Asymptotically inspired moment-closure approximation for adaptive networks

Maxim S. Shkarayev and Leah Shaw

Applied Sciences Department, College of William & Mary

Adaptive social networks, in which nodes and network structure co-evolve, are often described using a mean-field system of equations for the density of node and link types. These equations constitute an open system due to dependence on higher order topological structures. We propose a moment-closure approximation based on the analytical description of the system in an asymptotic regime. We apply the proposed approach to two examples of adaptive networks: recruitment to a cause model and adaptive epidemic model. We show a good agreement between the improved mean-field prediction and simulations of the full network system.

Keywords: adaptive networks, moment closure approximation, network dynamics, SIS

I. INTRODUCTION

In recent years we have seen a lot of progress in the field of network dynamics and dynamics on the networks. Strong interest in understanding the phenomena like disease spread on social networks, the social interaction via a number of online tools such as Facebook and Twitter, dynamics of neuronal networks and many others, encouraged development of mathematical tools necessary to analyze the behavior of such systems.

Often the first step in analyzing such systems is to represent them as networks, where an individual unit, e.g. a person, a user account, a neuron, is represented by a node, and possibility of interaction between any two units is represented by a link between them. Networks where dynamical processes on the nodes affect the network structure, which in turn affects the dynamics on the nodes, are called adaptive [1–5]. The dynamical processes on such networks are often characterized by their statistical properties via a mean-field approach. Mean-field equations constitute a hierarchy of equations, where the expected state of nodes, due to interaction via network, is coupled to the statistical description of links in the networks. The dynamical evolution of the links in turn depends on the evolution of statistical description of node triples, which in their turn depend on higher order structures and so on. In other words, mean-field description yields an infinite system of coupled equations, which often must be truncated in order to be solvable. The truncated system is open and has to be closed by introducing additional information about the system.

The dynamics on network systems are often closed at the level link equations, where the first order of network information appears [6]. Perhaps the simplest closure approach is based on the assumption of homogeneous distribution of different node types in the system, and that the probability of finding a particular type of node in the neighborhood of a given node is independent of what else can be found in that node’s neighborhood. This closure was shown to produce excellent results for many different systems [6–12]. Heterogeneous mean-field approach, where conditioning on the total degree of nodes is introduced may improve the accuracy of the approximation, although drastically increasing the number of equations in the description [13]. Often, additional information about the system, such as the expected clustering coefficient, may be used to improve the closure [9]. In other cases, assumptions about the shape of degree distribution functions, possibly guided by the numerical simulations or the physical observations, may lead to an improvement in closure [7, 11, 14]. Equation-free approaches may also be used when closing the mean-field equations [15].

All of the above closures often lead to an excellent approximation of the system dynamics. However, they all suffer either from the fact that no a priori knowledge of the accuracy of those approximations can be obtained or from resulting in an excessive number of equations that must be analyzed. In this paper, we propose a new method that may lead to accurate closures, which would also allow one to manage the expectation of the accuracy of the obtained closure. The proposed approach is based on simplification of the mean-field system of equations in some asymptotic regime. In the rest of the paper we demonstrate our approach by applying it to two adaptive network systems. In section II, we derive a closure for a system modeling recruitment to a cause [3]. In section III we derive an improved closure for an adaptive epidemic model [2].
The link rewiring, shown in Fig. (b), takes place at a transitions between the classes are marked by the arrows. The newborn nodes are born into the N-class at rate $\lambda_2$ and R-nodes. The system is open in the sense that nodes can die at rate $\gamma$. The expected number of $\mathrm{S}$-nodes in the network, $\rho_\alpha$, is given in Eqs. (A1a)-(A1i). We characterize the time evolution of the expected number of node chains that originate as a given N-node (S-node) that is connected to an S-node, which is turn is connected to an R-node. The terms $N_{X_1}$, $X_2$, etc. for example $N_S$ is the expected number of S-nodes in the network, while $N_{RS}$ is the expected number of links with S- and R-nodes at its ends. In our definition of node chains we require the ith and i plus first nodes to be different; however, i and i plus second nodes can in fact be the same node. In the example of a network presented in Fig. 1(c) there are 4 RSR triples, corresponding to the following node combinations: 1-2-1, 1-2-3, 3-2-1, 3-2-3. Note that the order at which nodes appear matters.

