Motif-based mean-field approximation of interacting particles on clustered networks

Kai Cui [1,*] Wasiur R. KhudaBukhsh [2,†] and Heinz Koeppl[1,‡]

[1]*Department of Electrical Engineering and Information Technology, Technische Universität Darmstadt, 64287 Darmstadt, Germany*
[2]*University of Nottingham, Nottingham, United Kingdom*

Interacting particles on graphs are routinely used to study magnetic behavior in physics, disease spread in epidemiology, and opinion dynamics in social sciences. The literature on mean-field approximations of such systems for large graphs typically remains limited to specific dynamics, or assumes cluster-free graphs for which standard approximations based on degrees and pairs are often reasonably accurate. Here, we propose a motif-based mean-field approximation that considers higher-order subgraph structures in large clustered graphs. Numerically, our equations agree with stochastic simulations where existing methods fail.

With applications in as disparate branches of science as statistical physics [1], epidemiology [2–9], chemistry and systems biology [10,11], social science [12,13], and computer science [14–16], interacting particles on complex networks constitute an important class of models in the mathematician's and physicist's toolkit [17–20]. They describe systems where individual entities (particles), endowed with local states, interact with a subset of other entities (neighbors) and transition from one state to another as time evolves. For instance, in epidemiology the local state space consists of immunological statuses, such as susceptible, infected, removed, etc. Who interacts with whom defines a graph with the particles as the vertices.

The time evolution of the ensemble of particle states is often described by a continuous-time Markov jump process, for which discrete-time analysis can be insufficient [21]. As the number of particles increases, the exponentially growing combinatorial state space renders exact stochastic analysis prohibitive. To this end, the standard mean-field theoretic approach has been to describe the nonequilibrium dynamics of interacting particle systems via ordinary differential equations (ODEs) for the proportions of particles in each state. Together with control [22], learning-based methods [23,24] and graph limit theory [25–28], mean-field models can enable analysis of otherwise intractable settings [29–31]. More advanced mean-field approximations, such as heterogeneous mean fields [32], pair approximations [33–37], or approximate master equations (AMEs) [38–41] and extensions thereof [42–45], acknowledge the heterogeneity of the particles' behaviors due to the graph structure and incorporate vertices' degrees and edge counts (pairs). Though they provide reasonable accuracy for a number of applications, they are generally not asymptotically exact in that they do not agree with the functional law of large numbers (FLLN)

limits of the corresponding stochastic processes, agreeing only in certain special cases [46,47]. Even for calculations of critical parameter values, standard mean-field approximations are often inaccurate [48]. Nevertheless, their simplicity and intuitiveness have commonly justified mean-field approaches despite their inexactness.

In this Letter, we propose a simple and elegant derivation of a general motif-based mean-field approximation for interacting particles on bounded-degree graphs to address two crucial shortcomings of the state of the art: (i) The implicit assumption of cluster-free graphs [49]. In practice, the encountered graphs are far from cluster free and exhibit complex structures [14,50–52] (e.g., neural and transportation networks [53]), which greatly affect, e.g., cascades in correlated networks [54]. Here, we go beyond correlation coefficients and account for arbitrary subgraph structures called motifs [55] beyond standard degree and edge-based calculations. (ii) The restriction to special cases (e.g., susceptible-infectious-recovered (SIR) epidemics [56]) or dynamics driven by simple neighborhood counts. For instance, infection rates are often assumed to depend only on the number of infected neighbors, while in practice shared connections among neighbors and the shape of the induced neighborhood subgraph are too important to neglect (e.g., simplicial dynamics [14,57]). Though there exists a multitude of works on the analysis of clustered graphs [49,58–61], here we provide a general approximation that takes into account both of these aspects into a single coherent mean-field framework. We now introduce the mathematical model before explaining how our approximation addresses the above two issues.

Model. A convenient way to generate random graphs is via the configuration model (CM) [20,62], which allows specifying either a degree sequence or probability law from which the degrees are sampled. Each vertex is assigned as many half edges as its degree. We may need to add or drop a parity edge if the degree sequence is not graphical, but its contribution is negligible in large graphs [20, Sec. 7.6, pp. 239]. The configuration model graph is then constructed by uniformly-at-random matching of all available half edges. As

*kai.cui@bcs.tu-darmstadt.de
†wasiur.khudabukhsh@nottingham.ac.uk
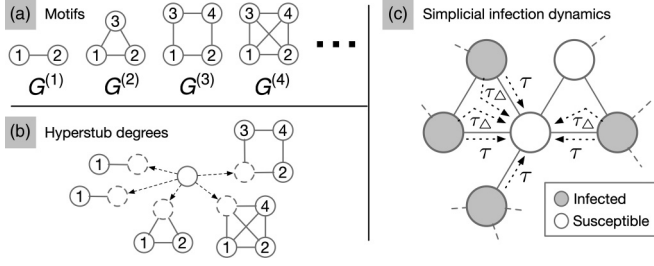‡heinz.koeppl@bcs.tu-darmstadt.de

FIG. 1. Schematic illustration of the model. (a) A finite number of network motifs occurring in the network. (b) An example ECM vertex with two motif participations as vertex 2 in $G^{(1)}$ and one each as vertices 3, 1, 3 in $G^{(2)}, G^{(3)}, G^{(4)}$ respectively, i.e., $d^1 \equiv (d^{1,1}, d^{1,2}) = (0,2), d^2 = (0,0,1), d^3 = (1,0,0,0), d^4 = (0,0,1,0)$. (c) Simplicial SIS dynamics [14] as an example of general neighborhood-dependent dynamics. Susceptible vertices are infected at rate τ by infected neighbors, and additionally at rate $\tau_\triangle$ for each shared neighbor.

