Supplement to "Bayesian modelling and computation utilising directed cycles in multiple network data"

Anastasia Mantziou[1*], Sally Keith[2], David M.P. Jacoby[2], Simón Lunagómez[3] and Robin Mitra[4]

[1*]Department of Statistics, University of Warwick, Coventry, CV4 7AL, U.K..
[2]Lancaster Environment Centre, Lancaster University, Lancaster, LA1 4YQ, U.K..
[3]Department of Statistics, ITAM, Río Hondo, México, 01080.
[4]Department of Statistical Science, University College London, Gower Street, London, WC1E 6BT, U.K..

*Corresponding author(s). E-mail(s): anastasia.mantziou@warwick.ac.uk;
Contributing authors: sally.keith@lancaster.ac.uk; d.jacoby@lancaster.ac.uk; simon.lunagomez@itam.mx; robin.mitra@ucl.ac.uk;

Keywords: Doubly intractable distributions, Importance Sampling, Object data analysis, Relational data.

1 Proof HS is a distance metric

The HS measure is the weighted sum of the Hamming distance and the symmetric difference of cycles between two graphs. The Hamming distance is a well-known distance metric, thus, to prove that the HS measure is also a distance metric, we need to prove that the symmetric difference between graphs' cycles is a distance metric.

Let $\mathcal{C}_n$ be the set of cycles for graphs of size n, and each $C_{\mathcal{G}_i}, C_{\mathcal{G}_j}, C_{\mathcal{G}_k} \in \mathcal{C}_n$ be the subset of cycles found in graphs $\mathcal{G}_i$, $\mathcal{G}_j$ and $\mathcal{G}_k$ respectively. Thence, the symmetric difference of the cycles of two graphs is $d_{symm} =\mid C_{\mathcal{G}_\cdot} \Delta C_{\mathcal{G}_\cdot} \mid$. The function $d_{symm} : \mathcal{C}_n \times \mathcal{C}_n \to [0, \infty)$ is a distance metric if the following conditions are satisfied:

1. $d_{symm}(C_{\mathcal{G}_i}, C_{\mathcal{G}_j}) = 0 \Leftrightarrow C_{\mathcal{G}_i} = C_{\mathcal{G}_j}$
2. $d_{symm}(C_{\mathcal{G}_i}, C_{\mathcal{G}_j}) = d_{symm}(C_{\mathcal{G}_j}, C_{\mathcal{G}_i})$
3. $d_{symm}(C_{\mathcal{G}_i}, C_{\mathcal{G}_j}) \leq d_{symm}(C_{\mathcal{G}_i}, C_{\mathcal{G}_k}) + d_{symm}(C_{\mathcal{G}_k}, C_{\mathcal{G}_j})$

Conditions 1 and 2 are clearly satisfied. Thus, we need to prove that the triangle inequality holds for the symmetric difference of cycles. The symmetric difference has the following property,

$$C_{\mathcal{G}_i} \Delta C_{\mathcal{G}_j} = (C_{\mathcal{G}_i} \Delta C_{\mathcal{G}_k}) \Delta (C_{\mathcal{G}_k} \Delta C_{\mathcal{G}_j}).$$

It follows that

$$\begin{gathered} C_{\mathcal{G}_i} \Delta C_{\mathcal{G}_j} \subseteq (C_{\mathcal{G}_i} \Delta C_{\mathcal{G}_k}) \cup (C_{\mathcal{G}_k} \Delta C_{\mathcal{G}_j}) \Rightarrow \\ \mid C_{\mathcal{G}_i} \Delta C_{\mathcal{G}_j} \mid \leq \mid C_{\mathcal{G}_i} \Delta C_{\mathcal{G}_k} \mid + \mid C_{\mathcal{G}_k} \Delta C_{\mathcal{G}_j} \mid . \end{gathered}$$

Thus condition 3 is satisfied for the symmetric difference of cycles between graphs.

2 Additional details for the Proposed Bayesian inference framework for the SNF model using Importance Sampling

We now present additional details on the inferential scheme used to obtain draws from the posterior distributions of the parameters of the SNF model, as discussed in Section 5.2 of the main article. Notably, we update the adjacency matrix of the centroid $A_{\mathcal{G}^m}$ using either of the following two proposals,

(I) We perturb the edges of the current centroid $A_{\mathcal{G}^m}^{(curr)}$ as follows:

$$A_{\mathcal{G}^m}^{(prop)}(i,j) = \begin{cases} 1 - A_{\mathcal{G}^m}^{(curr)}(i,j), & \text{with probability } \omega \\ A_{\mathcal{G}^m}^{(curr)}(i,j), & \text{with probability } 1-\omega \end{cases}.$$

(II) We propose a new network representative $A_{\mathcal{G}^m}^{(prop)}$, with each edge of the proposed representative being drawn independently from a Bernoulli distribution with parameter $\frac{1}{N}\sum_{l=1}^{N} A_{\mathcal{G}_l}(i,j)$, where $\{A_{\mathcal{G}_l}\}_{l=1}^{N}$ denoting the N observed networks.