The master equations are multidimensional and, therefore, are extremely difficult to analyze. A common way to analyze the dynamics of social networks is via mean-field equations. These can be generated by multiplying the master equations by $k_1^i k_2^i k_3^i$ for some non-negative integer values of $i_j$, and summing over k. Thus, the equations describing node dynamics are obtained by taking $i_1+i_2+i_3 = 0$, as given in Eqs. (A1a)-(A1c), while the description of the link dynamics is obtained by taking $i_1+i_2+i_3 = 1$, as given in Eqs. (A1d)-(A1i).
The hierarchy of equations generated in this manner must be truncated in order to obtain a finite dimensional description of the system. However, as evidenced by the system of equations in (A1), which contains dependence on statistical description of high-order structures, i.e. the terms \( N_{\text{NSR}}, N_{\text{SSR}} \) and \( N_{\text{RSR}} \), such truncation leaves the system open and in need of closure. The usual approach comes from the assumption of homogeneous distribution of the R-nodes throughout the network, which leads to the following closure equations:

\[
\frac{N_{\text{NSR}}}{N_S} = \frac{N_{\text{SSR}} N_{\text{RS}}}{N_S}, \quad (2a)
\]

\[
\frac{N_{\text{RSR}}}{N_S} = \left( \frac{N_{\text{RS}}}{N_S} \right)^2 + \frac{N_{\text{RS}}}{N_S}, \quad (2b)
\]

where we also assumed the the node degree distribution of susceptible nodes is Poisson. These closures are ad hoc, and may fail to capture the system behavior accurately if, for example, correlations are present. Here we develop an approach that derives the closure based on the system behavior in some asymptotic regime. In particular, we derive the equations that describe the evolution of node triples, take a steady-state relation and consider it in the asymptotic regime, where we are able to close the equations. Finally, we use the derived closures in other parameter regimes and outside of the steady-state and evaluate the performance of the newly derived closures.

A. Closing of NSR and SSR terms

We consider the evolution of the expected number of node triples in the limit of \( \gamma, \delta, \mu / (N_N + N_S + N_R) \ll w, \lambda_1, \lambda_2 \). We consider the following relation:

\[
\sum_k \left( \frac{N_{\text{RS}}}{N_N} \frac{\partial \rho_{S,k}}{\partial N_N} + \frac{N_{\text{RN}}}{N_N} \frac{\partial \rho_{N,k}}{\partial N_N} \right) (k_1 k_3 + k_2 k_3). \quad (3)
\]

This relation is evaluated at the steady state and using Eq. (A2a) and (A2b). Leaving the algebra to the reader, the above relation leads to

\[
\left( \frac{N_{\text{NSR}}}{N_S} + \frac{N_{\text{SSR}}}{N_S} - \frac{N_{\text{SR}} N_{\text{RS}}}{N_S} - \frac{N_{\text{SS}} N_{\text{RS}}}{N_S} \right) \times \left( \frac{N_{\text{RN}}}{N_N} - \frac{N_{\text{RS}}}{N_N} \right) = 0, \quad (4)
\]

a result that supports the closure in Eqs. (2a) for \( X = N \) and \( S \), although does not imply it.

![Fig. 2. Relative errors in \( N_{\text{NSR}} + N_{\text{SSR}} \) and \( N_{\text{NSR}} \) closures as a function of \( \gamma \) for \( w = 10^0 \) (circle, green online), \( w = 10^1 \) (triangle, red online), \( w = 10^2 \) (cross, black online). Here and in the rest of the paper, the relative error, \( \Delta \), is defined as the modulus of one minus the ratio of approximated to the measured quantity. The simulations are performed following the continuous time algorithm introduced in [17]. The other parameters are \( \delta = 1, \sigma = 10, \mu = 10^0 \).](image)

