Exploring network dynamics with a mathematical triple jump

Holly Silk*, Güven Demirel†, Martin Homer*, and Thilo Gross*
* *University of Bristol, Department of Engineering Mathematics and Bristol Centre for Complexity Sciences, Bristol, UK.* † *Max-Planck-Institute for the Physics of Complex Systems, Nöthnitzer Str. 38, Dresden, Germany*
(Dated: February 13, 2013)

In the investigation of network dynamics a prominent role is played by analytical coarse-graining schemes. A very powerful class of these schemes are the so-called heterogeneous approximations, which describe the dynamics of an agent-based network model by an infinite-dimensional system of differential equations. Here, we propose a widely applicable method for solving these systems, and illustrate it in the example of the adaptive voter model.

In the triple jump an athlete attempts to cross a distance by a sequence of three jumps: *hop*, *step*, and *jump*, that each employ different techniques. Here, we address the challenge of analyzing a dynamical network model with a mathematical triple jump. First using a *heterogeneous moment expansion* we approximate an agent-based model by a high-dimensional system of ordinary differential equations (ODEs). Then, using *generating functions* we convert the high-dimensional ODEs into a low-dimensional system of partial differential equations (PDEs). Finally, using the *method of characteristics* we convert the PDE into a low-dimensional set of ODEs that can be solved analytically.

In the mathematical exploration of a network model, finding a suitable approach that yields mathematically tractable equations is often a central challenge. One prominent method for approximating network dynamics are moment expansions [1–6]. The central idea is to write evolution equations for the abundance of certain motifs in the network. Different expansions can be distinguished by the the basis of motifs that they use. The most basic type of approximation, the homogeneous mean-field approximations, only track the abundance of motifs consisting of single nodes, such that they yield equations describing, for example in epidemic spreading, the abundance of infected and susceptible agents. More sophisticated approximations, such as the pair approximation [1–3] or triplet approximation [5, 6] track also the abundance of larger motifs, such as linked pairs or triplets of nodes.

A powerful class of approximations that have been recently proposed [7, 8] are the *heterogeneous approximations*. In these approximations the definition of a certain motif prescribes not only the state and connectivity of the network nodes in the motif, but also their outgoing links. For instance in the epidemic context one distinct motif could be a susceptible node connected to exactly two other susceptible nodes and exactly one infected node.

Heterogeneous approximations have been shown to yield excellent results in examples. However, they typically lead to infinite dimensional systems of ODEs. In practice, the number of equations is limited by a truncation, but the number of ODEs that are studied is still often of the order of 10^6. They are thus typically studied by numerical integration of these high-dimensional ODE systems.

Here, we use *generating functions* to map the infinite-dimensional ODE system from a heterogeneous active neighborhood approximation to a PDE. This mapping is exact and reversible and thus does not involve additional assumptions. Generating functions are a major tool of discrete mathematics and as such have been applied to network problems. However, they are typically used to capture the structure rather than the dynamics of networks [9]. For the present context, relevant work includes [10] where generating functions were used to explore specific processes in the dynamics of an adaptive epidemic model. Here, we solve the PDE analytically by the *method of characteristics* [11]. The application of characteristics constitutes another exact reversible transformation, so that no additional approximation is necessary. The triple jump approach can thus reveal full time-dependent solutions that describe the dynamics of network to a high degree of accuracy.

We illustrate the triple jump approach by considering the adaptive voter model [5, 12–16]. This extension of the seminal voter model [17] has been studied with a wide variety of techniques [6], and can thus be considered as a benchmark system. The model provides a "hard" example, where most approximations do not perform well [14]. In particular, it was already noted in [6] that this system is one of the rare cases where the heterogeneous approximation does not perform particularly well. We therefore do not expect the results of this particular application of the triple jump methodology to yield a very close match with those of agent-based models. We have nevertheless chosen to focus on the adaptive voter model as it allows for a concise presentation of the methodology.

Our main intention is to show that the systems of equations obtained from the heterogeneous approximation can be solved analytically using generating functions and characteristics. We believe that this triple jump approach will be valuable for the wide variety of models in which the heterogeneous approximation affords an almost exact description of the system.

