# Belief propagation on networks with cliques and chordless cycles

Peter Mann* and Simon Dobson
*School of Computer Science, University of St Andrews, St Andrews, Fife KY16 9SX, United Kingdom*

(Dated: April 16, 2023)

It is well known that tree-based theories can describe the properties of undirected clustered networks with extremely accurate results [S. Melnik, *et al.* Phys. Rev. E 83, 036112 (2011)]. It is reasonable to suggest that a motif based theory would be superior to a tree one; since additional neighbour correlations are encapsulated in the motif structure. In this paper we examine bond percolation on random and real world networks using belief propagation in conjunction with edge-disjoint motif covers. We derive exact message passing expressions for cliques and chordless cycles of finite size. Our theoretical model gives good agreement with Monte Carlo simulation and offers a simple, yet substantial improvement on traditional message passing showing that this approach is suitable to study the properties of random and empirical networks.

## I. INTRODUCTION

Belief propagation, also known as message passing, is an algorithm that is fundamental to many different disciplines including physics, computer science, epidemiology and statistics [1–8]. The belief propagation algorithm, which relies on the Bethe-Peierls approximation [9–11], is valid only for large and sparse treelike networks; becoming exact when the network *is* a tree. This is because the removal of a vertex, $i$, and its edges from a tree isolates its neighbours from one another; the only path connecting them has vanished and so too do inter-neighbour correlations. Such a graph is called the *cavity graph* of vertex $i$ and allows a message passing system of self-consistent equations to be written for the vertices of the network. Loops in networks introduce correlations between vertices that invalidates the Bethe-Peierls approximation as there is still a path between the neighbours in the cavity graph of a given vertex. Despite this drawback, belief propagation has a proven ability to analytically describe the structural properties of many real world networks, which are known to contain loops [12]. Belief propagation on clustered networks has been studied previously [1, 3, 4, 7, 8, 13–23] and an immense literature exists that tackles the problem in a variety of different manners.

Recently, Cantwell, Newman and Kirkley solved the belief propagation model for networks with arbitrary loop structure [3, 4]. They achieved this by considering messages from a collective local neighbourhood, of variable distance, about each vertex, rather than a partition into recognised motifs. In their framework, it is assumed that all inter-vertex correlations due to short range loops are encapsulated within the neighbourhood. For a given configuration of edges, the probability of the size of the component to which a vertex belongs is calculated, before being averaged over the probability of each edge configuration for the neighbourhood. The neighbourhood calculation is computationally expensive and so the authors introduced a Monte Carlo algorithm to sample the set of connected graphs.

It is not unreasonable, however, to desire a *model* of an empirical network to carry out further experimentation on. For instance, suppose that the empirical network represents a data set that is too small to apply some statistical or machine learning algorithm. If we had a model of the network, perhaps we could understand how to grow the network, adding new vertices and edges, whilst keeping the inherent topology and statistical properties fixed.

One way to obtain a model is to cover the network in motifs; specifically, a set of *edge-disjoint* motifs, such that a given edge belongs to only one motif. The simplest cover is to simply assume each edge is a 2-clique [12] - and this approximation certainly works well in some cases. However, given the large body of knowledge for random clustered networks [17, 24–33] it is tempting to apply covers with larger and more complicated motifs in the hope to obtain more accurate models. For sparse random graphs, this technique works well and the success lies in the locally treelike nature of the *factor graph* of the covered network. The factor graph is a bipartite graph that has two different sets of vertices. One set represents the vertices of the substrate network, whilst the other represents the motifs to which they belong. Edges connect vertices in the original graph to the vertices representing the motifs to which they belong, see Fig 1. If the factor graph is locally treelike then all of the short range loops in the network are encapsulated within the motif cover.

Unfortunately, naive motif covers of empirical networks often do not create representative models. Cantwell, Kirkley and Newman summarise that "*these techniques are not generally applicable to real world networks*" [4]. In fact, it is sometimes the case that tree-based theories give the closest match to the empirical network. This is despite a cover with larger motifs necessarily having a more treelike factor graph than a cover comprised of only 2-cliques. This indicates the presence of an additional driving force for a suitable motif cover beyond simply decreasing the number of loops in the factor graph.

In this paper, we introduce exact expressions for belief propagation on networks that are composed of edge-disjoint chordless cycles and cliques of finite size. These networks can be obtained by covering an empirical network with motifs or constructively using the configuration model [24, 31]. We then study message passing on empirical clique covered networks. We find that our model exhibits excellent agreement with Monte Carlo simulation beyond the results obtained from ordinary belief propagation [12] due to the arbitrarily large clique sizes we can analytically include in our model. We offer insight into why naive motif covers often fail to capture the properties of real world networks by considering the statistical bias introduced by the breaking of symmetry in certain covers. Finally, we relate our method to the generalised configuration model which is shown to be the ensemble average of the message passing model.

This work will lead to significant advances in the application of the theory of clustered graphs to the study of empirical networks. We hope that additional studies will be conducted with dynamics other than bond percolation to further these results.

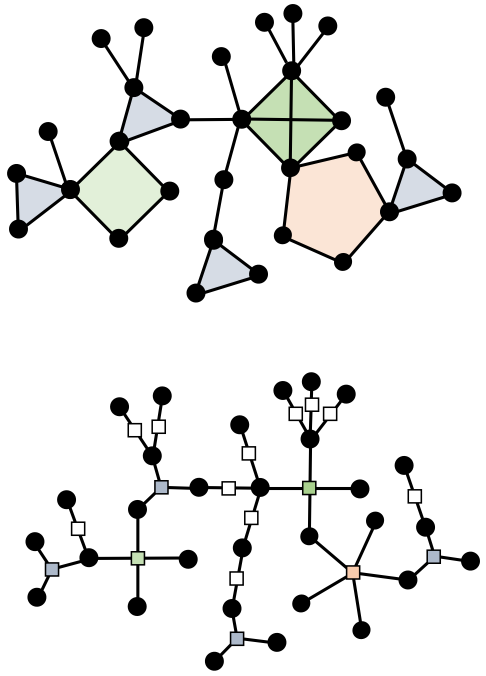


FIG. 1. Top: A network is covered with edge-disjoint chordless cycles and cliques. Bottom: the factor graph of the network. Figure inspired by Figure 1 of [18]. Such networks can have a high local density of loops that encapsulate the neighbour correlation, becoming increasingly sparse and treelike at long ranges.

## II. THEORETICAL

Let $G = (N, E)$ be a graph, and $\omega : E \to \{0, 1\}$ be a *edge configuration* on the graph that maps each edge to a value of either 0 or 1 with probabilities $1 - \phi$ and $\phi$, respectively for $\phi \in [0, 1]$. An edge is *occupied* if $\omega(e) = 1$ for $e \in E$, and *unoccupied* if $\omega(e) = 0$; as $\phi$ is increased a giant connected component of occupied edges emerges through a phase transition. The state space of the model is $\Omega = \{0, 1\}^E$ such that $\omega \in \Omega$ are $E$-dimensional vectors. Let $\eta(\omega) = \{e \in E : \omega(e) = 1\}$ denote the set of occupied edges. For a given network $G$ the probability measure of an edge configuration is

$$\mu(\omega \mid G) = \frac{1}{Z} \prod_{e\in E} \phi^{\omega(e)} (1 - \phi)^{\omega(e)}, \tag{1}$$

where $Z$ is the partition function

$$Z = \sum_{\omega\in\Omega} \prod_{e\in E} \phi^{\omega(e)} (1 - \phi)^{\omega(e)}. \tag{2}$$

This summation is over all possible edge configurations of a given graph, weighted by their probabilities and therefore, is equal to one. When $\phi$ is small, the graph is composed of many small clusters of occupied vertices and there is no long range connectivity. As $\phi$ is increased a giant percolating cluster emerges which incorporates a finite fraction of the vertices $\mathcal{O}(N)$.

We now derive the belief propagation formulation [34] for motif covered networks. Our model also applies to random configuration model networks that are composed of cliques or chordless cycles, such as those described in section I. Similar expressions appear in [3] and [17]; however these frameworks correspond to neighbourhood decomposed networks; and for the Ising model on motif covered networks, respectively. Let us select a vertex $i$ at random from the equilibrium of the percolation process and suppose that $i$ is the corner of a set of motifs $\boldsymbol{\tau}_i$. The probability $\pi_i(s)$ is the probability that $i$ belongs to a non-percolating component of size $s$. This can be written in terms of the probability $\pi_{i\leftarrow j}$ that a neighbour of $i$, vertex $j$, which is in motif $\tau$ together with $i$, leads to $s_j$ vertices if we were to follow all of its edges, other than those that point back to motif $\tau$

$$\begin{aligned}\pi_i(s) = \sum_{\{s_j : j\in\partial_\tau(i),\ \tau\in\boldsymbol{\tau}_i\}} & \left[ \prod_{\tau\in\boldsymbol{\tau}_i} \prod_{j\in\partial_\tau(i)} \pi_{i\leftarrow j}(s_j) \right] \\ & \times \delta\left( s - 1, \sum_{\tau\in\boldsymbol{\tau}_i} \sum_{j\in\partial_\tau(i)} s_j \right), \end{aligned} \tag{3}$$

where $\tau$ is a motif belonging to $\boldsymbol{\tau}_i$, $\partial_\tau(i)$ is the set of edges vertex $i$ has within motif $\tau$ and where $\delta(x, y)$ is the Kronecker delta. This expression averages over every combination of the total number of reachable vertices $s_j$ along each edge $j$ of each motif $\tau$ in the set of motifs $\boldsymbol{\tau}_i$ that $i$ belongs to. This summation is then filtered by the Kronecker delta to retain only those terms that collectively sum to $s - 1$, to which we add 1 for vertex $i$ itself to yield the component of overall size $s$. We can generate this probability by defining

$$G_i(z) = \sum_{s=1}^{\infty} \pi_i(s) z^s \tag{4}$$

such that

$$G_i(z) = \sum_{s=1}^{\infty} z^s \sum_{\{s_j : j \in \partial_\tau(i),\ \tau \in \boldsymbol{\tau}_i\}} \left[ \prod_{\tau \in \boldsymbol{\tau}_i} \prod_{j \in \partial_\tau(i)} \pi_{i \leftarrow j}(s_j) \right] \times \delta\left( s-1, \sum_{\tau \in \boldsymbol{\tau}_i} \sum_{j \in \partial_\tau(i)} s_j \right). \tag{5}$$