N, the number of vertices, grows to infinity, the numbers of self-loops and multiple edges have independent Poisson limits whose means depend only on the first two moments of the degree distribution [19, Theorem 3.1.2]. Therefore, their contributions to the limits of various counts scaled by $1/N$ (standard mean-field scaling) vanish in the limit.

To introduce higher-order structure, we adopt the extended configuration model (ECM) [63]—also known as the hyperstub configuration model [56,64]. Denoting vertices and edges of graphs H by $V(H)$ and $E(H)$, respectively, and given M graphical network motifs $G^{(1)}, \ldots, G^{(M)}$ with $N_1, \ldots, N_M$ vertices, we construct an ECM on N vertices by specifying higher-order motif participation counts (hyperstub degrees) $(d_1, \ldots, d_N)$, where $d_v \equiv (d^1, \ldots, d^M) \in \mathcal{D}$, $d^i \equiv (d^{i,1}, \ldots, d^{i,N_i})$, and $d^{i,j} \in \mathbb{N}_0$ denotes the number of participations (hyperstubs) as the jth vertex (role) in the motif $G^{(i)}$ (see Fig. 1). As in the standard CM, hyperstubs are first generated for each node in accordance with a limiting hyperstub degree distribution $P(d)$. Subsequently, for each possible motif, we iteratively sample hyperstubs of each motif vertex role and add edges wherever the underlying motif has an edge, repeating until no hyperstubs are left.

To describe the dynamics of the interacting system, we will consider time-indexed colored ECM graphs $\{G_t\}_{t \geq 0}$. Each vertex is endowed with a local (finite) state space $\mathcal{X}$. Denote the state of vertex v at time t—interpreted as color—as $G_t[v]$, and define the colored neighborhoods $\mathcal{N}_t^{(v)}$ as colored subgraphs of G_t with fixed center vertex v, induced by the set of all vertices participating in motifs with v. Treated as a stochastic process, G_t is a Markov jump process with infinitesimal rates $\lambda^{x \to y}(\mathcal{N}_t^{(v)})$, depending on v only via its colored neighborhood configuration, i.e., the rate for vertex v to jump from state x to $y (\neq x)$ is given by

$$\mathbb{P}(G_{t+h}[v] = y \mid G_t[v] = x, \mathcal{N}_t^{(v)})$$
$$= \lambda^{x \to y}(\mathcal{N}_t^{(v)})h + o(h). \quad (1)$$

Note that the rate functions $\lambda^{x \to y}$ depend on the entire subgraph and its coloring up to isomorphism (not only neighbor state counts), and therefore generalize those considered in

standard mean-field approximations. To illustrate this, define the neighbor evaluation function

$$\psi(\mathcal{N}^{(v)}, f) \equiv \sum_{n \in V(\mathcal{N}^{(v)}):(v,n)\in E(\mathcal{N}^{(v)})} f(\mathcal{N}^{(v)}[n]) \quad (2)$$

for any $f : \mathcal{X} \to \mathbb{R}$ and colored neighborhood $\mathcal{N}^{(v)}$. Then, the simplicial susceptible-infected-susceptible (SIS) model [14], which imposes additional higher-order terms on the infection rates of vertices, can be modeled as

$$\lambda^{S \to I}(\mathcal{N}^{(v)}) = \tau \psi(\mathcal{N}^{(v)}, \mathbb{1}_{\{I\}}) + \tau_\triangle \sum_{(v,n,n')\in\Delta_v}$$
$$\times [\mathbb{1}_{\{I\}}(\mathcal{N}^{(v)}[n]) + \mathbb{1}_{\{I\}}(\mathcal{N}^{(v)}[n'])], \quad (3)$$

$$\lambda^{I \to S}(\mathcal{N}^{(v)}) = \gamma, \quad (4)$$

for pairwise infection rate τ, triangle (clique) infection rate $\tau_\triangle$, recovery rate γ, and indicator function $\mathbb{1}_A$. Here, the summation is over all unique triangles Δ_v involving v. This model is more realistic than the standard SIS model when shared acquaintances meet more often (see Fig. 1, [50]). In our experiments, we also consider the standard SIS model where $\tau_\triangle = 0$, which can also be understood as a result of microscopic contact processes [65], for which we similarly imagine higher-order interactions to be of interest. For a susceptible-infected-removed (SIR) model, $\lambda^{I \to S}$ is replaced by jumps to a third terminal state R. Finally, we consider the Ising Glauber dynamics [1] with states $\{U, D\}$ and

$$\lambda^{U \to D}(\mathcal{N}^{(v)}) = 1 - \lambda^{D \to U}(\mathcal{N}^{(v)})$$
$$= \left\{ 1 + \exp\left[\frac{2J}{T} \psi(\mathcal{N}^{(v)}, (-1)^{\mathbb{1}_{\{D\}}}) \right] \right\}^{-1} \quad (5)$$

for interaction strength $J > 0$ and temperature $T > 0$.

Mean-field approximation. While the exact colored graphs G_t can be evolved through their probability laws or their associated operator semigroup M, an exact analysis is typically prohibitive due to the combinatorial state space. In the limit of large graphs ($N \to \infty$), our aim will thus be to approximate, by a system of ODEs M', the evolution of certain population fractions, obtained by aggregating the colored graphs via some aggregation function φ—e.g., densities of different colors $\varphi_x(G) \equiv \frac{1}{N} \sum_v \mathbb{1}_{\{x\}}(G[v])$—such that the diagram

$$
\begin{array}{ccc}
G_0 & \xrightarrow{\ M\ } & G_t \\
\downarrow{\scriptstyle \varphi} & & \downarrow{\scriptstyle \varphi} \\
\rho_0 & \xrightarrow{\ M'\ } & \rho_t
\end{array}
\quad (6)
$$

commutes: The goal is to find a system of ODEs M' that accurately models the evolution of population fractions ρ, such that aggregating population fractions through φ and then applying M' is equivalent to first exactly evolving the system and then aggregating.