Under case (I), we accept the proposed network representative $A_{\mathcal{G}^m}^{(prop)}$ with probability

$$\min\left\{1, \frac{\widehat{Z}(A_{\mathcal{G}^m}^{(prop)}, \gamma^{(curr)})^{-N} \exp\{-\gamma^{(curr)} \sum_{i=1}^N d_{\mathcal{G}}(A_{\mathcal{G}_i}, A_{\mathcal{G}^m}^{(prop)})\}}{\widehat{Z}(A_{\mathcal{G}^m}^{(curr)}, \gamma^{(curr)})^{-N} \exp\{-\gamma^{(curr)} \sum_{i=1}^N d_{\mathcal{G}}(A_{\mathcal{G}_i}, A_{\mathcal{G}^m}^{(curr)})\}} \cdot \frac{\exp\left\{-\gamma_0 d_{\mathcal{G}}(A_{\mathcal{G}^m}^{(prop)}, A_{\mathcal{G}_0})\right\}}{\exp\left\{-\gamma_0 d_{\mathcal{G}}(A_{\mathcal{G}^m}^{(curr)}, A_{\mathcal{G}_0})\right\}}\right\},$$

while under case (II), we accept the proposed network representative $A_{\mathcal{G}^m}^{(prop)}$ with probability

$$\min\left\{1, \frac{\frac{\exp\{-\gamma^{(curr)} \sum_{i=1}^N d_{\mathcal{G}}(A_{\mathcal{G}_i}, A_{\mathcal{G}^m}^{(prop)})\}}{\widehat{Z}(A_{\mathcal{G}^m}^{(prop)}, \gamma^{(curr)})^N} \exp\left\{-\gamma_0 d_{\mathcal{G}}(A_{\mathcal{G}^m}^{(prop)}, A_{\mathcal{G}_0})\right\}}{\frac{\exp\{-\gamma^{(curr)} \sum_{i=1}^N d_{\mathcal{G}}(A_{\mathcal{G}_i}, A_{\mathcal{G}^m}^{(curr)})\}}{\widehat{Z}(A_{\mathcal{G}^m}^{(curr)}, \gamma^{(curr)})^N} \exp\left\{-\gamma_0 d_{\mathcal{G}}(A_{\mathcal{G}^m}^{(curr)}, A_{\mathcal{G}_0})\right\}} \cdot \frac{Q(A_{\mathcal{G}^m}^{(curr)} \mid A_{\mathcal{G}^m}^{(prop)})}{Q(A_{\mathcal{G}^m}^{(prop)} \mid A_{\mathcal{G}^m}^{(curr)})}\right\},$$

We note here that the proposal distribution under case (I) is symmetric, and thus it cancels out from the Metropolis ratio, while under case (II) the proposal distribution $Q(A_{\mathcal{G}^m}^{(\cdot)} \mid A_{\mathcal{G}^m}^{(\cdot)})$ does not cancel.

Accordingly, we use a mixture of K random walks to propose values for the dispersion parameter γ, as follows:

1. Draw a uniform random variable $u \sim \text{Unif}(-v_k, v_k)$, with k indicating the k^{th} proposal.
2. Perturb the current state $\gamma^{(curr)}$ by the uniform random variable drawn, $y = \gamma^{(curr)} + u$.
3. The newly proposed value for γ is $\gamma^{(prop)} = \begin{cases} y, \text{if } y > 0 \\ -y, \text{if } y < 0 \end{cases}$,

which we accept with probability

$$\min\left\{1, \frac{\widehat{Z}(A_{\mathcal{G}^m}^{(curr)}, \gamma^{(prop)})^{-N} \exp\{-\gamma^{(prop)} \sum_{i=1}^N d_{\mathcal{G}}(A_{\mathcal{G}_i}, A_{\mathcal{G}^m}^{(curr)})\}}{\widehat{Z}(A_{\mathcal{G}^m}^{(curr)}, \gamma^{(curr)})^{-N} \exp\{-\gamma^{(curr)} \sum_{i=1}^N d_{\mathcal{G}}(A_{\mathcal{G}_i}, A_{\mathcal{G}^m}^{(curr)})\}} \cdot \frac{P(\gamma^{(prop)} \mid \alpha_0)}{P(\gamma^{(curr)} \mid \alpha_0)}\right\}.$$

Under this scheme, in each iteration of the MCMC algorithm, we draw a new sample from the IS density to calculate $\widehat{Z}$ in the numerator and denominator of the MH ratio, as detailed in Sections 5.1 and 5.2 of the main article.

