In Proceedings of the ACM SIGKDD Workshop on Mining Data Semantics, New York, NY, USA, 12–16 August 2012; MDS '12.
166. Dinh, K.; Xiang, Z.; Liu, Z.; Tavaré, S. Approximate Bayesian Computation sequential Monte Carlo via random forests. *arXiv* **2024**, arXiv:2406.15865.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.