With the following manipulation

$$\sum_{s=1}^{\infty} z^s = z \sum_{s=1}^{\infty} z^{s-1} \overset{\delta}{=} z \sum_{s=1}^{\infty} z^{\sum_\tau \sum_{j \in \partial_\tau(i)} s_j} = z \prod_{\tau \in \boldsymbol{\tau}_i} \prod_{j \in \partial_\tau(i)} \sum_{s_j=0}^{\infty} z^{s_j}, \tag{6}$$

we obtain

$$G_i(z) = z \prod_{\tau \in \boldsymbol{\tau}_i} \prod_{j \in \partial_\tau(i)} \sum_{s_j=0}^{\infty} \pi_{i \leftarrow j}(s_j) z^{s_j}. \tag{7}$$

The summation limits over $s_j$ are $0 \to \infty$; since, the original summation is over $s \in [1, \infty]$ and so $s-1$ is $s_j \in [0, \infty]$. For each neighbour $j$ we can write a generating function for the probability that the number of vertices that can be reached, other than through the motif itself, is $s_j$ as

$$H_{i \leftarrow j}(z) = \sum_{s_j=0}^{\infty} \pi_{i \leftarrow j}(s_j) z^{s_j}. \tag{8}$$

The probability for the total number of vertices that can be reached via motif $\tau$ is simply the product

$$H_{i \leftarrow \tau}(z) = \prod_{j \in \partial_\tau(i)} H_{i \leftarrow j}(z). \tag{9}$$

Inserting these expressions in Eq 7 we find

$$G_i(z) = z \prod_{\tau \in \boldsymbol{\tau}_i} H_{i \leftarrow \tau}(z). \tag{10}$$

This expression is a general result and holds for arbitrary motif topologies. If we had knowledge of each $H_{i \leftarrow \tau}(z)$ this generating function would yield the distribution of component sizes that vertex $i$ belongs to. From this, for instance, we can find the size of the percolating component $S$ as one minus the probability that all vertices $i \in N$ belong to finite components

$$S = 1 - \frac{1}{N} \sum_i \prod_{\tau \in \boldsymbol{\tau}_i} H_{i \leftarrow \tau}(1). \tag{11}$$

The average size of the finite components $\langle s_i \rangle$ is given by

$$\langle s_i \rangle = \frac{\sum_s s \pi_i(s)}{\sum_s \pi_i(s)} = \frac{G_i'(1)}{G_i(1)}, \tag{12}$$

where we have used Eq 4. Inserting Eq 10 we have

$$\langle s_i \rangle = 1 + \sum_{\tau \in \boldsymbol{\tau}_i} \frac{H_{i \leftarrow \tau}'(1)}{H_{i \leftarrow \tau}(1)}, \tag{13}$$

where we have used

$$\frac{d}{dz}\left[ \prod_\tau H_{i \leftarrow \tau}(z) \right] = \sum_\tau \left[ \left( \frac{d}{dz} H_{i \leftarrow \tau}(z) \right) \prod_{\nu \neq \tau} H_{i \leftarrow \nu}(z) \right] = \left( \prod_\tau H_{i \leftarrow \tau}(z) \right) \left( \sum_\tau \frac{H_{i \leftarrow \tau}'(z)}{H_{i \leftarrow \tau}(z)} \right). \tag{14}$$

To progress we require the derivative of $H_{i \leftarrow \tau}(z)$. However, the functional form of $H_{i \leftarrow \tau}(z)$, crucial to enumerating $G_i(z)$, depends on the structure of the motif and the configuration of its edges among the occupied and unoccupied states. We will now examine $H_{i \leftarrow \tau}(z)$ for ordinary edges, chordless cycles of arbitrary length and cliques of arbitrary size.

### A. Calculating $H_{i \leftarrow \tau}$ for ordinary edges

Let us for a moment assume that the network contains only 2-cliques - ordinary edges; which was previously studied by Karrer *et al* [12]. In this case Eq 10 reduces to

$$G_i(z) = z \prod_j H_{i \leftarrow j}(z), \tag{15}$$

where the product over $j$ accounts for each distinct 2-clique $i$ belongs to. Because there is only a single edge in a 2-clique, this index equivalently runs over the neighbour vertices $H_{i \leftarrow \tau}(z) = H_{i \leftarrow j}(z)$. If the edge is unoccupied with probability $1 - \phi$, then $\pi_{i \leftarrow j}(s)$ is zero. In this case, the edge does not contribute to the component size of $i$ and so $s = 0$. If the edge is occupied with probability $\phi$, then $\pi_{i \leftarrow j}(s \geq 1)$ is non-zero. Therefore, the summation in Eq 8 has the zero term $\pi_{i \leftarrow j}(0) = 1 - \phi$ extracted, and for $s \geq 1$ we have

$$\pi_{i \leftarrow j}(s) = \phi \sum_{\{s_k : k \in \partial_\nu(j) \backslash i\}} \left[ \prod_{k \in \partial(j) \backslash i} \pi_{j \leftarrow k}(s_k) \right] \times \delta\left( s-1, \sum_{k \in \partial(j) \backslash i} s_k \right), \tag{16}$$

where the notation $\partial(j) \backslash i$ denotes the set edges to which vertex $j$ belongs to excluding the one connecting it to $i$. Substituting this expression into Eq 8 we have

$$H_{i \leftarrow j}(z) = 1 - \phi + \phi \sum_{s=1}^{\infty} z^s \sum_{\{s_k : k \in \partial(j) \backslash i\}} \left[ \prod_{k \in \partial(j) \backslash i} \pi_{j \leftarrow k}(s_k) \right] \times \delta\left( s-1, \sum_{k \in \partial(j) \backslash i} s_k \right). \tag{17}$$

Application of the Kronecker delta yields

$$H_{i \leftarrow j}(z) = (1 - \phi) + \phi z \prod_{k \in \partial(j) \backslash i} \sum_{s_k=0}^{\infty} \pi_{j \leftarrow k}(s_k) z^{s_k}. \tag{18}$$

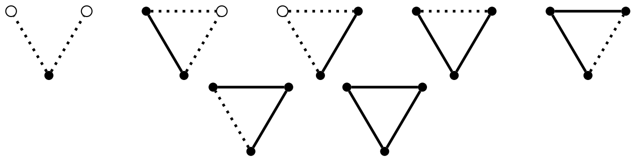


FIG. 2. The seven edge configurations $\mathcal{C}$ that a focal vertex $i$ (bottom) can encounter when it belongs to a triangle with vertices $j$ (top left vertex) and $k$ (top right vertex). Solid lines represent occupied edges whilst dashed lines are unoccupied. Solid vertices belong to the same component as $i$, whilst unfilled vertices are not connected to $i$. When both of $i$'s edges are unoccupied (top left triangle) the state of the edge between the neighbours is inconsequential to the percolation properties of $i$.

Noticing that the final summation is the generating function of $\pi_{j\leftarrow k}$ (Eq 8) we have a self-consistent expression for $H_{i\leftarrow j}(z)$

$$H_{i\leftarrow j}(z) = (1-\phi) + \phi z \prod_{k\in\partial(j)\setminus i} H_{j\leftarrow k}(z). \tag{19}$$

This expression can be solved for all edges by fixed point iteration from a suitable starting point and substituted into Eq 10 to find $G_i(z)$. The distribution for $\pi_i(s)$ can then be found by differentiating the series.

To evaluate the finite components, we require the derivative of $H_{i\leftarrow j}(z)$ with respect to $z$. Following [12] we find

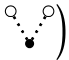
$$H'_{i\leftarrow j}(1) = \phi\left[1 + \sum_{k\in\partial(j)\setminus i} \frac{H'_{j\leftarrow k}(1)}{H_{j\leftarrow k}(1)}\right] \prod_{k\in\partial(j)\setminus i} H_{j\leftarrow k}(1). \tag{20}$$

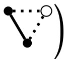
To find the average size of the finite components we solve Eqs 19 and 20 and insert them into Eq 13.

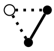
## B. Calculating $H_{i\leftarrow\tau}$ for 3-cliques

Let us now examine the case where vertices belong to triangles as well as ordinary edges. Each triangle $\tau$ that vertex $i$ belongs to connects it to vertices $j$ and $k$; and, our aim is to calculate the generating function $H_{i\leftarrow\tau}(z)$ for the probability $\pi_{i\leftarrow\tau}(s_\tau)$ that the number of vertices that are reachable from vertex $i$ due to its membership in triangle $\tau$ is $s_\tau$. However, unlike ordinary edges, which contain just 1 edge, a triangle has 3 edges. Each of these edges may be occupied $\omega(e) = 1$, or unoccupied $\omega(e) = 0$ and there are 7 distinct edge configurations $\mathcal{C}$ for a given triangle that bear impact on the connectivity of vertex $i$, see Fig 2.