Since the degrees are bounded, the jumps of G_t are also bounded. Therefore, one expects the jumps of various $(1/N)$-scaled fractions to vanish in the limit because their quadratic variations (e.g., the running sum of squared jump sizes) vanish over finite time horizons. Consequently, even though the scaled proportions are not necessarily Markovian, their large-
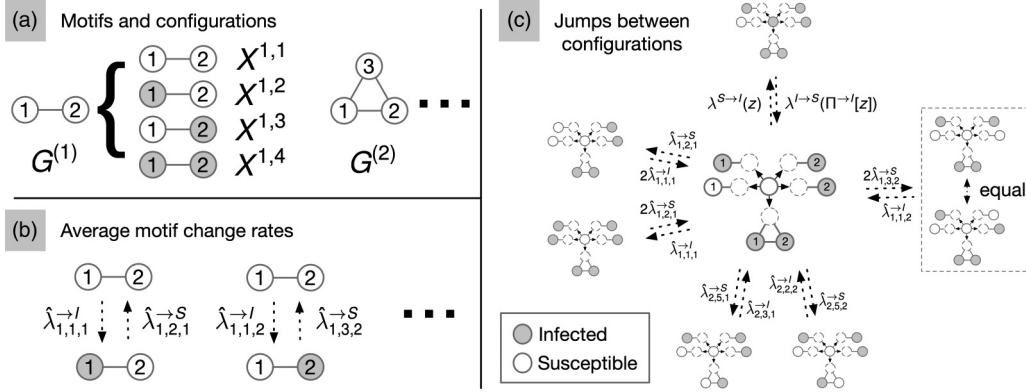
FIG. 2. Schematic illustration of the MMF equations for two states. (a) A fixed, finite number of network motifs with associated motif colorings. (b) Common, shared jump rates between motif colorings. (c) A visualization of all jumps from and to configuration z (center), where the jump rates are given by the number of motif configurations times their shared motif jump rates. Here, $d^1 = (2, 2)$, $d^2 = (0, 0, 1)$ and $z^{1,1} = (1, 1, 0, 0)$, $z^{1,2} = (0, 0, 2, 0)$, $z^{2,3} = (0, 1, 0, 0, 0, 0, 0, 0)$.

graph limits have continuous paths and can be described using ODEs by first performing the Doob-Meyer decomposition, which intuitively separates out a stochastic process that captures the mean of the scaled proportions and a zero-mean martingale (a stochastic process that acts like an error process or fluctuations around the mean process), and then invoking the FLLN for martingales [66,67] to claim that the fluctuations around the mean process vanish in the limit.

Denote the set of non-negative integer solutions to the Diophantine equation $y_1 + y_2 + \cdots + y_n = k$ by $\Theta(n, k)$. It is useful to think of $k \mapsto \Theta(n, k)$ as the equivalence class of a vector in $\mathbb{N}_0^n$ whose elements sum up to k (where two vectors are equivalent if their elements have the same sum). For motifs $G^{(i)}$, consider their set of distinct colorings $\mathcal{G}^{(i)}$ and $C_i \equiv |\mathcal{G}^{(i)}| = |\mathcal{X}|^{N_i}$. For a vertex with hyperstub degree d, the

possible counts of each neighboring motif coloring where the vertex participates as the jth vertex role in a motif $G^{(i)}$ are elements of $\Theta(C_i, d^{i,j})$.

Therefore, all colored neighborhoods $\mathcal{N}^{(v)}$ will belong to an equivalence class corresponding to a count vector (configuration) $z \in \mathcal{Z} \subset \times_{i=1}^{M} \times_{j=1}^{N_i} \mathbb{N}_0^{C_i}$ under an appropriate equivalence relation $\sim$, such that $z \equiv (z^{1,1}, \ldots, z^{1,N_1}, z^{2,1}, \ldots, z^{M,1}, \ldots, z^{M,N_M}) \in \mathcal{Z}$, $z^{i,j} \equiv (z_k^{i,j}, \ldots, z_{C_i}^{i,j})$, and $z_k^{i,j} \in \{0, 1, \ldots, d^{i,j}\}$ denotes the number of participations as role j in neighboring motifs $G^{(i)}$ currently in the kth motif coloring $X^{i,k} \in \mathcal{X}^{V(G^{(i)})}$. Denote the set of such equivalence classes that are compatible with d and the center vertex state x by $[x, d]$. To each $z \in [x, d]$ corresponds injectively some $z' = \Pi^{\to y}[z] \in [y, d]$, where the color of the center vertex is changed from x to y, and,
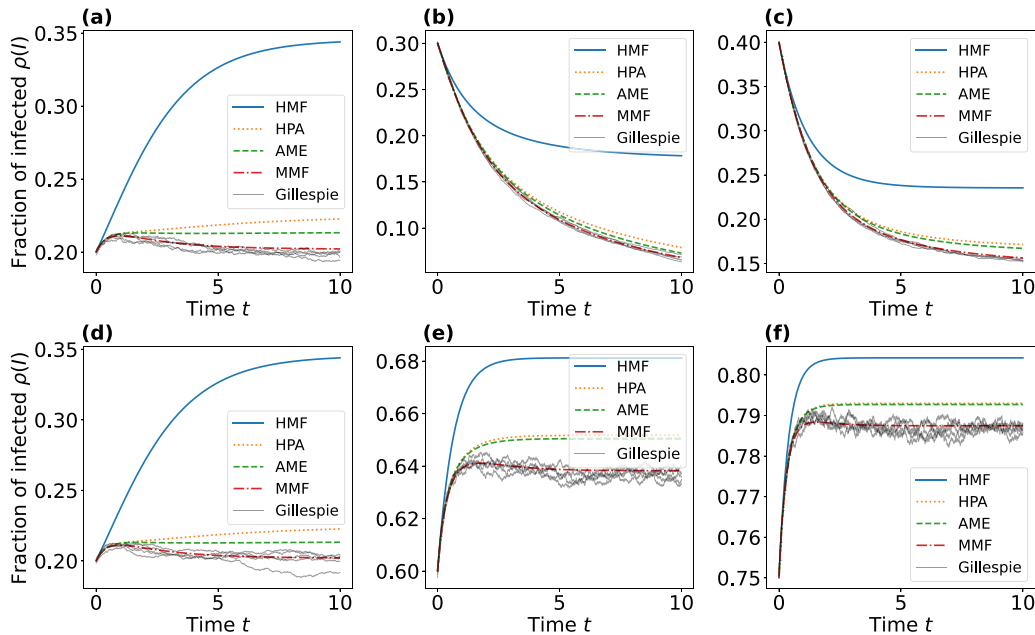