We can see in Fig. 2(a) that the expected number of NSR and SSR triples per susceptible node, \( N_{\text{NSR}}/N_S + N_{\text{SSR}}/N_S \), is well approximated by the closure of Eq. (4) in the large \( \lambda_1 \) and \( \lambda_2 \) limit, further improving as \( w \) is increased. According to Fig. 2(c), the closure of \( N_{\text{NSR}}/N_S + N_{\text{NSR}}/N_S \) continues to hold even in the parameter regime, outside of the considered limit. As for the individual closures, Fig. 2(b) shows that the closure of \( N_{\text{SSR}}/N_S \) holds as well in the large \( \lambda_1 \) and \( \lambda_2 \) regime. However, as we can see in Fig. 2(d), the closure completely fails apart for small \( \lambda_1 \) and \( \lambda_2 \), especially as \( \gamma \) becomes dominant. Note that the error in closure of \( N_{\text{SSR}}/N_S + N_{\text{NSR}}/N_S \) is controlled and, as expected, performs well, while \( N_{\text{SSR}}/N_S \) closure is ad hoc and, therefore, deviation from simulations is not unexpected.

B. Closure of RSR term

In order to develop a closure for the \( N_{\text{RSR}} \) term, we consider the evolution of the expected number of
RSR triples:
\[
\begin{align*}
\delta_t N_{\text{RSR}} &= -\lambda_2 N_{\text{RSR}} - \gamma \sum_k k^3 \rho_{S,k}(k) + \\
\lambda_1 N_{\text{RSR}} &= \delta N_{\text{RSR}} + \gamma (2 N_{\text{RSR}} + N_{\text{RSSR}}) + w \frac{N_{\text{RSR}}}{N_S}(2 N_{\text{RS}} + N_S) + \delta (-2 N_{\text{RSR}} + N_{\text{RS}}).
\end{align*}
\]

Analyzing the steady state of this equation in the limit where \(\gamma, \delta, \mu/(N_N + N_S + N_R) \ll w, \lambda_1, \lambda_2\), and utilizing the relations in Eqs. (A2a) and (A2b), we are able to solve this equation and obtain the following relation:
\[
\frac{N_{\text{RSR}}}{N_S} = \left(\frac{N_{\text{RS}}}{N_S} - \frac{N_{\text{RN}}}{N_N}\right) \left(2 \frac{N_{\text{RS}}}{N_S} + 1\right) + \frac{N_{\text{RRN}}}{N_N}.
\]

At last, we close the \(N_{\text{RRN}}\) term by assuming that the R-nodes are distributed homogeneously in the neighborhood of N-nodes:
\[
\frac{N_{\text{RRN}}}{N_N} = \left(\frac{N_{\text{RN}}}{N_N}\right)^2 + \frac{N_{\text{RRN}}}{N_N},
\]

which finally leads to the RSR closure:
\[
\frac{N_{\text{RSR}}}{N_S} = \left(\frac{N_{\text{RS}}}{N_S}\right)^2 + \frac{N_{\text{RS}}}{N_S} + \left(\frac{N_{\text{RS}}}{N_S} \frac{N_{\text{RN}}}{N_N}\right)^2.
\]

In Fig. 3, we compare the performance of the new closure of \(N_{\text{RSR}}\) in Eq. (8) to the ad hoc, homogeneity based closure of Eq. (2b). We consider the solution of the mean-field equations, found in Appendix A, closed according to the two methods, and comparing those to the steady-state size of the recruiting class measured in the direct network simulations. Thus, in Fig. 3(a) we see that in the considered limit, i.e. when \(\lambda_1, \lambda_2\) and \(w\) are large, the mean-field closed using our approach is in a much better agreement with the simulations than the ad hoc assumption of Eq. (2b). The reason for the superior performance lies in the better approximation of the \(N_{\text{RSR}}/N_S\) term, shown in Fig. 3(b).

The appeal of this approach is evident when we test it outside of the derivation limit. In Figs. 3(c) and 3(d) we see that, when we reduce \(w\), the mean-field and the closure continue to be in a good match with the simulations. Curiously, we also note that in this limit the new closure approaches the homogeneity closure. In Fig. 3(e), as \(\lambda_1\) and \(\lambda_2\) are reduced, the mean-field solution appear to be less consistent with the simulations, which is also reflected on the closure in Fig. 3(f). Finally, in Fig. 3(g) all of the parameter are about the same order, and yet the mean-field and closure are very much consistent with the simulations.