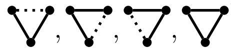
The probability that the triangle connects $i$ to $s_\tau$ vertices is dependent on the edge configuration that the triangle is in. Therefore, in order to compute $\pi_{i\leftarrow\tau}(s_\tau)$, we must calculate conditional probabilities $\pi_{i\leftarrow\tau}(s_\tau \mid \mathcal{C})$ of leading to $s_\tau$ vertices for a given edge configuration $\mathcal{C}$. We find

$\pi_{i\leftarrow\tau}\Big(s_\tau \mid$  $\Big) = \delta(s_\tau, 0),$ (21a)

$\pi_{i\leftarrow\tau}\Big(s_\tau \mid$  $\Big) = \displaystyle\sum_{\{s_\ell:\ell\in\partial_\nu(j),\ \nu\in\boldsymbol{\nu}_j\setminus\tau\}} \left[\prod_{\nu\in\boldsymbol{\nu}_j\setminus\tau}\prod_{\ell\in\partial_\nu(j)} \pi_{j\leftarrow\ell}(s_\ell)\right] \delta\left(s_\tau - 1, \sum_{\nu\in\boldsymbol{\nu}_j\setminus\tau}\sum_{\ell\in\partial_\nu(j)} s_\ell\right),$ (21b)

$\pi_{i\leftarrow\tau}\Big(s_\tau \mid$  $\Big) = \displaystyle\sum_{\{s_r:r\in\partial_\varphi(k),\ \varphi\in\boldsymbol{\varphi}_k\setminus\tau\}} \left[\prod_{\varphi\in\boldsymbol{\varphi}_k\setminus\tau}\prod_{r\in\partial_\varphi(k)} \pi_{k\leftarrow r}(s_r)\right] \delta\left(s_\tau - 1, \sum_{\varphi\in\boldsymbol{\varphi}_k\setminus\tau}\sum_{r\in\partial_\varphi(k)} s_r\right),$ (21c)

where the notation $\boldsymbol{\nu}_j\setminus\tau$ in Eq 21b denotes the set of motifs that vertex $j$ belongs to excluding $\tau$; and similarly for $\boldsymbol{\varphi}_k\setminus\tau$ in Eq 21c. When the focal vertex is connected to *both* $j$ and $k$ in $\tau$ via occupied edges, we must average over both sets of neighbours, see Fig 3.

$\pi_{i\leftarrow\tau}\Big(s_\tau \mid$  $\Big) = \displaystyle\sum_{\{s_\ell:\ell\in\partial_\nu(j),\ \nu\in\boldsymbol{\nu}_j\setminus\tau\}}\ \sum_{\{s_r:r\in\partial_\varphi(k),\ \varphi\in\boldsymbol{\varphi}_k\setminus\tau\}}$

$\displaystyle\times\left[\prod_{\nu\in\boldsymbol{\nu}_j\setminus\tau}\prod_{\ell\in\partial_\nu(j)} \pi_{j\leftarrow\ell}(s_\ell) \prod_{\varphi\in\boldsymbol{\varphi}_k\setminus\tau}\prod_{r\in\partial_\varphi(k)} \pi_{k\leftarrow r}(s_r)\right]$

$\displaystyle\times\,\delta\left(s_\tau - 2, \sum_{\nu\in\boldsymbol{\nu}_j\setminus\tau}\sum_{\ell\in\partial_\nu(j)} s_\ell + \sum_{\varphi\in\boldsymbol{\varphi}_k\setminus\tau}\sum_{r\in\partial_\varphi(k)} s_r\right).$ (22)

The total probability that triangle $\tau$ leads to $s_\tau$ vertices is found by summing each conditional probability with the

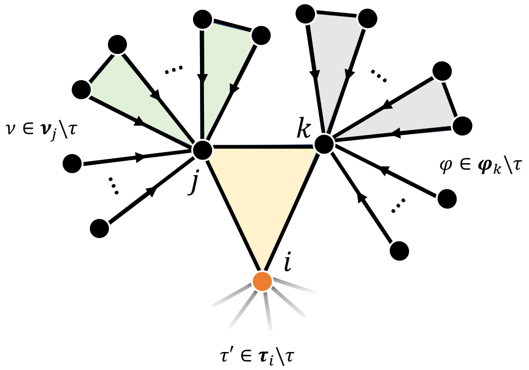

$\nu \in \boldsymbol{\nu}_j \backslash \tau$
$k$
$j$
$\varphi \in \boldsymbol{\varphi}_k \backslash \tau$
$i$
$\tau' \in \boldsymbol{\tau}_i \backslash \tau$

FIG. 3. The probability that triangle $\tau$ connects vertex $i$ to $s_\tau$ vertices, when both $j$ and $k$ belong to the same component as $i$, is equal to the probability that the number of neighbouring vertices of $j$ and $k$, apart from those in $\tau$ itself, collectively connect to $s_\tau - 2$ vertices, see Eq 22.

probability of each edge configuration [24, 26, 29]

$$\pi_{i\leftarrow\tau}(s_\tau) = (1-\phi)^2 \pi_{i\leftarrow j}\left(0 \mid \triangledown\right) + \phi(1-\phi)^2\left[\pi_{i\leftarrow j}\left(s_\tau \mid \triangledown\right) + \pi_{i\leftarrow k}\left(s_\tau \mid \triangledown\right)\right] \\ + \left[3\phi^2(1-\phi) + \phi^3\right]\pi_{i\leftarrow\tau}\left(s_\tau \mid \triangledown, \triangledown, \triangledown, \triangledown\right). \tag{23}$$

Extracting the $s_\tau = 0$ term, we generate this expression as

$$H_{i\leftarrow\tau}(z) = (1-\phi)^2 + \sum_{s_\tau=1}^{\infty} z^{s_\tau}\pi_{i\leftarrow\tau}(s_\tau). \tag{24}$$

Summing over $s_\tau$, expressions 21b and 21c follow similar manipulations to Eq 6; whilst Eq 22 becomes

$$\sum_{s_\tau=1}^{\infty} z^{s_\tau}\pi_{i\leftarrow\tau}\left(s_\tau \mid \triangledown, \triangledown, \triangledown, \triangledown\right) = \sum_{s_\tau=1}^{\infty} z^{s_\tau} \sum_{\{s_\ell : \ell\in\partial_\nu(j),\ \nu\in\boldsymbol{\nu}_j\backslash\tau\}} \ \sum_{\{s_r : r\in\partial_\varphi(k),\ \varphi\in\boldsymbol{\varphi}_k\backslash\tau\}} \\ \times\left[\prod_{\nu\in\boldsymbol{\nu}_j\backslash\tau}\prod_{\ell\in\partial_\nu(j)}\pi_{j\leftarrow\ell}(s_\ell)\prod_{\varphi\in\boldsymbol{\varphi}_k\backslash\tau}\prod_{r\in\partial_\varphi(k)}\pi_{k\leftarrow r}(s_r)\right] \\ \times\,\delta\left(s_\tau - 2, \sum_{\nu\in\boldsymbol{\nu}_j\backslash\tau}\sum_{\ell\in\partial_\nu(j)} s_\ell + \sum_{\varphi\in\boldsymbol{\varphi}_k\backslash\tau}\sum_{r\in\partial_\varphi(k)} s_r\right). \tag{25}$$

The Kronecker delta evaluates as follows

$$\sum_{s_\tau=1}^{\infty} z^{s_\tau} \overset{\delta}{=} z^2 \sum_{s=1}^{\infty} z^{\sum\limits_{\nu\in\boldsymbol{\nu}_j\backslash\tau}\sum\limits_{\ell\in\partial_\nu(j)} s_\ell}\, z^{\sum\limits_{\varphi\in\boldsymbol{\varphi}_k\backslash\tau}\sum\limits_{r\in\partial_\varphi(k)} s_r}, \tag{26}$$

$$= z^2 \prod_{\nu\in\boldsymbol{\nu}_j\backslash\tau}\prod_{\ell\in\partial_\nu(j)}\sum_{s_\ell=0}^{\infty}\prod_{\varphi\in\boldsymbol{\varphi}_k\backslash\tau}\prod_{r\in\partial_\varphi(k)}\sum_{s_r=0}^{\infty} z^{s_\ell+s_r}. \tag{27}$$

We find

$$\sum_{s_\tau=1}^{\infty} z^{s_\tau}\pi_{i\leftarrow\tau}\left(s_\tau \mid \triangledown, \triangledown, \triangledown, \triangledown\right) = \left(z\prod_{\nu\in\boldsymbol{\nu}_j\backslash\tau}\prod_{\ell\in\partial_\nu(j)}\sum_{s_\ell=0}^{\infty}\pi_{j\leftarrow\ell}(s_\ell)z^{s_\ell}\right)\left(z\prod_{\varphi\in\boldsymbol{\varphi}_k\backslash\tau}\prod_{r\in\partial_\varphi(k)}\sum_{s_r=0}^{\infty}\pi_{k\leftarrow r}(s_r)z^{s_r}\right), \tag{28}$$

$$= \left(z\prod_{\nu\in\boldsymbol{\nu}_j\backslash\tau} H_{j\leftarrow\nu}(z)\right)\left(z\prod_{\varphi\in\boldsymbol{\varphi}_k\backslash\tau} H_{k\leftarrow\varphi}(z)\right). \tag{29}$$

Inserting this result into Eq 24 together with the other expressions in Eqs 21b and 21c we have

$$\begin{aligned}H_{i\leftarrow\tau}(z) = (1-\phi)^2 + \phi(1-\phi)^2\left(z\prod_{\nu\in\boldsymbol{\nu}_j\backslash\tau} H_{j\leftarrow\nu}(z) + z\prod_{\varphi\in\boldsymbol{\varphi}_k\backslash\tau} H_{k\leftarrow\varphi}(z)\right)\\ + \left(3\phi^2(1-\phi)+\phi^3\right)\left(z\prod_{\nu\in\boldsymbol{\nu}_j\backslash\tau} H_{j\leftarrow\nu}(z)\right)\left(z\prod_{\varphi\in\boldsymbol{\varphi}_k\backslash\tau} H_{k\leftarrow\varphi}(z)\right).\end{aligned} \tag{30}$$

The result is a polynomial in powers of generating functions $H_{\ell\leftarrow\sigma}(z)$ where $\ell$ is a vertex in motif $\tau$ and $\sigma$ is a motif that $\ell$ belongs to, other than motif $\tau$.