ADAPTIVE VOTER MODEL

We consider a network of N nodes, representing agents, and K bidirectional links, representing social contacts. Each agent i, is associated with a binary variable $s_i \in \{A, B\}$ representing the agent's opinion. We initialize the network of agents as an Erdős-Rényi random graph and assign opinions to the agents randomly with equal probability. The network is then evolved in time by consecutive update steps. In each step a link connecting two agents i and j is selected randomly. If the agents hold identical opinions, $s_i = s_j$, then the link is said to be *inert* and nothing happens. If $s_i \neq s_j$, then the link is said to be *active* and one agent, $a \in \{i, j\}$, is chosen with probability 1/2 to resolve the conflict. With probability p, the link connecting i and j is cut and a establishes a new link to a randomly chosen agent, k with $s_k = s_a$. Otherwise, with probability $1 - p$, agent a changes its opinion such that $s_i = s_j$. We say that in the former case the conflict is resolved by a *rewiring event*, and in the latter case by an *opinion adoption event*.

One can easily verify that the rules of the adaptive voter model are unbiased, such that there is no net drift in the number of nodes holding a particular opinion. Thus, the fraction of nodes holding opinion A, will remain constant in time except for stochastic fluctuations that become negligible in the thermodynamic limit $N \to \infty$ (with constant K/N). For simplicity, we can thus focus on the symmetric case where the number of nodes holding opinions A and B are equal.

HOP: HETEROGENEOUS MOMENT EXPANSION

In the first step, we convert the stochastic agent-based model into an infinite-dimensional system of ordinary differential equations (ODEs) by a heterogeneous expansion, known as the active neighborhood approximation. Following [7] we define $A_{k,l}$ to be the normalized number of agents of opinion A who have l active and and k inert links. In the thermodynamic limit we can treat the $A_{k,l}$ as continuous variables and capture their dynamics by the differential equations

$$\frac{\mathrm{d}A_{k,l}}{\mathrm{d}t} = \frac{\bar{p}}{2}\left[kA_{l,k} - lA_{k,l}\right] \tag{1i}$$
$$+ \frac{\bar{p}}{2}\left[(l+1)A_{k-1,l+1} - lA_{k,l}\right] \tag{1ii}$$
$$+ \frac{\bar{p}}{2}\frac{\sum_{k,l}(l-1)lA_{k,l}}{\sum_{k,l} lA_{kl}}\left[(l+1)A_{k-1,l+1} - lA_{k,l}\right] \tag{1iii}$$
$$+ \frac{\bar{p}}{2}\frac{\sum_{k,l} lkA_{k,l}}{\sum_{k,l} kA_{kl}}\left[(k+1)A_{k+1,l-1} - kA_{k,l}\right] \tag{1iv}$$
$$+ \frac{p}{2}\left[(l+1)A_{k-1,l+1} - lA_{k,l}\right] \tag{1v}$$
$$+ \frac{p}{2}\left[(l+1)A_{k,l+1} - lA_{k,l}\right] \tag{1vi}$$
$$+ \frac{p}{2}\frac{\sum_{k,l} lA_{k,l}}{\sum_{k,l} A_{k,l}}\left[A_{k-1,l} - A_{k,l}\right], \tag{1vii}$$

where $\bar{p} = (1 - p)$. The terms in Eq. (1) describe the change experienced by a focal node $A_{k,l}$ due to opinion adoption ((i) to (iv)) and rewiring ((v) to (vii)) events. Specifically, contributions arise from (i) adopting the opinion of a neighbor of type B, (ii) a neighbor of type B adopting the opinion of the focal node, (iii) a neighbor of type B adopting the opinion of another node of type A, (iv) a neighbor of type A adopting the opinion of another node of type B, (v) the focal node rewiring one of its links away from a neighbor of type B (acquiring a new neighbor of type A), (vi) a neighbor of type B rewiring a link away from the focal node, and (vii) a node of type A rewiring one of its links to the focal node.

The events (iii), (iv), and (vii) involve nodes outside the direct neighborhood of the focal node. They are thus dependent on the number of next-nearest neighbors (iii, iv) or active links existing elsewhere (vii). The corresponding rates then depend on longer-ranged correlations that are not captured by the $A_{k,l}$ alone. We therefore need to estimate these rates based on the available information from the nearest-neighbor correlations captured. This approximation is called moment closure, and is known to be the main source of inaccuracy in moment expansions for networks [6].