FIG. 3. Mean-field approximations in the SIS model using the edge and triangle graphs as motifs. (a)–(c) Results for $\tau = 0.3$, $\gamma = 0.9$. (a) $P^{a,3}$, $P_0(I) = 0.2$, (b) $P^{u,3}$, $P_0(I) = 0.3$, (c) $P^{d,2}$, $P_0(I) = 0.4$. (d)–(f) Results for $P^{a,3}$. (d) $\tau = 0.3$, $\gamma = 0.9$, $P_0(I) = 0.2$, (e) $\tau = 0.5$, $\gamma = 0.7$, $P_0(I) = 0.6$, (f) $\tau = 0.65$, $\gamma = 0.55$, $P_0(I) = 0.75$.
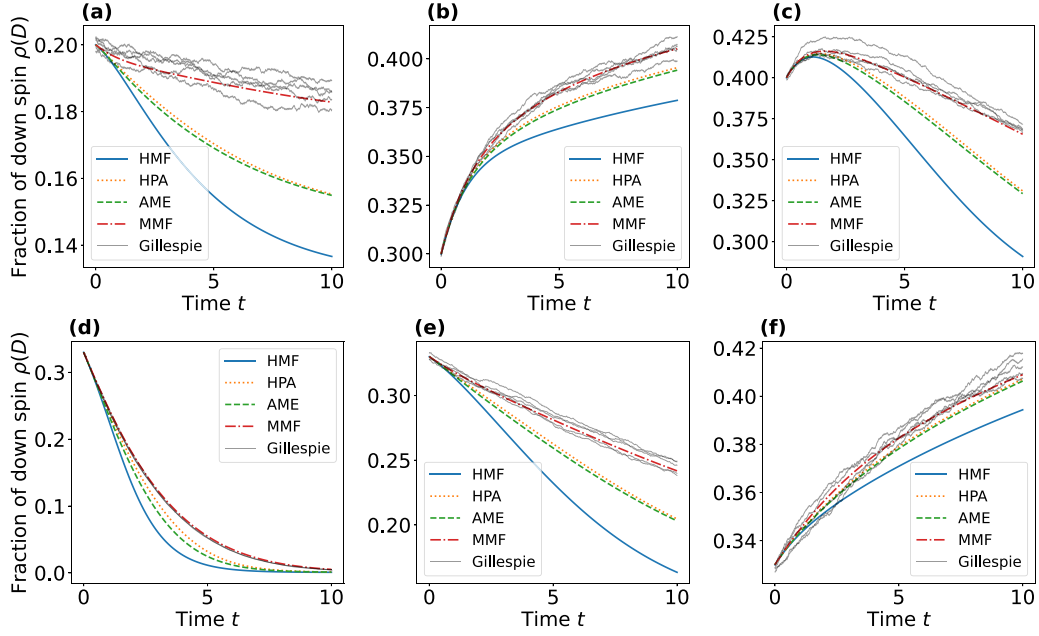
FIG. 4. Mean-field approximations in the Ising Glauber model using the edge and triangle graphs as motifs. (a)–(c) Results for $TJ^{-1} = 3$. (a) $P^{a,3}$, $P_0(D) = 0.2$, (b) $P^{u,3}$, $P_0(D) = 0.3$, (c) $P^{d,2}$, $P_0(D) = 0.4$. (d)–(f) Results for $P^{a,3}$, $P_0(D) = 0.33$. (d) $TJ^{-1} = 1$, (e) $TJ^{-1} = 3$ (f) $TJ^{-1} = 4$.

analogously, $z'' = \Pi_{i,k,v}^{\rightarrow y}[z] \in [x, d]$, where the color of a neighboring vertex that participates as role v in motif $G^{(i)}$ with current motif coloring k, $X^{i,k} \in \mathcal{X}^{V(G^{(i)})}$, is changed from $X_v^{i,k}$ to y. Moreover, each $z \in [x, d]$ determines the colored neighborhood (up to isomorphism) of a center vertex with color x and hyperstub degree d.

Aggregating colored ECMs over equivalence classes from the quotient space $\mathcal{G}/\sim$, where $\mathcal{G}$ is the space of all colored ECMs, is tantamount to keeping track of proportions $\rho_t(x, d, z)$ of vertices in G_t with color $x \in \mathcal{X}$, hyperstub degree d, and counts of neighboring motif colorings $z \in [x, d]$. Note that although z already contains all information about x, d, for notational convenience we track proportions of (x, d, z). As $N \rightarrow \infty$, these proportions can be described by deterministic ODEs, which we shall call the motif-based mean-field (MMF) equations.