### C. Calculating $H_{i\leftarrow\tau}$ for chordless cycles of size $n\geq 3$

Perhaps the simplest family of subgraphs to consider is the set of chordless cycles of increasing vertex count. In this section we will derive a closed form expression for $H_{i\leftarrow\tau}$ where $\tau$ is a cycle of length $n\geq 3$. As with the triangle in section II B we choose a vertex in the cycle $i$ and calculate the probability that $s_\tau$ other vertices can be reached due to membership in motif $\tau$. We achieve this by writing the conditional probabilities that the cycle has a given bond configuration; since, each bond configuration occurs with a different probability. Let all edges of the cycle be occupied and so vertex $i$ belongs to the same component as the other $n-1$ vertices. Without loss of generality, let us also label the vertices from 1 to $n$ and set $i=1$. In this case we can write

$$\pi_{i\leftarrow\tau}(s_\tau\mid\mathcal{C}_n) = \sum_{\{s_\ell:\ell\in\partial_\nu(j),\ \nu\in\boldsymbol{\nu}_j\backslash\tau,\ j\in\tau\backslash i\}}\left[\prod_{j\in\tau\backslash i}\prod_{\nu\in\boldsymbol{\nu}_j\backslash\tau}\prod_{\ell\in\partial_\nu(j)}\pi_{j\leftarrow\ell}(s_\ell)\right]\delta\left(s_\tau-n+1,\sum_{j\in\tau\backslash i}\sum_{\nu\in\boldsymbol{\nu}_j\backslash\tau}\sum_{\ell\in\partial_\nu(j)} s_\ell\right). \tag{31}$$

The set $\{\ell\in\partial_\nu(j),\ \nu\in\boldsymbol{\nu}_j\backslash\tau,\ j\in\tau\backslash i\}$ accounts (from left to right) for all edges $\ell$ in motif $\nu$ that vertex $j$ belongs to, for all motifs $\nu$ that vertex $j$ belongs to, apart from $\tau$, for all vertices $j$ in motif $\tau$ apart from $i$. The Kronecker delta filters those terms that do not sum to $s_\tau-n+1$, accounting for the vertices that belong to the $n$-cycle other than $i$. If one of the edges around the cycle was set to the unoccupied state a path between the vertices through the cycle would still be present. Subsequent removal of an edge from the cycle has the potential to isolate a vertex and so, all of the states with a single edge removed have been exhausted. We can write the coefficient of this term as

$$P(\mathcal{C}_n) = \phi^n + n\phi^{n-1}(1-\phi). \tag{32}$$

We now examine the configuration $\mathcal{C}_{n-1}$ where a single neighbour vertex $k$ within the cycle belongs to a different component to vertex $i$. For this to occur, both of the edges that connect this vertex to the cycle must be unoccupied with probability $(1-\phi)^2$ and so the probability of this edge configuration is

$$P(\mathcal{C}_{n-1}) = \phi^{n-1-1}(1-\phi)^2. \tag{33}$$

To exclude $k$ from the set of vertices in motif $\tau$ that contribute to the summation in Eq 31, we simply modify the set notation to $\mathcal{S}(k)=\{\ell\in\partial_\nu(j),\ \nu\in\boldsymbol{\nu}_j\backslash\tau,\ j\in\tau\backslash\{i,k\}\}$. However, given that the identity of $k$ can be any neighbour in $\tau$ apart from $i$, we must account for all possible identities by summing $k$ from 2 to $n$ to obtain

$$\pi_{i\leftarrow\tau}(s_\tau\mid\mathcal{C}_{n-1}) = \sum_k\sum_{\mathcal{S}(k)}\left[\prod_{j\in\tau\backslash\{i,k\}}\prod_{\nu\in\boldsymbol{\nu}_j\backslash\tau}\prod_{\ell\in\partial_\nu(j)}\pi_{j\leftarrow\ell}(s_\ell)\right]\delta\left(s_\tau-n+2,\sum_{j\in\tau\backslash\{i,k\}}\sum_{\nu\in\boldsymbol{\nu}_j\backslash\tau}\sum_{\ell\in\partial_\nu(j)} s_\ell\right). \tag{34}$$

No edges can be removed from this motif without further isolation of a vertex, indicating that this edge configuration has been fully calculated. Considering the next term $\mathcal{C}_{n-2}$, where the removal of another edge isolates a second vertex. Vertex $i$ now belongs to an induced subgraph of $n-2$ vertices in $\tau$. We cannot set an edge of our choice to be unoccupied to arrive at this state, however. It happens that the removed edge *must* be one of the edges that a neighbour of $k$ has within the component of vertex $i$. If we had chosen a different edge, we would not account for all combinations for connected components of this length as we would have isolated vertices prematurely from the chain of removed vertices. We can write

$$\pi_{i\leftarrow\tau}(s_\tau\mid\mathcal{C}_{n-2}) = \sum_{k=2}^{n-1}\sum_{\mathcal{S}(k,\tau_{k+1})}\left[\prod_{j\in\tau\backslash\{i,k,\tau_{k+1}\}}\prod_{\nu\in\boldsymbol{\nu}_j\backslash\tau}\prod_{\ell\in\partial_\nu(j)}\pi_{j\leftarrow\ell}(s_\ell)\right]\delta\left(s_\tau-n+3,\sum_{j\in\tau\backslash\{i,k,\tau_{k+1}\}}\sum_{\nu\in\boldsymbol{\nu}_j\backslash\tau}\sum_{\ell\in\partial_\nu(j)} s_\ell\right), \tag{35}$$

where $\mathcal{S}(k, \tau_{k+1}) = \{\ell \in \partial_\nu(j),\ \nu \in \boldsymbol{\nu}_j \backslash \tau,\ j \in \tau \backslash \{i, k, \tau_{k+1}\}\}$. To picture this, we are essentially moving a pair of connected vertices around the motif as a sliding window. The probability of this edge configuration is

$$P(\mathcal{C}_{n-2}) = \phi^{n-1-2}(1-\phi)^2. \tag{36}$$

The logic behind this expression is that the chain of removed vertices must fail to connect to the same component that vertex $i$ belongs to, which occurs with probability $(1-\phi)^2$. Otherwise, all edges between the vertices that are connected to $i$ are occupied with probability $\phi$, of which, there are $n-1-2$.

Generalising this process, the conditional probability that a chordless cycle with $n$ vertices, of which $r$ belong to a different component to vertex $i = 1$, leads to $s_\tau$ reachable vertices is

$$\begin{aligned}\pi_{i\leftarrow\tau}(s_\tau \mid \mathcal{C}_{n-r})P(\mathcal{C}_{n-r}) = \phi^{n-1-r}(1-\phi)^2 \sum_{k=2}^{n-r+1} \sum_{\mathcal{S}([k,\tau_{k+r-1}])} \left[ \prod_{j\in\tau\backslash\{i,[k,\tau_{k+r-1}]\}} \prod_{\nu\in\boldsymbol{\nu}_j\backslash\tau} \prod_{\ell\in\partial_\nu(j)} \pi_{j\leftarrow\ell}(s_\ell) \right] \\ \times \delta\left( s_\tau - (n-1-r), \sum_{j\in\tau\backslash\{i,[k,\tau_{k+r-1}]\}} \sum_{\nu\in\boldsymbol{\nu}_j\backslash\tau} \sum_{\ell\in\partial_\nu(j)} s_\ell \right),\end{aligned} \tag{37}$$

where there are $n-1-r$ vertices in $\tau$ that are connected to $i$ and with the notational understanding that all vertices in the range $k$ to $\tau_{k+r-1}$ (inclusive) are excluded from the set in addition to vertex $i$ such that

$$\mathcal{S}([k, \tau_{k+r-1}]) = \{\ell \in \partial_\nu(j),\ \nu \in \boldsymbol{\nu}_j \backslash \tau,\ j \in \tau \backslash \{i, k, \tau_{k+1}, \dots, \tau_{k+r-2}, \tau_{k+r-1}\}\}, \tag{38}$$

and similarly for the summation in the Kronecker delta

$$j \in \tau \backslash \{i, [k, \tau_{k+r-1}]\} = j \in \tau \backslash \{i, k, \tau_{k+1}, \dots, \tau_{k+r-2}, \tau_{k+r-1}\}. \tag{39}$$

We can then sum Eq 37 from $r = 1$ to $n-1$ to account for chains of isolated vertices for all lengths. Together with the expression for the fully connected cycle we finalise the expression

$$\pi_{i\leftarrow\tau}(s_\tau) = \pi_{i\leftarrow\tau}(s_\tau \mid \mathcal{C}_n)P(\mathcal{C}_n) + \sum_{r=1}^{n-1} \pi_{i\leftarrow\tau}(s_\tau \mid \mathcal{C}_{n-r})P(\mathcal{C}_{n-r}). \tag{40}$$

This expression is then generated as

$$\begin{aligned}H_{i\leftarrow\tau}(z) = \left(\phi^n + n\phi^{n-1}(1-\phi)\right) \prod_{j\in\tau\backslash i} \left( z \prod_{\nu\in\boldsymbol{\nu}_j\backslash\tau} H_{j\leftarrow\nu}(z) \right) \\ + \sum_{r=1}^{n-1} \phi^{n-1-r}(1-\phi)^2 \sum_{k=2}^{n-r+1} \prod_{j\in\tau\backslash\{i,[k,\tau_{k+r-1}]\}} \left( z \prod_{\nu\in\boldsymbol{\nu}_j\backslash\tau} H_{j\leftarrow\nu}(z) \right).\end{aligned} \tag{41}$$

The first term in Eq 41 accounts for the edge configuration when all vertices in the $n$-cycle are connected. The second term in this expression accounts for the various edge configurations that the cycle may exhibit when a chain of $r$ vertices are removed from the cycle due to bond percolation. From right to left the indices of the second term account for: for all motifs $\nu$ (apart from $\tau$) that vertex $j$ belongs to, for all $j$ in the chain of vertices that are connected to $i$ (which we enumerate by removing vertices $\{i, [k, \tau_{k+r-1}]\}$ from the set of vertices in $\tau$), for all possible starting points $k$ of the $r$-chain, for all possible lengths $r$. This expression is the main result of this section; as a consistency check, we find Eq 30 for triangles $n = 3$ can be recovered.