For instance, the rate (iv) at which a typical neighbor of type A (A-neighbor) of the focal node adopts the opinion B depends on the average number of B-neighbors of the A-neighbor, and thus on the next-nearest-neighborhood of the focal node. To approximate this number, one considers all potential A-neighbors, based on the known distribution $A_{k,l}$, but takes into account that a node that has k links to other A-nodes is k times more likely to be an A-neighbor of the focal node. The distribution of A-neighbors of the focal node is thus $kA_{k,l}/(\sum kA_{kl})$, where the denominator arises from normalization. Based on this distribution we can then estimate the number of B-neighbors of a typical A-neighbor of the focal node as $(\sum lkA_{k,l})/(\sum kA_{kl})$, which appears as a factor in the corresponding term (iv). Finally, to estimate the typical density of A-neighbors of a B-neighbor of the focal node we exploit the symmetry of the system, which implies $A_{k,l} = B_{k,l}$.

STEP: GENERATING FUNCTIONS

The heterogeneous expansion results in an infinite system of ODEs, which we transform into a single scalar PDE by use of generating functions [18]. We start by defining the generating function $Q(x, y, t) = \sum_{k,l} A_{k,l}(t)x^k y^l$, where x and y are abstract spatial variables that do not have any physical interpretation.

The time evolution of Q is given by

$$\frac{\partial Q}{\partial t} = Q_t = \sum_{k,l} \frac{\mathrm{d}A_{k,l}}{\mathrm{d}t} x^k y^l. \qquad (2)$$

Substituting Eq. (1) into this expression results in a number of terms on the right-hand-side that can be interpreted as functions of Q and its derivatives, for which we use the notation $Q_x = \partial Q/\partial x$. For instance, the process (ii) described above results in a term proportional to

$$\sum_{k,l} \left[(l+1)A_{k-1,l+1} - lA_{kl}\right] x^k y^l = xQ_y - yQ_y, \qquad (3)$$

where we separated the two terms in the square bracket and used a shift of indices $k-1 \to k$, $l+1 \to l$ on the first. Proceeding analogously, we find that Q satisfies

$$Q_t = \frac{\bar{p}\beta}{2}(y-x)Q_x + \left[\frac{\bar{p}}{2}(x-y) + \frac{\bar{p}\alpha}{2}(x-y) + \frac{p}{2}(x-y) + \frac{p}{2}(1-y)\right] Q_y + \frac{p\gamma}{2}Q(x-1), \qquad (4)$$

where

$$\alpha = \frac{Q_{yy}(1,1)}{Q_y(1,1)},\ \beta = \frac{Q_{xy}(1,1)}{Q_x(1,1},\ \gamma = \frac{Q_y(1,1)}{Q(1,1)} \qquad (5)$$

are the transformed factors from the moment closure approximation. These factors are unknown, but can be found by demanding self-consistency. Similarly, transforming the initial condition yields, at $t = 0$, $Q = \mathrm{e}^z$ with $z = \langle k\rangle(x+y-2)/2$.

JUMP: CHARACTERISTICS

The generating function PDE (4) is a first-order scalar quasilinear equation, of the form

$$\sum_{i=1}^{m} a_i(x_1,\ldots,x_m,u)\frac{\partial u}{\partial x_i} = c(x_1,\ldots,x_m,u) \qquad (6)$$

Such PDEs can be solved by the method of characteristics [11]. The central idea of this method is to describe the solution surface parametrically, and hence capture the dynamics with a low-dimensional set of ODEs in $(x_1,\ldots,x_m,u)$-space: the *bicharacteristic equations*

$$\frac{\mathrm{d}x_i}{\mathrm{d}\eta} = a_i, \quad \text{and} \quad \frac{\mathrm{d}u}{\mathrm{d}\eta} = c, \qquad (7)$$

with the initial condition $u = u_0(\xi_1,\ldots,\xi_{m-1})$, $x_i = x_{i0}(\xi_1,\ldots,\xi_{m-1})$ at $\eta = \eta_0$.

The bicharacteristic equations for Eq. (4) are

$$\frac{\mathrm{d}t}{\mathrm{d}\eta} = 1 \qquad (8)$$

$$\frac{\mathrm{d}x}{\mathrm{d}\eta} = \frac{\bar{p}\beta}{2}(x-y) \qquad (9)$$

$$\frac{\mathrm{d}y}{\mathrm{d}\eta} = \frac{\bar{p}(1+\alpha)}{2}(y-x) + \frac{p}{2}(y-1) \qquad (10)$$

$$\frac{\mathrm{d}Q}{\mathrm{d}\eta} = \frac{p\gamma}{2}(x-1)Q. \qquad (11)$$

with $t = 0$, $x = \xi_1$, $y = \xi_2$, $Q = \exp\{\langle k\rangle(\xi_1+\xi_2-2)/2\}$ at $\eta = 0$. For simplicity, we assume that α, β, γ are constant; while this may not be true in general, it is true close to stationary states, as the longer-ranged correlations in adaptive networks equilibrate in general faster than shorter-ranged ones. Solving Eqs. (8)–(11) we find