This leads us to our main result: The MMF master equations for the limiting proportions ρ_t constitute a system M' of ODEs in (6) with an accuracy going beyond existing mean-

field approximations, and are given by

$$\dot{\rho}_t(x, d, z) = \sum_{y \in \mathcal{X}} (\Lambda^{\leftarrow y} - I) \rho_t(x, d, z) \lambda^{x \rightarrow y}(z)$$
$$+ \sum_{y \in \mathcal{X}} \sum_{i,j,k,v \neq j} (\Lambda_{i,j,k,v}^{\leftarrow y} - I) \rho_t(x, d, z) z_k^{i,j} \hat{\lambda}_{i,k,v}^{\rightarrow y},$$
(7)

where we aggregate rates $\lambda^{x \rightarrow y}(z)$ and $z_k^{i,j} \hat{\lambda}_{i,k,v}^{\rightarrow y}$ over equivalence classes corresponding to each center vertex configuration z (since z uniquely defines the colored neighborhood up to isomorphism) and each coloring k of neighboring motifs $G^{(i)}$, respectively. Here, we defined unit operators I and influx step operators $\Lambda^{\leftarrow y}$, $\Lambda_{i,j,k,v}^{\leftarrow y}$ acting on functions $f(x, d, z, y)$, $f(x, d, z, y, k)$ such that we have influx by center vertex jumps from configurations $\Pi^{\rightarrow y}[z]$,

$$[\Lambda^{\leftarrow y} f](x, d, z, y) = f(y, d, \Pi^{\rightarrow y}[z], x), \quad (8)$$
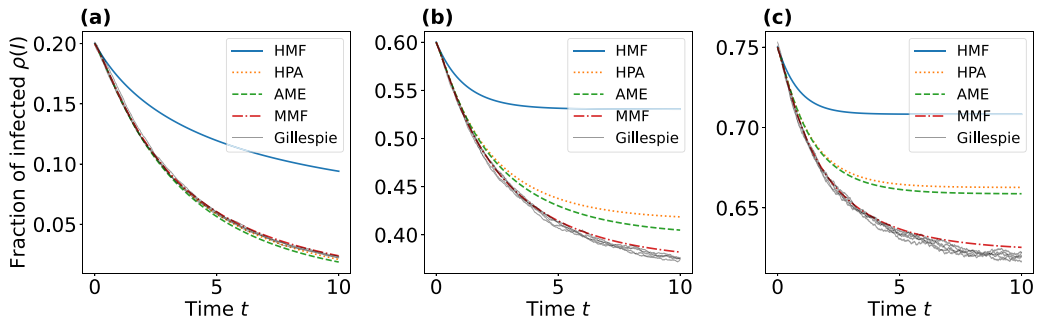

FIG. 5. Mean-field approximations in the SIS model using the edge and square graphs as motifs and $P^{a,2}$. (a) $\tau = 0.3$, $\gamma = 0.9$, $P_0(I) = 0.2$, (b) $\tau = 0.5$, $\gamma = 0.7$, $P_0(I) = 0.6$, (c) $\tau = 0.65$, $\gamma = 0.55$, $P_0(I) = 0.75$.
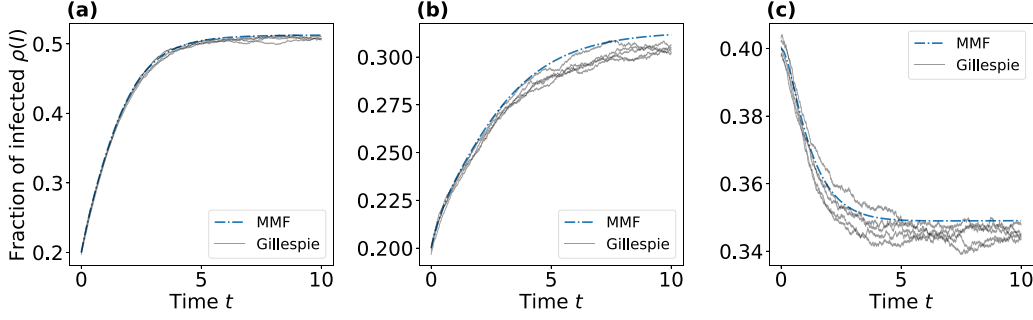
FIG. 6. A comparison between MMF and numerical Gillespie simulation in the simplicial SIS model using $\tau_\triangle = \tau$, the edge and triangle graphs as motifs and $\tau = 0.3$, $\gamma = 0.9$. (a) $P^{a,3}$, $P_0(I) = 0.2$, (b) $P^{u,3}$, $P_0(I) = 0.2$, (c) $P^{d,2}$, $P_0(I) = 0.4$.

and similarly influx by jumps of all neighboring motifs' vertices that are not the center vertex ($v \neq j$),

$$\left[\Lambda_{i,j,k,v}^{\leftarrow y} f\right](x, d, z, y, k) = f\left(x, d, \Pi_{i,k,v}^{\rightarrow y}[z], X_v^{i,k}, \Omega_{i,v}^{\rightarrow y}[k]\right),$$
(9)

where $\Omega_{i,v}^{\rightarrow y}[k]$ denotes the motif coloring resulting from changing the color of vertex v to y in motif i with coloring k. The jump rates of any neighbors in role v of motif $G^{(i)}$ with coloring k from the corresponding state $\tilde{x} = X_v^{i,k}$ to y are approximated by the averaged jump rate over all such colored motif occurrences,