## III. CALCULATING $H_{i\leftarrow\tau}$ FOR CLIQUES

We will now derive an exact expression for belief propagation on clique motifs, extending the ensemble expressions derived by Mann *et al* [29]. The generating function for the probability that a focal vertex $i$, that is member of a clique $\tau$, can reach $s_\tau$ vertices through its membership in $\tau$ is $H_{i\leftarrow\tau}(z)$. Following bond percolation, each edge in $\tau$ can be occupied with probability $\phi$ or unoccupied with probability $1-\phi$. Depending on the bond configuration of the clique, vertex $i$ might not be connected to all neighbours via occupied edges. To calculate $H_{i\leftarrow\tau}(z)$ we must average over each possible bond configuration that the focal vertex might observe and the various states of connectivity among the neighbours and $i$. The expression takes the form of a polynomial whose terms are the powerset, denoted by $2^{\{j\in\tau\backslash i\}}$, of the set of vertices in the clique apart from $i$, including the empty set $\{\varnothing\}$. For instance, consider the form of the expression when $\tau$ is a 4-clique, with vertices labelled 0 to 3 and with vertex $i$ arbitrarily being labelled as 0, we have

$$H_{0\leftarrow\tau}(z) = \mathcal{P}\left(\varnothing\right) + \mathcal{P}\left(1\right) + \mathcal{P}\left(2\right) + \mathcal{P}\left(3\right) + \mathcal{P}\left(1,2\right) + \mathcal{P}\left(1,3\right) + \mathcal{P}\left(2,3\right) + \mathcal{P}\left(1,2,3\right). \tag{42}$$

Each term $\mathcal{P}(j,\dots,k)$ generates the probability that membership in $\tau$ leads to $s_\tau$ reachable vertices, given $i$ is connected to vertices $j,\dots,k$ (a path of occupied edges exits within the motif). Each term can be decomposed into the probability that the neighbours collectively lead to $s_\tau$ vertices (including themselves), multiplied by the probability that the bond percolation process yielded that edge configuration $P(j,\dots,k)$

$$\mathcal{P}(j,\dots,k) = \sum_{s_\tau=1}^{\infty} z^{s_\tau}\pi_{i\leftarrow\tau}(s_\tau \mid j,\dots,k)P(j,\dots,k). \tag{43}$$

It happens, that because of the symmetry of a clique, $P(j,\dots,k)$ only depends on the number of vertices $|j,\dots,k|$, not their identity [35] Given this logic, we group terms that have equal numbers of connected vertices in $\tau$ and the polynomial in Eq 42 becomes a summation over the subsets $a_\kappa \in A_\kappa$ of the set of neighbours $\{j\in\tau\backslash i\}$ of a given length $\kappa$

$$H_{i\leftarrow\tau}(z) = \sum_{\kappa=0}^{|\tau|-1} P(\kappa)\sum_{a_\kappa\in A_\kappa}\sum_{s_\tau=0}^{\infty} z^{s_\tau}\pi_{i\leftarrow\tau}(s_\tau \mid a_\kappa), \tag{44}$$

Concentrating on the probability $\pi_{i\leftarrow\tau}(s_\tau \mid a_\kappa)$ first, we can expand this as a sum over all sizes $s_\ell$

$$\pi_{i\leftarrow\tau}(s_\tau \mid a_\kappa) = \sum_{\{s_\ell:\ell\in\partial_\nu(j),\ \nu\in\boldsymbol{\nu}_j\backslash\tau,\ j\in a_\kappa\}}\left[\prod_{j\in a_\kappa}\prod_{\nu\in\boldsymbol{\nu}_j\backslash\tau}\prod_{\ell\in\partial_\nu(j)}\pi_{j\leftarrow\ell}(s_\ell)\right]\delta\left(s_\tau-\kappa-1,\sum_{j\in a_\kappa}\sum_{\nu\in\boldsymbol{\nu}_j\backslash\tau}\sum_{\ell\in\partial_\nu(j)} s_\ell\right). \tag{45}$$

where $\ell$ is an index over the edges of the neighbours that $i$ is connected to in $\tau$, excluding those that point back to $\tau$. Generating this expression we obtain

$$\sum_{s_\tau=0}^{\infty} z^{s_\tau}\pi_{i\leftarrow\tau}(s_\tau \mid a_\kappa) = \prod_{j\in a_\kappa}\left(z\prod_{\nu\in\boldsymbol{\nu}_j\backslash\tau} H_{j\leftarrow\nu}(z)\right). \tag{46}$$

The next step is to find the coefficient $P(\kappa)$ for each size $\kappa$ component of neighbours within $\tau$ that can attach to $i$, see [29]. Within this calculation, we must account for all possible connected graphs that can occur among $\kappa+1$ vertices in $\tau$. For a given $\kappa$ there are $\frac{1}{2}(\kappa+1)\kappa$ total edges among vertices in the connected component. Letting $r = |\tau|-\kappa-1$, there are $\frac{1}{2}r(r-1)$ many edges between the removed vertices (ones that are not in the component with $i$). Finally, there are

$$\omega(r) = \sum_{v=1}^{r}(|\tau|-v) - \frac{1}{2}r(r-1), \tag{47}$$

edges that connect vertices that belong to the same component as $i$ to those that don't (they interface the two components).

For instance, consider a $|\tau| = 6$ clique that has $\kappa = 3$ neighbours that are connected to focal vertex $i$ and therefore $r = 2$ vertices that do not have a path of occupied edges to $i$ as depicted in Fig 4. There are $\frac{1}{2}(3+1)3 = 6$ edges among the vertices in the same component as $i$ and $\omega(2) = 8$ interface edges. The bond occupancy probability of this configuration is given by $\phi^{\frac{1}{2}(\kappa+1)\kappa}(1-\phi)^{\omega(r)}$. However, we can remove up to $\frac{1}{2}\kappa(\kappa-1)$ of the $\frac{1}{2}(\kappa+1)\kappa$ occupied edges from the component that $i$ belongs to and still retain connectivity among the $\kappa+1$ vertices. Therefore, indexing the number of removed edges by $m$ we have

$$P(\kappa) = \sum_{m=0}^{\frac{1}{2}\kappa(\kappa-1)} \mathcal{Q}_{\kappa+1,\frac{1}{2}(\kappa+1)\kappa-m}\phi^{\frac{1}{2}(\kappa+1)\kappa-m}(1-\phi)^{\omega(r)+m}, \tag{48}$$

where $\mathcal{Q}_{n,k}$ is the number of connected graphs that can be made that have $n$ vertices and $k$ edges, see Appendix A. Inserting Eqs 46 and 48 into Eq 44 we arrive at the main result of this section

$$H_{i\leftarrow\tau}(z) = \sum_{\kappa=0}^{|\tau|-1}\sum_{m=0}^{\frac{1}{2}\kappa(\kappa-1)} \mathcal{Q}_{\kappa+1,\frac{1}{2}(\kappa+1)\kappa-m}\phi^{\frac{1}{2}(\kappa+1)\kappa-m}(1-\phi)^{\omega(r)+m}\sum_{a_\kappa\in A_\kappa}\prod_{j\in a_\kappa}\left(z\prod_{\nu\in\boldsymbol{\nu}_j\backslash\tau} H_{j\leftarrow\nu}(z)\right), \tag{49}$$

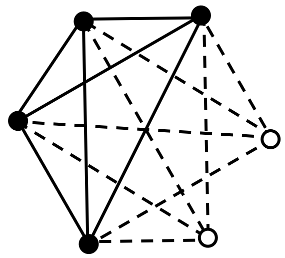


FIG. 4. A bond occupancy configuration of a 6-clique in which two vertices do not belong to the giant component (unfilled) whilst the remaining 4 vertices do. Solid edges are occupied whilst dashed edges are unoccupied. The occupation state of the edge linking the two unfilled vertices is inconsequential to the percolation properties of the 4 vertices in the giant component. There are $0.5(3+1)3 = 6$ occupied edges among the filled vertices, of which 3 can be set to unoccupied and connectivity retained. There are $\omega(2) = 8$ edges that must be unoccupied if the two unfilled vertices are to remain outside of the giant component.

where $r = |\tau| - \kappa - 1$. Unpacking this expression when $\tau$ is a 3-clique reproduces Eq 30. When $\tau$ is a 4-clique we obtain

$$
\begin{aligned}
H_{i\leftarrow\tau}(z) = {} & (1-\phi)^3 + z\phi(1-\phi)^4\left[\prod_{\nu\in\boldsymbol{\nu}_{\tau_1}\setminus\tau} H_{\tau_1\leftarrow\nu}(z) + \prod_{\nu\in\boldsymbol{\nu}_{\tau_2}\setminus\tau} H_{\tau_2\leftarrow\nu}(z) + \prod_{\nu\in\boldsymbol{\nu}_{\tau_3}\setminus\tau} H_{\tau_3\leftarrow\nu}(z)\right] \\
& + z^2\left[\phi^3(1-\phi)^3 + 3\phi^2(1-\phi)^4\right]\left[\prod_{\nu\in\boldsymbol{\nu}_{\tau_1}\setminus\tau} H_{\tau_1\leftarrow\nu}(z)\prod_{\nu\in\boldsymbol{\nu}_{\tau_2}\setminus\tau} H_{\tau_2\leftarrow\nu}(z) + \prod_{\nu\in\boldsymbol{\nu}_{\tau_1}\setminus\tau} H_{\tau_1\leftarrow\nu}(z)\prod_{\nu\in\boldsymbol{\nu}_{\tau_3}\setminus\tau} H_{\tau_3\leftarrow\nu}(z)\right. \\
& \left. + \prod_{\nu\in\boldsymbol{\nu}_{\tau_2}\setminus\tau} H_{\tau_2\leftarrow\nu}(z)\prod_{\nu\in\boldsymbol{\nu}_{\tau_3}\setminus\tau} H_{\tau_3\leftarrow\nu}(z)\right] \\
& + z^3[\phi^6 + 6\phi^5(1-\phi) + 15\phi^4(1-\phi)^2 + 16\phi^3(1-\phi)^3]\prod_{\nu\in\boldsymbol{\nu}_{\tau_1}\setminus\tau} H_{\tau_1\leftarrow\nu}(z)\prod_{\nu\in\boldsymbol{\nu}_{\tau_2}\setminus\tau} H_{\tau_2\leftarrow\nu}(z)\prod_{\nu\in\boldsymbol{\nu}_{\tau_3}\setminus\tau} H_{\tau_3\leftarrow\nu}(z).
\end{aligned}
\tag{50}
$$