Thus far we have shown that out method has produced a closure that is in a good match with the simulated system in steady-state. We further test...
the performance of our closure by using outside of steady-state. Figures 4(a) and 4(b) show that our closure continues to be consistent with the simulations even during the transient period.

III. ADAPTIVE EPIDEMIC MODEL

The other example that we consider is the epidemic spread on an adaptive social network [2]. In this model, the disease spread is described using the susceptible-infected-susceptible model, where each individual in the society is in one of the two states—sick or infected, and healthy, but susceptible to infection. In the framework of networks, we refer to these as I- and S-nodes respectively. The infected individuals become susceptible at rate $r$. The disease can spread at a rate $p$ per edge between a susceptible node and an infected node, and the infected nodes spontaneously become susceptible at a rate $r$. The link rewiring takes place at a rate $w$ as shown in figure (b).

$$r_3 = (1, -1), \quad r_4 = (1, 0), \quad r_5 = (-1, 0), \quad \text{and} \quad \Omega_{I,k}(r_1) = \Omega_{S,k}(r_1) = r k_1, \quad \Omega_{S,k}(r_2) = p K(k),$$

$$\Omega_{S,k}(r_3) = w k_2, \quad \Omega_{S,k}(r_4) = w N_{IS}/N_S, \quad \Omega_{I,k}(r_5) = p M(k), \quad \Omega_{I,k}(r_5) = w k_1.$$

The function $K$ (function $M$) corresponds to the expected number of node chains that originate at an S-node (I-node), which connects to an S-node, which in turn connects to an I-node. The value of $\rho_{S,k}$ (value of $\rho_{I,k}$) corresponds to the number of S-nodes (I-nodes) with $k_1$ of S-nodes and $k_2$ of I-nodes in their neighborhoods. The functions $N_{X_1, \ldots, X_n}$ are defined in the same way as in the Section II.

The mean-field equations are generated by multiplying the master equations by $k_1^{i_1} k_2^{i_2}$ and summing over $k$. Thus, two node equations are generated for $i_1 + i_2 = 0$, and three distinct link equations are generated for $i_1 + i_2 = 1$. These equations, presented in the Eqs. (B1a)-(B1e), are open, due to dependence on the terms describing to the expected number of ISI triples, $N_{ISI}$, and SSI triples, $N_{SSI}$. In order to close this system of equations, additional information is required. The usual approach is to make an assumption that infected nodes are homogeneously distributed throughout the network, an assumption that leads to the following closure:

$$\frac{N_{SSI}}{N_S} = \frac{N_{IS} N_{IS}}{N_S^2 N_S},$$

$$\frac{N_{ISI}}{N_S} = \left( \frac{N_{IS}}{N_S} \right)^2 + \frac{N_{IS}}{N_S},$$

where we make an additional assumption that the total degree distribution of susceptible nodes is Poisson.

We derive a new closure of the ISI term by considering the evolution of the sum of ISI and III triples. Adding equations (9a) and (9b), multiplying the sum by $k_2^2$ and summing over $k$ we obtain the

![Schematic representation of the coarse-graining of the nodes in the epidemic adaptive network. Figure (a) shows that the infection is transmitted at a rate $p$ per edge between a susceptible node and an infected node, and the infected nodes spontaneously become susceptible at a rate $r$. The link rewiring takes place at a rate $w$ as shown in figure (b).]
Following evolution equation:

\[
\partial_t (N_{ISI} + N_{III}) = r(-2N_{ISI} + N_{IS}) + p(2N_{SSI} + N_{SSI}) + w(-2N_{ISI} + N_{IS}) \tag{11}
\]

\[
+ r(-2N_{III} + N_{II}) + p(2N_{ISII} + N_{ISI}) + w(2N_{ISSI} + N_{ISSI}) \tag{10b}
\]

where the four-point terms ISSI and IISI correspond to total number of node configurations shown in Fig.s. 6(a) and 6(b) respectively. We solve for \(N_{ISI}/N_S\) at the steady-state and obtain the following relation:

\[
\frac{N_{ISI}}{N_S} = \frac{N_{IS}}{N_S} + \frac{N_{IS}N_{SSI}}{N_S N_{SSI}} \bigg(\frac{r}{w + r} \left(\frac{N_{II}}{N_I} - \frac{N_{II}}{N_I} \right) - \frac{N_{III}N_{II}}{N_{III}N_{II}}\bigg). \tag{12}
\]