3 Additional details for real data application

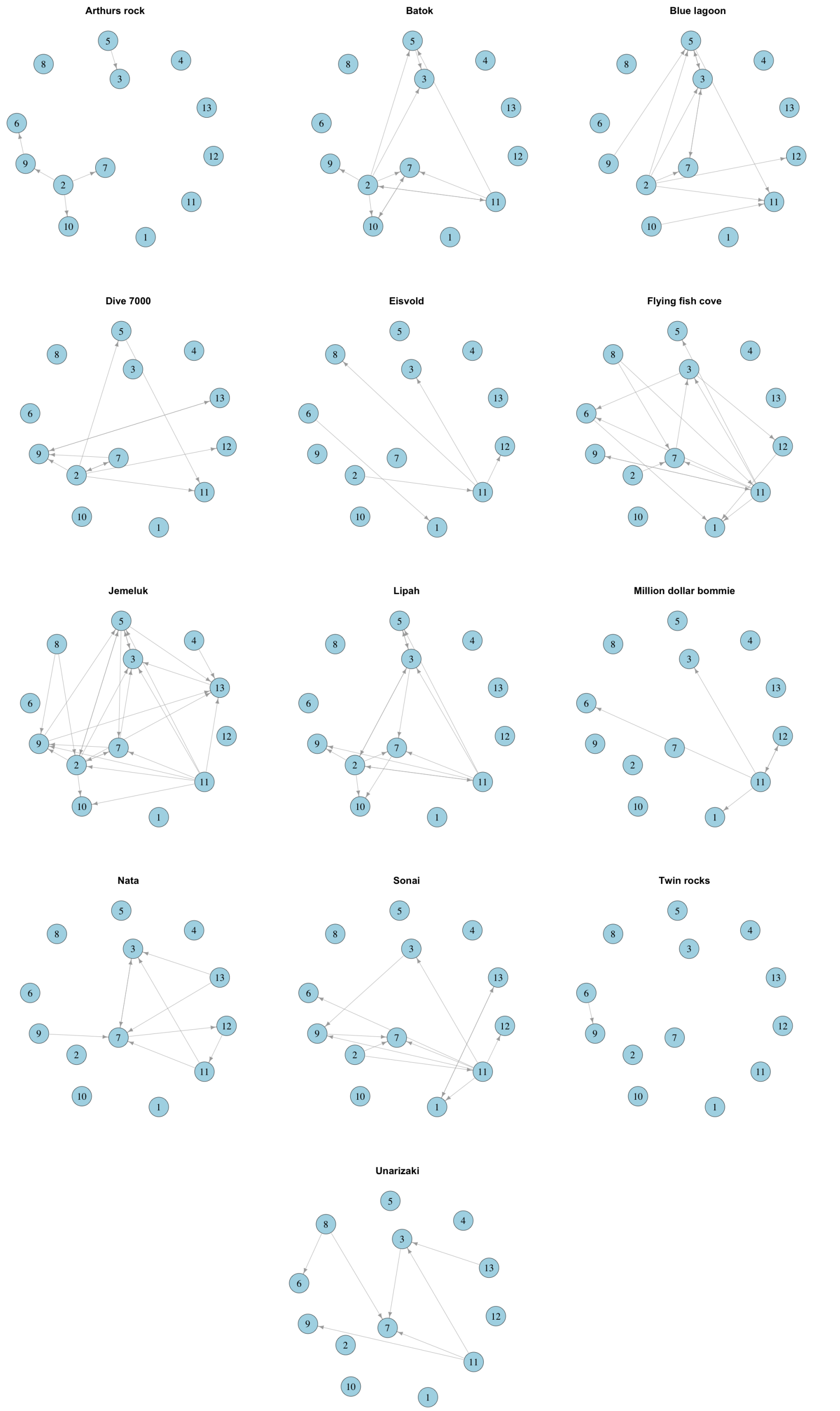

Fig. 1 Network population of fish aggressive interactions with each network representing a reef, at different regions in the Indo-Pacific ocean.