The coefficients of larger cliques are in agreement with the exact expressions previously found by [36] and [29]. The derivative of Eq 49 that is required for the calculation of the finite components is given by

$$
H'_{i\leftarrow\tau}(z) = \sum_{\kappa=0}^{|\tau|-1}\sum_{m=0}^{\frac{1}{2}\kappa(\kappa-1)} \mathcal{Q}_{\kappa+1,\frac{1}{2}(\kappa+1)\kappa-m}\,\phi^{\frac{1}{2}(\kappa+1)\kappa-m}(1-\phi)^{\omega(r)+m}\sum_{a_\kappa\in A_\kappa}\frac{d}{dz}\left[\prod_{j\in a_\kappa}\left(z\prod_{\nu\in\boldsymbol{\nu}_j\setminus\tau} H_{j\leftarrow\nu}(z)\right)\right],
\tag{51}
$$

where we can apply Eq 14 to find

$$
\begin{aligned}
\frac{d}{dz}\left[\prod_j\left(z\prod_\nu H_{j\leftarrow\nu}(z)\right)\right] = {} & \prod_j\left(z\prod_\nu H_{j\leftarrow\nu}(z)\right)\left[\sum_j\left(z\prod_\nu H_{j\leftarrow\nu}(z)\right)^{-1}\right. \\
& \left.\times\left[\prod_\nu H_{j\leftarrow\nu}(z) + z\left(\prod_\nu H_{j\leftarrow\nu}(z)\right)\left(\sum_\nu\frac{H'_{j\leftarrow\nu}(z)}{H_{j\leftarrow\nu}(z)}\right)\right]\right].
\end{aligned}
\tag{52}
$$

The fixed point of these expressions can be found to yield the derivatives and Eq 13 can be solved.

## IV. RELATION TO THE GENERALISED CONFIGURATION MODEL

The message passing formulation calculates the properties of a given graph realisation $G$. Often, the properties of an *ensemble* of networks $G \in \mathcal{G}$ with equivalent statistics are required, rather than a single instance. An ensemble of random networks can be created according to the generalised configuration model algorithm [24–26, 30, 31]. In this model, a set of motifs are defined and each vertex is assigned a tuple of integers, called its *joint degree* that represents the number of edge-disjoint motifs of a given topology that it belongs to. For instance, if a vertex belongs to three ordinary edges, one triangle, one 4-cycle and two 5-cycles the joint degree is $(k_2, k_3, k_4, k_5) = (3, 1, 1, 2)$. The distribution of joint degrees are fixed by the joint degree distribution $p(k_2, k_3, k_4, k_5)$ which is the fraction of vertices with a given joint degree.

During the construction process, the configuration model randomises the identity of the vertices that belong to a given motif. Therefore, for a given vertex $i$, the joint degree of its neighbours can vary drastically and so the neighbourhood beyond the nearest neighbours becomes a mean field quantity. To see this, consider an edge-disjoint triangle cover of a network; ordinary edges being present also. All of the edges in the graph have been assigned to a *motif topology* by the cover that relates to the type of motif to which they belong. Let us now break each edge into two parts whilst retaining the topology labels; isolating each vertex. To create the required number of ordinary edges and triangles, the configuration model selects vertices at random and connects their stubs together, matching the topologies. For instance, to create a triangle, three vertices are selected at random that have free unmatched triangle stubs and are then connected together appropriately. In this way, each random graph that is constructed by the configuration model belongs to an equivalent set of motif topologies, but the identity of the vertices that comprise those motifs will be different. Representing this stochasticity by an average over all graphs in the ensemble, the product of the average message that vertex $j \in \tau$ receives from each of its other motifs $\nu \in \boldsymbol{\nu}_j \backslash \tau$ can be written as an average over the product of messages

$$\prod_{\nu \in \boldsymbol{\nu}_j \backslash \tau} \langle H_\nu(z) \rangle \le \left\langle z \prod_{\nu \in \boldsymbol{\nu}_j \backslash \tau} H_{j \leftarrow \nu}(z) \right\rangle, \tag{53}$$

where we have used a version of the Chebyshev integral inequality [34] for $k$ monotonic functions of the same monotonicity

$$\prod_{i=1}^{n} \langle f_i(x_1, \dots, x_k) \rangle \le \left\langle \prod_{i=1}^{n} f_i(x_1, \dots, d_k) \right\rangle. \tag{54}$$

In general, the average of a product is not equal to the product of the average. Only when the covariance between the messages is zero is this true. For any two motifs to be correlated in the cavity graph, there must be loops in the factor graph. However, in the limit that the number of vertices $N$ becomes infinite, the shortest short range cycles are expected to be at least length $\mathcal{O}(\log N)$. Assuming that the factor graph is locally treelike, then the messages that arrive at a cavity from independent motifs can be treated as though they are uncorrelated with one another; in this case Eq 53 becomes an equality.

As an example, let us partition the product over the motifs of the neighbour into a product over its motif topologies. For instance, for ordinary edges and triangles we have

$$\prod_{s \in \boldsymbol{s}_j \backslash \tau} \langle H_\perp \rangle \prod_{t \in \boldsymbol{t}_j} \langle H_\Delta \rangle = \left\langle z \prod_{s \in \boldsymbol{s}_j \backslash \tau} H_{j \leftarrow s} \prod_{t \in \boldsymbol{t}_j} H_{j \leftarrow t} \right\rangle, \tag{55}$$

and

$$\prod_{s \in \boldsymbol{s}_j} \langle H_\perp \rangle \prod_{t \in \boldsymbol{t}_j \backslash \tau} \langle H_\Delta \rangle = \prod_{j \in t \backslash i} \left\langle z \prod_{s \in \boldsymbol{s}_j} H_{j \leftarrow s} \prod_{t \in \boldsymbol{t}_j \backslash t} H_{j \leftarrow t} \right\rangle, \tag{56}$$

where $\boldsymbol{s}_j$ $(\boldsymbol{t}_j)$ is the set of ordinary edges (triangles) to which vertex $j$ belongs. There are two expressions since motif $\tau$ could have been an ordinary edge or a triangle and the probabilities associated with each one are not equivalent in general. However, since the average message is the same for each motif of a given topology, we can simplify this as a power

$$\prod_{s \in \boldsymbol{s}_j \backslash \tau} \langle H_\perp \rangle \prod_{t \in \boldsymbol{t}_j} \langle H_\Delta \rangle = \langle H_\perp \rangle^{k_2} \langle H_\Delta \rangle^{k_3}, \tag{57}$$

where $k_2$ is the excess number of ordinary edges that $j$ has (given that $\tau$ is an ordinary edge) and $k_3$ is the number of triangles $j$ belongs to. In other words, $k_2 = \mathrm{card}(\boldsymbol{s}_j \backslash \tau)$ and $k_3 = \mathrm{card}(\boldsymbol{t}_j)$. Similarly, if $\tau$ had instead been a triangle we can write

$$\prod_{s \in \boldsymbol{s}_j} \langle H_{\perp,s} \rangle \prod_{t \in \boldsymbol{t}_j \backslash \tau} \langle H_{\perp,t} \rangle = \langle H_\perp \rangle^{k_2} \langle H_\Delta \rangle^{k_3}, \tag{58}$$

where $k_2 = \mathrm{card}(\boldsymbol{s}_j)$ is the number of ordinary edges and $k_3 = \mathrm{card}(\boldsymbol{t}_j \backslash \tau)$ is the excess triangle degree of vertex $j$, respectively. This expression is the product of the average messages along ordinary edges and triangles. The excess degree distributions are distributed as $q_\tau(k_2, k_3)$ and so averaging over the distribution we can write a generating function for the message that a neighbour receives in the random graph ensemble as

$$G_{1,\tau}(z) = \sum_{k_2} \sum_{k_3} q_\tau \langle H_\perp(z) \rangle^{k_2} \langle H_\Delta(z) \rangle^{k_3}. \tag{59}$$

An equivalent argument can be followed to define another fundamental generating function $G_0(x, y) = \sum_{k_2} \sum_{k_3} p(k_2, k_3) x^{k_2} y^{k_3}$ from Eq 10. This logic can be extended to all motifs that are included in the model. With these two generating functions defined, the mapping between the message passing formulation and the generalised configuration model is complete.