Next we note that the term \(N_{ISI}/N_{SSI}\) corresponds to the expected number of I-nodes attached to the first S-node in the SSI structure. This relation is well approximated by the homogeneity assumption, that the information about the neighborhood of the 3rd node in Fig. 6(a) has no effect on the information about the neighborhood of the 2nd node. In other words, the following approximation is considered

\[
\frac{N_{ISI}}{N_{SSI}} = \frac{N_{SSI}}{N_{SS}} \tag{13}
\]

Note that here the homogeneity assumption is expected to be more accurate than in the Eqs. (10b) and (10b). Now, we can argue that in the limit of small \(w\) and \(p\), when the system is approaching the epidemic threshold and the infection levels are low, the expected number of ISI triples formed at an S-node is well approximated by

\[
\frac{N_{ISI}}{N_S} = \frac{N_{IS}}{N_S} + \frac{N_{IS}N_{SSI}}{N_S N_{SSI}} \tag{14}
\]

where the term proportional to \(N_I/N_S\) was dropped.

Thus, we have derived a new closure of \(N_{ISI}\) that relies on our ability to close the \(N_{SSI}\) term, and this brings us one step closer to finding an accurate closure of the system (B1). Curiously, the homogeneity closure of \(N_{SSI}\) in Eq. (10a) leads to the homogeneity closure in Eq. (10b). However, as is suggested by Figs. 7 and 8, improving closure of \(N_{SSI}\) leads to the improvement of \(N_{ISI}\) closure. In fact, Figs 7(b), 7(d), 7(f) as well as 8(b), 8(d), 8(f) show the relative deviation of the closure relations from the approximated quantity and suggest that the new approximation, where the measured value of \(N_{SSI}\) is used, is superior to the relation in Eq. (10b). Note that the only time the homogeneity closure appears to perform better is when it intersects the measured value of \(N_{ISI}/N_S\), and, therefore, its superiority over the performance of the new closure is rather coincidental. Further consideration of the results in Fig. 7
shows that, as we move away from the derivation regime of low \( w \) values, the performance of the closure diminishes, though it is still superior to the old approximation. Surprisingly, as shown in Fig. 8, the performance of the new approximation improves for the larger values of \( p \), i.e. as \( N_I \) is increased and as we move further away from the bifurcation, where the derivation was made.

Finally, we test our new closure outside of the steady-state. Thus, Fig. 9(a) compares the performance of the newly derived approximation to that of the homogeneity approximation, and shows that, unlike the homogeneity closure, the new closure follows the measured values of \( N_{ISI}/N_S \) very accurately. Furthermore, as shown in Fig. 9(b), the closure of Eq. (14) performs better as we approach the steady-state, while the opposite is observed for the Eq. (10b) closure.

FIG. 8. Number of ISI triples per S-node as a function of infection rate, for several rewiring rates: simulations compared to two approximations. The curves and circles are defined as in Fig. 7, with the same network size and number of links.

FIG. 9. Time evolution of ISI closure. Fig. (a) shows the time evolution of the number of ISI triples per S-node (circles, red online) and the approximate value as obtained from the relation in Eq. (10b) (light curve, green online) and from Eq. (14) (dark curve, black online). Fig. (b) shows the relative error due to the two approximations. The simulations are performed for \( w = 10^{-1} \), \( p = 10^0, 10^5 \) nodes and \( 5 \times 10^5 \) links.

IV. DISCUSSION

We presented an approach of closing mean-field description of adaptive social networks. In our approach we proposed exploiting the possible simplification of master equations in some asymptotic regime. We applied this approach to two examples of adaptive networks: recruitment to a cause model and a model of epidemic spread on adaptive network. Using the two examples, we successfully developed closures that perform as well as or better than the usual, homogeneous distributions of nodes based, closures.