$$\hat{\lambda}_{i,k,v}^{\rightarrow y} \equiv \frac{\sum_{d,z} \rho_t(\tilde{x}, d, z) z_k^{i,v} \lambda^{\tilde{x} \rightarrow y}(z)}{\sum_{d,z} \rho_t(\tilde{x}, d, z) z_k^{i,v}},$$
(10)

since a vertex in configuration z participates $z_k^{i,v}$ times in the considered motif coloring. See Fig. 2 for a visualization. Finally, sampling independent and identically distributed (i.i.d.) initial states from some $P_0 : \mathcal{X} \rightarrow [0, 1]$, the initial conditions are given by

$$\rho_0(x, d, z) = P_0(x) P(d) \prod_{i,j,k,v \neq j} \left[P_0\left(X_v^{i,k}\right) \cdot \mathbb{1}_{\{x\}}\left(X_j^{i,k}\right)\right]^{z_k^{i,j}},$$
(11)

where $0^0 \equiv 1$. The fractions of vertices in any state x are then given by $\rho_t(x) = \sum_{d,z} \rho_t(x, d, z)$.

The biggest appeal of the MMF equations (7) is their simplicity and intuitiveness. While they may generally not be asymptotically exact, experimentally we find that they are

quite accurate. Note that as a special case, we obtain classical approximations such as AMEs [38] and thereby coarser approximations [39] for degree distributions $\tilde{P} : \mathbb{N}_0 \rightarrow [0, 1]$ by considering only the edge motif $G^{(1)}$, assuming binomial role distributions and aggregating equivalent terms, i.e., $P(d^{1,1}, d^{1,2}) = \tilde{P}(d^{1,1} + d^{1,2}) \cdot \binom{d^{1,1} + d^{1,2}}{d^{1,1}} (\frac{1}{2})^{d^{1,1}} (\frac{1}{2})^{d^{1,2}}$.

Numerical evaluation. For numerical purposes, we generate equations only for P-supported hyperstub degrees d and simulate rescaled proportions $\rho_t(x, z \mid d) \equiv \rho_t(x, d, z)/P(d)$. For fast ECM graph generation, we drop leftover hyperstubs (in our experiments, this amounts to less than 0.5% of all generated stubs, leading to only slight inaccuracies) instead of resampling until cardinality constraints are satisfied and allow, but ignore, self-loops and multi-edges. We use a third-order numerical integrator and compare MMF against the approximate master equations (AMEs) [38], the heterogeneous pair approximation (HPA) [35], the heterogeneous mean-field approximation (HMF) [32], and exact Gillespie simulations on graphs of size $N = 100\,000$. For use by the wider community, PYTHON code is available at [68].

For two given, arbitrary network motifs $G^{(1)}$, $G^{(2)}$, we consider the three parametrized families of antidiagonal, uniform, and diagonal hyperstub degree distributions $P^{a,\theta}$, $P^{u,\theta}$, and $P^{d,\theta}$ with parameter $\theta \in \mathbb{N}$. For $P^{a,\theta}$, we put uniform mass $1/(\theta + 1)$ on each case, where $\sum_j d^{1,j} = k$ and $\sum_j d^{2,j} = \theta - k$ for $k = 0, 1, \ldots, \theta$. In each case, we shall assume a uniform distribution over motif roles, resulting in a product of multinomials $P^{a,\theta}(d) \equiv \frac{1}{(\theta+1)} \sum_{k=0}^{\theta} \mathbb{1}_{\{k\}}(\sum_j d^{1,j} + d^{2,j}) \cdot \prod_{i \in \{1,2\}} \text{Mult}(d^i \mid k, \frac{1}{N_i} 1_{N_i})$, where 1_{N_i} is the N_i-dimensional
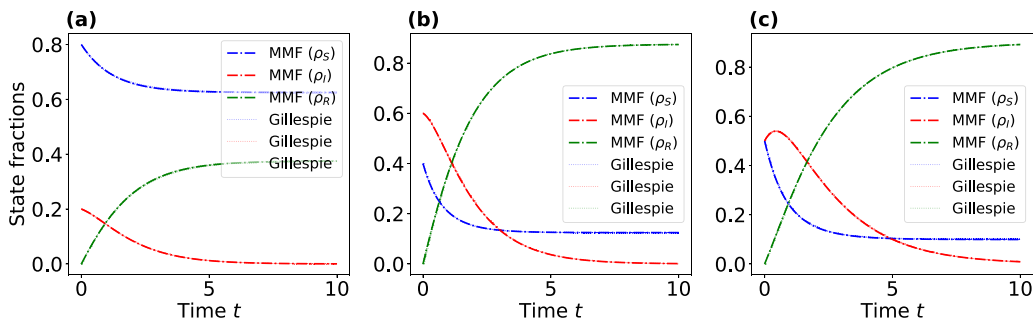

FIG. 7. The MMF approximation is almost indiscernible from the numerical simulation in the SIR model using the edge and triangle graphs as motifs and $P^{a,2}$. (a) $\tau = 0.3$, $\gamma = 0.9$, $1 - P_0(S) = P_0(I) = 0.2$, (b) $\tau = 0.5$, $\gamma = 0.7$, $1 - P_0(S) = P_0(I) = 0.6$, (c) $\tau = 0.6$, $\gamma = 0.5$, $1 - P_0(S) = P_0(I) = 0.5$.

one-vector. For $P^{u,\theta}$ and $P^{d,\theta}$, we similarly put equal probability mass whenever $\sum_j d^{1,j} + \sum_j d^{2,j} \leqslant \theta$ and $\sum_j d^{1,j} = \sum_j d^{2,j} = \theta$, respectively.