## V. DISCUSSION

The exact expressions we have derived work well for random graph models, such as the generalised configuration model [26, 30, 31]. In Fig 5 we apply these equations to a random graph model comprising 2-, 3- and 4-cliques that has been constructed according to the GCM algorithm; observing excellent agreement between Monte Carlo bond percolation and our theoretical model. In or-

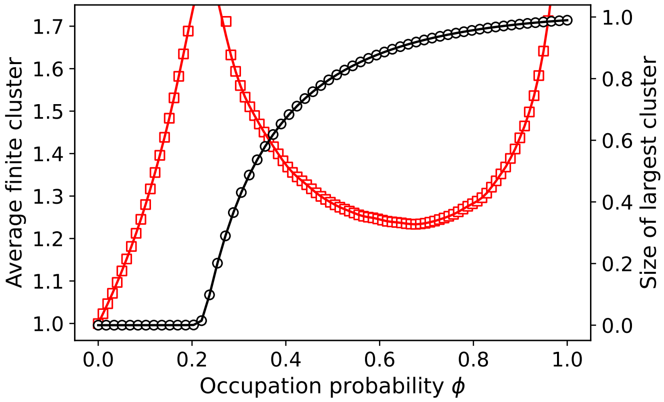


FIG. 5. A random GCM graph with 2- 3- and 4-cliques. Solid lines are the results of the message passing calculations from Eq 11 with Eq 49 and Eq 13 with Eq 51. The scatter points are the average of Monte Carlo simulations of bond percolation on the same network. Circles are the average size of the largest cluster; whilst squares are the average finite component sizes.

der to apply our expressions to the study of real world networks, we must find a way to cover the network in motifs. In what amounts to community detection [37], there are perhaps a near infinite number of ways to cover a network with motifs; and some will inherently lead to better models of the empirical network than others [28, 32, 38]. The problem of network dismantling is closely related [39–41].

A trivial solution is to simply cover the network in 2-cliques. In doing this, we are assuming that the network is treelike; this would lead to the traditional belief propagation model, which is known to suffer from statistical errors. To progress beyond this model, the simplest loop we might add to a cover is a triangle; and beyond this chordless cycles and cliques of all orders. However, restricting the permissible motifs to cliques and cycles is unlikely to lead to a locally treelike factor graph. At the other extreme of this logic, one might try to define motifs that contain $\mathcal{O}(N)$ vertices; perhaps the largest Eulerian cycle for instance. Whilst this cover might lead to a treelike factor graph, we could not hope to write the message passing equations in closed form. Therefore, we must find a balance between motifs that are arbitrary enough to make the factor graph *sufficiently* treelike, yet are small enough to be analytically tractable such that their message passing equations can be calculated in reasonable time.

Let us suppose that we have defined a set of motifs, perhaps cliques of all sizes. We now have to decide a strategy of how to place the motifs on the network. By far the most simple strategy is to simply search for the presence of a pre-defined set of higher-order motifs and greedily add them to the network cover. This approach is usually stochastic if different starting locations (and subsequent search patterns) are chosen. Burgio *et al* [32] suggest that a *maximal* cover is the best strategy to retain as much higher-order structure of the empirical networks as possible. In their heuristic, the cover is chosen such that the fewest cliques possible are placed on the network; minimising the local disruption caused by placing a motif. Mann *et al* [28] showed that preferentially including the *largest* cliques can be beneficial for retaining the degree correlations among the more well-connected vertices; although this approach is often not maximal. Both of these methods can be generalised to placing motifs other than cliques.

Finally, we would like to highlight another concern that should be accounted for when choosing the motif cover - symmetry. Consider the neighbourhood of a vertex that has been covered by 2- and 3-cliques in Fig 6 a). Due to the edge-disjoint property of the cover, only half of the edges between the neighbours are included in triangles (and therefore, motifs surrounding the focal vertex). When calculating the message passing equations, the focal vertex only observes the effective neighbourhood depicted in Fig 6 b). Because of the asymmetry in how the edges between 1st-order neighbours are accounted for, this cover introduces a statistical bias due to the preferential inclusion of only *some* of the edges between the 1st-order neighbours and not others. To remove the bias we must either choose the 2-clique cover or, choose the entire motif to be part of the cover. This means that clique covers, which almost always involve breaking symmetry between the neighbours, are inferior to covering methods that account for the entire neighbourhood of a vertex [3]. Despite these drawbacks, our model gives good results for

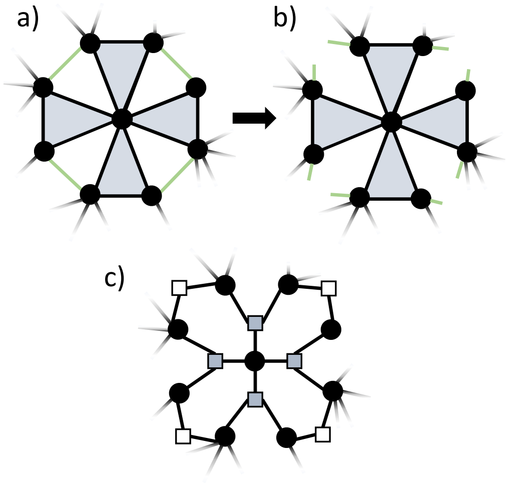


FIG. 6. a) The neighbourhood of a focal vertex $i$ covered with ordinary edges (green) and triangles (shaded). Only half of the connections between the neighbours are included in the neighbourhood of $i$ due to the edge-disjoint property of the cover. b) When we calculate the cavity equations $H_{i\leftarrow\tau(z)}$ for $i$, the ordinary edges are included in the product over the motifs that vertex $j \in \tau\backslash i$ belongs to $\prod_{\nu} H_{j\leftarrow\nu}(z)$; introducing bias and therefore statistical error. c) The factor graph contains short range loops.

real world networks, see Fig 7. In Fig 7 a), we show the size of the largest connected component for a social network of coauthorship relations between 13,861 scientists [42] in the field of condensed-matter physics; and b) a network of 10,680 users of the PGP encryption software [43]. These networks are known to contain a high-density of short loops; and therefore, the standard message passing equations, based on a 2-clique cover (dashed green), fail to correctly predict the size of the percolating cluster. To cover the networks in cliques, we use the so-called *motif preserving clique cover* (MPCC) [28]. This heuristic attempts to include the largest cliques in the edge-disjoint cover in order to preserve as many contacts among the vertices within a given motif. We find a significant improvement over the 2-clique cover, therefore highlighting the importance of retaining higher-order structure for the message passing theory. These networks were previously studied by Cantwell and Newman [3] and have been chosen for comparison of the different theoretical methods. The neighbourhood method of [3] offers closer agreement to the Monte Carlo simulation than the motif cover approach. This is due to the edge-disjoint conditions of the covering method. However, our model offers a vast improvement over the traditional message passing theory. Our model, by definition, works on a representation of the underlying empirical network in the form of a collection of motifs. This model can then be used to garner a wide range of information regarding the substrate network including its relation to the ensemble of joint-degree equivalent graphs. In turn, this deepens our understanding of precisely *which* connective motifs are important to a network to yield its properties. Cliques are particularly favourable due to the wide range of algorithms and exact theoretical results that have been developed for these motifs, including those of section III as well as other recent work in this area [27, 28, 32]. We comment that finding the clique cover and solving the message passing equations for these networks takes just over 1 minute on 16 GB Apple M1 Silicone, which is very fast compared to other techniques.

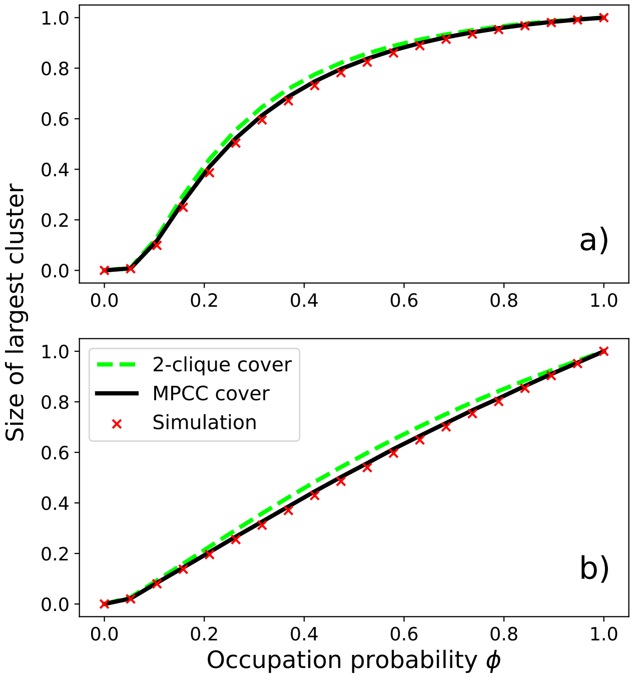


FIG. 7. The size of the largest cluster of: a) a coauthorship network of 13,861 scientists [42] and (b) a network of 10,680 users of the PGP encryption software [43]. The scatter points are the average of Monte Carlo simulation of bond percolation. The dashed (green) line is the result of the message passing equations with the 2-clique cover; whilst the solid (black) line is the result for the MPCC cover and Eq 49.

## VI. CONCLUSION

Belief propagation on loopy networks is a topic that has received much attention in the literature. In this paper, we have studied a bond percolation model on networks that contain simple cycles and cliques by deriving a message passing model. We assume that the factor graph of these networks is locally treelike in order to reduce the correlations between the messages each vertex observes. We examined our theoretical framework for random graphs constructed according to the generalised configuration model as well as a covered empirical networks. We found excellent agreement between Monte Carlo simulation of the average cluster size distribution and the expected size of the percolating component and our equations. Our model offers significant advantage over the traditional belief propagation framework when applied to empirical networks that contain a non-trivial density of loops.

The choice of motif cover is influential to the success of the model and we highlighted some considerations around this. However, it is non-trivial to derive the conditional message probabilities for a given edge configuration of each motif in a closed form expression. Further investigation of the impact of cover choice should be conducted including: the choice of permissible motifs in the model, different cover rules for mesoscale network structures, placing large or small motifs preferentially, the hardness of the constraints in the search for the optimally treelike factor graph and constraints to find optimally symmetric neighbourhood covers.

Finally, we showed how this message passing formulation reproduces the generalised configuration model [26] when it is averaged over all networks in an ensemble of graphs that have an equivalent joint degree distribution.

These results give important insight into how the belief propagation algorithm can be applied to both random and empirical networks. Further work should be carried out to find the critical point of models that contain chordless cycles and cliques [12, 19].