The closure we developed for the recruitment model showed significant improvement of the mean-field description from the one where all of the high order terms were approximated using the homogeneity closure. We also showed that for some other terms, the homogeneity closure may in fact lead to an accurate approximation.

In case of epidemic model, our approach resulted in an improved closure, contingent on improvements of a closure for another term. Therefore, we were unable to solve the newly closed system. However, the results of the simulations clearly showed that the new closure more accurately reflected the relation between the measured terms.

It is very important to note that the closures that we derived in some asymptotic regimes proved to be more accurate than the homogeneity closures even outside of the derivation limit. For example, even though in both cases the closures were derived at steady-state, they showed excellent results during the transient state of the system as well. The additional benefit from using this approach is that it allows one to deduce the level of error introduced by
the derived closure. In other words, we can expect that the closures will perform well in the derived regimes, and be less reliable away from it.

The developed method can be further applied to a more general class of networks problems.

Appendix A: Recruitment with Adaptation

Mean-Field Equations

\[
\begin{align*}
\partial_t N &= \mu - \lambda_1 N + \lambda_2 S - \delta N, \\
\partial_t S &= \lambda_1 N - \lambda_2 S - \gamma_1 R - \delta S, \\
\partial_t R &= \gamma_1 R - \delta R, \\
\partial_t N_{SN} &= 2\sigma \mu \frac{N}{N + S + N_R} + 2\lambda_2 N_{SN} - 2(\lambda_1 + \delta) N_{SN}, \\
\partial_t N_{SN} &= \mu \frac{N}{N + S + N_R} + \lambda_2 N_{SS} - \gamma N_{NSR} - (\lambda_1 + \lambda_2) N_{SN} + \lambda_1 N_{NN}, \\
\partial_t N_{SS} &= -2\gamma N_{SS} + 2\lambda_1 N_{SN} - 2(\lambda_2 + \delta) N_{SS}, \\
\partial_t N_{RN} &= \mu \frac{N}{N + S + N_R} + \gamma N_{NSR} - (\lambda_1 + 2\delta + w) N_{RN} + \lambda_2 N_{RS}, \\
\partial_t N_{RS} &= \gamma N_{NSR} - \gamma N_{SSR} - (\lambda_2 + 2\delta) N_{RS} + (\lambda_1 + w) N_{RN}, \\
\partial_t N_{RR} &= 2\gamma N_{RSS} - 2\delta N_{RR}.
\end{align*}
\]

At steady state, in the limit where \(\gamma, \delta, \mu/(N_N + N_S + N_R) \ll w, \lambda_1, \lambda_2\), the node and link equations lead to the following relations:

\[\begin{align*}
\lambda_1 N_N &= \lambda_2 N_S, \\
(\lambda_1 + w) N_{RN} &= \lambda_2 N_{RS}.
\end{align*}\] (A2a, A2b)

Appendix B: SIS with Adaptation Mean-Field Equations

The mean-field equations generated from the master equations (9a) and (9b) are

\[\begin{align*}
\partial_t N &= r N_I - p N_{IS}, \\
\partial_t I &= -r N_I + p N_{IS}, \\
\partial_t S &= -2p N_{SSI} + 2(r + w) N_{IS}, \\
\partial_t N_{IS} &= r N_{II} - p N_{IS} + p N_{SSI} - (r + w) N_{IS}, \\
\partial_t N_{II} &= -2r N_{II} + 2p N_{IS}.
\end{align*}\] (B1a, B1b, B1c, B1d, B1e)

with the initial conditions satisfying the conservation of nodes and links equations:

\[\begin{align*}
N_S + N_I &= N, \\
N_{SS} + N_{IS} + N_{II} &= N \sigma,
\end{align*}\] (B2a, B2b)

where \(N\) is the total number of nodes, and \(\sigma\) is the mean degree. At steady-state the mean-field equations give the following important relations used in the body of the paper:

\[\begin{align*}
p N_{IS} &= (r + w) N_{IS}, \\
p N_{IS} &= r N_{II}.
\end{align*}\] (B3a, B3b)