On the ECM graphs with edge and triangle motifs ($G^{(1)}$, $G^{(2)}$ from Fig. 1), we find that our approximation matches well with the numerical Gillespie simulation. For the SIS dynamics (3) and (4) in Fig. 3, our approximation outperforms other approximation methods over a range of (hyperstub) degree distributions and dynamics parameters. Similar assertions hold for the Ising Glauber dynamics (5) in Fig. 4, where existing mean-field approximations become highly inaccurate near the critical point due to the high clustering of the considered graphs. Furthermore, our approximations remain quite accurate also, e.g., for graphs with edge and square motifs ($G^{(1)}$, $G^{(3)}$ in Fig. 1) as seen in Fig. 5. For the simplicial version of the SIS dynamics, in Fig. 6 we find that the accuracy of our approximations is acceptable, while existing degree-based approximations are unable to handle simplicial dynamics by design. Finally, we verify the accuracy of our proposed framework on the SIR dynamics model in Fig. 7 with nonbinary states, where the Gillespie simulation for $N = 100\,000$ is almost indiscernible from the predicted mean-field proportions, showing the generality of our approach.

Discussion. We have proposed motif-based mean-field equations for arbitrary neighborhood-dependent jump dynamics on a highly adjustable random graph model, considering both higher-order graph structures and dynamics. Numerical examples show that our approximations are quite accurate. Potential extensions include the consideration of general k-hop neighborhoods with $k > 1$, control and lumping of equations [69,70] under additional assumptions on motif roles to improve tractability. Finally, for applications, estimating hyperstub degree distributions constitutes another important problem, as an identifiability problem arises from counting larger motifs that include smaller motifs.

[1] R. J. Glauber, J. Math. Phys. **4**, 294 (1963).

[2] I. Z. Kiss, J. C. Miller, and P. L. Simon, *Mathematics of Epidemics on Networks: From Exact to Approximate Models*, Vol. 46 (Springer, New York, 2017).

[3] F. Darabi Sahneh, C. Scoglio, and P. Van Mieghem, IEEE/ACM Trans. Network. **21**, 1609 (2013).

[4] R. Pastor-Satorras and A. Vespignani, Phys. Rev. Lett. **86**, 3200 (2001).

[5] R. Pastor-Satorras, C. Castellano, P. Van Mieghem, and A. Vespignani, Rev. Mod. Phys. **87**, 925 (2015).

[6] J. Sanz, C.-Y. Xia, S. Meloni, and Y. Moreno, Phys. Rev. X **4**, 041005 (2014).

[7] J. S. Juul and S. H. Strogatz, Phys. Rev. Research **2**, 033005 (2020).

[8] B. F. Nielsen, L. Simonsen, and K. Sneppen, Phys. Rev. Lett. **126**, 118301 (2021).

[9] J. Hindes and I. B. Schwartz, Phys. Rev. Lett. **117**, 028302 (2016).

[10] D. F. Anderson and T. D. Nguyen, arXiv:1910.12723.

[11] D. F. Anderson and T. D. Nguyen, J. Math. Chem. **59**, 2063 (2021).

[12] K. Sugishita, M. A. Porter, M. Beguerisse-Díaz, and N. Masuda, Phys. Rev. Research **3**, 023249 (2021).

[13] J. S. Juul and M. A. Porter, Phys. Rev. E **99**, 022313 (2019).

[14] D. Bruneo, M. Scarpa, A. Bobbio, D. Cerotti, and M. Gribaudo, Perform. Eval. **69**, 135 (2012).

[15] M. Gribaudo, D. Cerotti, and A. Bobbio, in *2008 Sixth Annual IEEE International Conference on Pervasive Computing and Communications (PerCom)* (Curran Associates, Inc., New York, 2008), pp. 300–305

[16] P. Van Mieghem, J. Omic, and R. Kooij, IEEE/ACM Trans. Network. **17**, 1 (2009).

[17] R. Albert and A.-L. Barabási, Rev. Mod. Phys. **74**, 47 (2002).

[18] A. Barrat, M. Barthlemy, and A. Vespignani, *Dynamical Processes on Complex Networks* (Cambridge University Press, Cambridge, England, 2008).

[19] R. Durrett, *Random Graph Dynamics* (Cambridge University Press, New York, 2006).

[20] R. van der Hofstad, *Random Graphs and Complex Networks Vol. 1* (Cambridge University Press, New York, 2016).

[21] P. G. Fennell, S. Melnik, and J. P. Gleeson, Phys. Rev. E **94**, 052125 (2016).

[22] J.-M. Lasry and P.-L. Lions, Jpn. J. Math. **2**, 229 (2007).

[23] X. Guo, A. Hu, R. Xu, and J. Zhang, in *Advances in Neural Information Processing Systems*, edited by H. Wallach, H. Larochelle, A. Beygelzimer, F. d' Alché-Buc, E. Fox, and R. Garnett (Curran Associates, Inc., New York, 2019), Vol. 32.

[24] K. Cui and H. Koeppl, in *Proceedings of The 24th International Conference on Artificial Intelligence and Statistics, Proceedings of Machine Learning Research*, edited by A. Banerjee and K. Fukumizu (PMLR, 2021), pp. 1909–1917, Vol. 130.

[25] L. Lovász, *Large Networks and Graph Limits* (American Mathematical Society, Providence, Rhode Island, 2012).

[26] P. E. Caines and M. Huang, in *2019 IEEE 58th Conference on Decision and Control (CDC)* (Curran Associates, Inc., New York, 2019), pp. 286–292.

[27] E. Bayraktar, S. Chakraborty, and R. Wu, arXiv:2003.13180.

[28] K. Cui and H. Koeppl, arXiv:2112.01280.

[29] L. Laguzet and G. Turinici, Bull. Math. Biol. **77**, 1955 (2015).