### Appendix A: Number of connected graphs $\mathcal{Q}_{n,k}$

The number of connected graphs $\mathcal{Q}_{n,k}$ with $n$ vertices and $k$ edges has been discussed previously in the context of evaluating the percolation formulas for cliques [29, 36]. There are at least three well-known approaches to evaluating this quantity including an asymptotic expansion by Flajolet *et al* [44], a closed-form expression [29] and a fast recursive formula [45] due to Harary and Palmer. The recursion is given by

$$\mathcal{Q}_{n,k} = \begin{cases} 0 & k < n-1, \quad \text{or} \quad k > n(n-1)/2, \\ n^{n-2} & k = n-1, \\ Q(n,k) & \text{otherwise}. \end{cases} \tag{A1}$$

where

$$Q(n,k) = \binom{\frac{1}{2}n(n-1)}{k} - \sum_{m=0}^{n-2} \binom{n-1}{m} \sum_{p=0}^{k} \times \binom{\frac{1}{2}(n-1-m)(n-2-m)}{p} \mathcal{Q}_{m+1,k-p}. \tag{A2}$$

This recursive formula can yield the coefficients of cliques with size larger than 100 vertices in fractions of a second running on 16 GB Apple M1 Silicone using Python 3.10.

### Appendix B: ACKNOWLEDGMENTS

This work was partially supported by the UK Engineering and Physical Sciences Research Council under grant number EP/N007565/1 (Science of Sensor Systems Software).

### References

[*] pm78@st-andrews.ac.uk

[1] J. Yedidia, W. Freeman, and Y. Weiss, "Constructing free-energy approximations and generalized belief propagation algorithms," *IEEE Transactions on Information Theory*, vol. 51, no. 7, p. 2282–2312, 2005.

[2] M. E. Newman, *Networks.* Oxford University Press, 2019.

[3] G. T. Cantwell and M. E. Newman, "Message passing on networks with loops," *Proceedings of the National Academy of Sciences*, vol. 116, no. 47, p. 23398–23403, 2019.

[4] A. Kirkley, G. T. Cantwell, and M. E. Newman, "Belief propagation for networks with loops," *Science Advances*, vol. 7, no. 17, 2021.

[5] M. Mezard, G. Parisi, and R. Zecchina, "Analytic and algorithmic solution of random satisfiability problems," *Science*, vol. 297, no. 5582, p. 812–815, 2002.

[6] R. Monasson and R. Zecchina, "Statistical mechanics of the random $k$-satisfiability model," *Phys. Rev. E*, vol. 56, pp. 1357–1370, Aug 1997.

[7] M. Bronstein, "Beyond message passing: a physics-inspired paradigm for graph neural networks," *The Gradient*, 2022.

[8] C. Bodnar, F. Frasca, Y. Wang, N. Otter, G. F. Montufar, P. Lio, and M. Bronstein, "Weisfeiler and lehman go topological: Message passing simplicial networks," in *International Conference on Machine Learning*, pp. 1026–1037, PMLR, 2021.

[9] H. A. Bethe, "Statistical theory of superlattices," *Proceedings of the Royal Society of London. Series A - Mathematical and Physical Sciences*, vol. 150, no. 871, p. 552–575, 1935.

[10] D. Straszak and N. K. Vishnoi, "Belief propagation, bethe approximation and polynomials," in *2017 55th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pp. 666–671, Oct 2017.

[11] A. Weller, K. Tang, D. Sontag, and T. Jebara, "Understanding the bethe approximation: When and how can it go wrong?," in *Proceedings of the Thirtieth Conference on Uncertainty in Artificial Intelligence (UAI-14)*, 2014.

[12] B. Karrer, M. E. J. Newman, and L. Zdeborová, "Percolation on sparse networks," *Phys. Rev. Lett.*, vol. 113, p. 208702, Nov 2014.

[13] P. D. Gujrati, "Role of loop activity in correlated percolation: results on a husimi cactus and relationship with a bethe lattice," *Journal of Physics A: Mathematical and General*, vol. 34, pp. 9211–9230, oct 2001.

[14] M. Eckstein, M. Kollar, K. Byczuk, and D. Vollhardt, "Hopping on the bethe lattice: Exact results for densities of states and dynamical mean-field theory," *Phys. Rev. B*, vol. 71, p. 235119, Jun 2005.

[15] Y. Shiraki and Y. Kabashima, "Cavity analysis on the robustness of random networks against targeted attacks: Influences of degree-degree correlations," *Phys. Rev. E*, vol. 82, p. 036101, Sep 2010.

[16] F. L. Metz, I. Neri, and D. Bollé, "Spectra of sparse regular graphs with loops," *Phys. Rev. E*, vol. 84, p. 055101, Nov 2011.

[17] S. Yoon, A. V. Goltsev, S. N. Dorogovtsev, and J. F. F. Mendes, "Belief-propagation algorithm and the ising model on networks with arbitrary distributions of motifs," *Phys. Rev. E*, vol. 84, p. 041144, Oct 2011.

[18] M. E. J. Newman, "Spectra of networks containing short loops," *Phys. Rev. E*, vol. 100, p. 012314, Jul 2019.

[19] K. E. Hamilton and L. P. Pryadko, "Tight lower bound for percolation threshold on an infinite graph," *Phys. Rev. Lett.*, vol. 113, p. 208701, Nov 2014.

[20] A. Coolen, "Replica methods for loopy sparse random graphs," *Journal of Physics: Conference Series*, vol. 699,

p. 012022, mar 2016.

[21] B. J. Frey and D. MacKay, "A revolution: Belief propagation in graphs with cycles," in *Advances in Neural Information Processing Systems* (M. Jordan, M. Kearns, and S. Solla, eds.), vol. 10, MIT Press, 1997.

[22] M. E. J. Newman, X. Zhang, and R. R. Nadakuditi, "Spectra of random networks with arbitrary degrees," *Phys. Rev. E*, vol. 99, p. 042309, Apr 2019.

[23] M. Chertkov and V. Y. Chernyak, "Loop calculus in statistical physics and information science," *Phys. Rev. E*, vol. 73, p. 065102, Jun 2006.

[24] M. E. J. Newman, "Random graphs with clustering," *Phys. Rev. Lett.*, vol. 103, p. 058701, Jul 2009.

[25] J. C. Miller, "Percolation and epidemics in random clustered networks," *Phys. Rev. E*, vol. 80, p. 020901, Aug 2009.

[26] B. Karrer and M. E. J. Newman, "Random graphs containing arbitrary distributions of subgraphs," *Physical Review E*, vol. 82, no. 6, 2010.

[27] T. Hasegawa and S. Mizutaka, "Structure of percolating clusters in random clustered networks," *Phys. Rev. E*, vol. 101, p. 062310, Jun 2020.

[28] P. Mann, V. A. Smith, J. B. O. Mitchell, and S. Dobson, "Degree correlations in graphs with clique clustering," *Phys. Rev. E*, vol. 105, p. 044314, Apr 2022.

[29] P. Mann, V. A. Smith, J. B. O. Mitchell, C. Jefferson, and S. Dobson, "Exact formula for bond percolation on cliques," *Phys. Rev. E*, vol. 104, p. 024304, Aug 2021.

[30] P. Mann, V. A. Smith, J. B. O. Mitchell, and S. Dobson, "Percolation in random graphs with higher-order clustering," *Phys. Rev. E*, vol. 103, p. 012313, Jan 2021.

[31] P. Mann, V. A. Smith, J. B. O. Mitchell, and S. Dobson, "Random graphs with arbitrary clustering and their applications," *Phys. Rev. E*, vol. 103, p. 012309, Jan 2021.

[32] G. Burgio, A. Arenas, S. Gomez, and J. T. Matamalas, "Network clique cover approximation to analyze complex contagions through group interactions," *Communications Physics*, vol. 4, no. 1, 2021.

[33] T. Hasegawa and Y. Iwase, "Observability transitions in clustered networks," *Physica A: Statistical Mechanics and its Applications*, vol. 573, p. 125970, 2021.

[34] B. Karrer and M. E. J. Newman, "Message passing approach for general epidemic models," *Phys. Rev. E*, vol. 82, p. 016101, Jul 2010.

[35] This is certainly not true for motifs whose vertices do not belong to the same site (or more correctly - orbit). In that case, the bond occupancy probability depends on the identity of the vertices in the connected component (see [29]).

[36] M. E. J. Newman, "Properties of highly clustered networks," *Phys. Rev. E*, vol. 68, p. 026121, Aug 2003.

[37] S. Fortunato and D. Hric, "Community detection in networks: A user guide," *Physics Reports*, vol. 659, p. 1–44, 2016.

[38] M. Welling, "On the choice of regions for generalized belief propagation," in *Proceedings of the 20th Conference on Uncertainty in Artificial Intelligence*, UAI '04, (Arlington, Virginia, USA), p. 585–592, AUAI Press, 2004.

[39] T. Ritmeester and H. Meyer-Ortmanns, "Belief propagation for supply networks: Efficient clustering of their factor graphs," *The European Physical Journal B*, vol. 95, no. 5, 2022.

[40] T. Li, "Network dismantling on factor graphs: Break long loops and spare local structures," *New Journal of Physics*, vol. 23, no. 10, p. 103014, 2021.

[41] T. Li, P. Zhang, and H.-J. Zhou, "Long-loop feedback vertex set and dismantling on bipartite factor graphs," *Phys. Rev. E*, vol. 103, p. L061302, Jun 2021.

[42] M. E. J. Newman, "The structure of scientific collaboration networks," *Proceedings of the National Academy of Sciences*, vol. 98, no. 2, pp. 404–409, 2001.

[43] M. Boguñá, R. Pastor-Satorras, A. Díaz-Guilera, and A. Arenas, "Models of social networks based on social distance attachment," *Phys. Rev. E*, vol. 70, p. 056122, Nov 2004.

[44] P. Flajolet, B. Salvy, and G. Schaeffer, "Airy phenomena and analytic combinatorics of connected graphs," *The Electronic Journal of Combinatorics*, vol. 11, no. 1, 2004.

[45] F. Harary and E. M. Palmer, *Graphical enumeration.* Academic Press, 1973.