[30] B. Djehiche, A. Tcheukam, and H. Tembine, AIMS Electron. Electric. Eng. **1**, 18 (2017).

[31] S. Perrin, M. Laurière, J. Pérolat, M. Geist, R. Élie, and O. Pietquin, *Proceedings of the Thirtieth International Joint*

Conference on Artificial Intelligence, IJCAI-21, edited by Z.-H. Zhou (IJCAI, 2022), PP. 356–362.

[32] V. Sood and S. Redner, Phys. Rev. Lett. **94**, 178701 (2005).

[33] D. A. Rand, Correlation equations and pair approximations for spatial ecologies, in *Advanced Ecological Theory* (Wiley, New York, 1999), Chap. 4, pp. 100–142.

[34] F. Vazquez and V. M. Eguíluz, New J. Phys. **10**, 063011 (2008).

[35] E. Pugliese and C. Castellano, Europhys. Lett. **88**, 58004 (2009).

[36] D. H. Silva, S. C. Ferreira, W. Cota, R. Pastor-Satorras, and C. Castellano, Phys. Rev. Research **1**, 033024 (2019).

[37] L. Pellis, T. House, and M. J. Keeling, J. Theor. Biol. **382**, 160 (2015).

[38] J. P. Gleeson, Phys. Rev. Lett. **107**, 068701 (2011).

[39] J. P. Gleeson, Phys. Rev. X **3**, 021004 (2013).

[40] P. G. Fennell and J. P. Gleeson, SIAM Rev. **61**, 92 (2019).

[41] J. Lindquist, J. Ma, P. Van den Driessche, and F. H. Willeboordse, J. Math. Biol. **62**, 143 (2011).

[42] A. F. Peralta and R. Toral, Chaos: Interdisciplin. J. Nonlinear Sci. **28**, 106303 (2018).

[43] A. F. Peralta and R. Toral, Phys. Rev. Research **2**, 043370 (2020).

[44] F. Farkhooi and W. Stannat, Phys. Rev. Lett. **119**, 208301 (2017).

[45] C.-R. Cai, Z.-X. Wu, M. Z. Q. Chen, P. Holme, and J.-Y. Guan, Phys. Rev. Lett. **116**, 258301 (2016).

[46] K. A. Jacobsen, M. G. Burch, J. H. Tien, and G. A. Rempała, J. Biol. Dynam. **12**, 746 (2018).

[47] W. R. KhudaBukhsh, C. Woroszylo, G. Rempała, and H. Koeppl, Adv. Appl. Probab. (2022), (to be published), arXiv:1703.06328.

[48] S. Chatterjee and R. Durrett, Ann. Prob. **37**, 2332 (2009).

[49] J. P. Gleeson, S. Melnik, J. A. Ward, M. A. Porter, and P. J. Mucha, Phys. Rev. E **85**, 026106 (2012).

[50] F. Battiston, G. Cencetti, I. Iacopini, V. Latora, M. Lucas, A. Patania, J.-G. Young, and G. Petri, Phys. Rep. **874**, 1 (2020).

[51] G. St-Onge, H. Sun, A. Allard, L. Hébert-Dufresne, and G. Bianconi, Phys. Rev. Lett. **127**, 158301 (2021).

[52] G. St-Onge, V. Thibeault, A. Allard, L. J. Dubé, and L. Hébert-Dufresne, Phys. Rev. Lett. **126**, 098301 (2021).

[53] A. R. Benson, D. F. Gleich, and J. Leskovec, Science **353**, 163 (2016).

[54] J. P. Gleeson, Phys. Rev. E **77**, 046117 (2008).

[55] A. C. Schwarze and M. A. Porter, SIAM J. Appl. Dynam. Sys. **20**, 2516 (2021).

[56] M. Ritchie, L. Berthouze, and I. Z. Kiss, J. Math. Biol. **72**, 255 (2016).

[57] G. St-Onge, V. Thibeault, A. Allard, L. J. Dubé, and L. Hébert-Dufresne, Phys. Rev. E **103**, 032301 (2021).

[58] N. Persky, R. Ben-Av, I. Kanter, and E. Domany, Phys. Rev. E **54**, 2351 (1996).

[59] V. Vlasov and A. Bifone, Sci. Rep. **7**, 10403 (2017).

[60] C. G. Assisi, V. K. Jirsa, and J. A. Scott Kelso, Phys. Rev. Lett. **94**, 018106 (2005).

[61] A. Fronczak, P. Fronczak, and J. A. Hołyst, Phys. Rev. E **68**, 046126 (2003).

[62] M. Newman, *Networks* (Oxford University Press, Oxford, 2018).

[63] B. Karrer and M. E. J. Newman, Phys. Rev. E **82**, 066118 (2010).

[64] M. Ritchie, L. Berthouze, and I. Z. Kiss, J. Complex Networks **5**, 1 (2017).

[65] S. Gómez, A. Arenas, J. Borge-Holthoefer, S. Meloni, and Y. Moreno, Europhys. Lett. **89**, 38009 (2010).

[66] T. G. Kurtz, *Approximation of Population Processes* (Society for Industrial and Applied Mathematics, Philadelphia, 1981).

[67] J. Jacod and A. Shiryaev, *Limit Theorems for Stochastic Processes* (Springer, Berlin, 2003).

[68] K. Cui, W. R. KhudaBukhsh, and H. Koeppl, Motif-based mean-field equations,.

[69] G. Großmann, C. Kyriakopoulos, L. Bortolussi, and V. Wolf, in International Conference on Quantitative Evaluation of Systems (Springer, New York, 2018) pp. 157–172.

[70] W. R. KhudaBukhsh, A. Auddy, Y. Disser, and H. Koeppl, J. Appl. Probab. **56**, 647 (2019).