$$Q = \exp\left\{\frac{\langle k\rangle}{2}(\xi_1+\xi_2-2) - \frac{p\gamma}{2|\Lambda|}\left[\frac{v_1^1\left(v_2^2(\xi_1-1) - v_2^1(\xi_2-1)\right)}{\lambda_1}\left(\mathrm{e}^{\lambda_1 t}-1\right) + \frac{v_2^1\left(v_1^1(\xi_2-1) - v_1^2(\xi_1-1)\right)}{\lambda_2}\left(\mathrm{e}^{\lambda_2 t}-1\right)\right]\right\}, \qquad (12)$$

where $\lambda_{1,2}$ and $v_{1,2}^{1,2}$ are eigenvalues and eigenvector components of the homogeneous linear operator defined by Eqs. (9) and (10), which relate (x,y) to (ξ_1,ξ_2).

Figs. 1 and 2 show the analytic solution, Eq. (12), compared with numerical integration of the high-dimensional ODE system, Eq. (1). They illustrate the well-known dynamics of the adaptive voter model. If the rewiring rate p is sufficiently low then the system approaches an active state that is stable in the thermodynamic limit [13]. If the rewiring rate exceeds a threshold p_c, however, the system rapidly approaches a fragmented state, where the network breaks into two disconnected components that hold different opinions, but are internally in consensus. Further, the figures show that the analytic and numeric results are in good agreement. The minor discrepancies are likely due to finite cut-off in the numerics, and (in Fig. 1) due to the assumption of constant α, β, and γ.

Two conclusions can immediately be drawn from the analytic solution Eq. (12). First, the consistency conditions Eq. (5) imply that $\alpha = \beta = \gamma$. This is interesting because it shows that correlations beyond the nearest neighbors are not captured by the approximation, i.e. the probability of finding an active link attached to a node is independent of the other neighbors of the node. This failure to capture longer-ranged correlations was already suspected to be the main source of error in the heterogeneous approximation of the adaptive voter model in [6]. Second, we can analytically determine the critical rewiring rate, p_c. As $t \to \infty$, $Q_x(1,1) =$

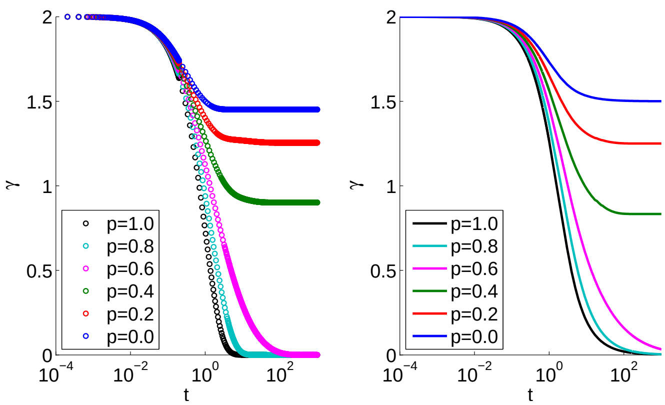

FIG. 1. (Color Online) Timeseries. Number of active links γ as a function of time t, for different rewiring rates p. Shown is a comparison of numerical integration of the high-dimensional ODE (left) and triple jump analytic solution (right).

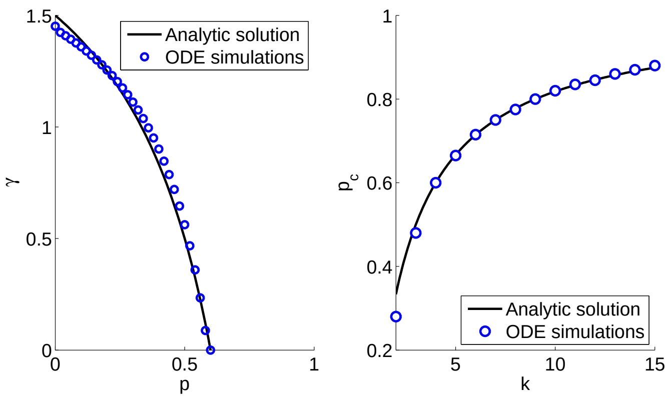

FIG. 2. (Color Online) Fragmentation transition. Shown are number of active links γ as a function of the rewiring rate p (left), and critical rewiring rate p_c as a function of mean degree k. Both exhibit a good agreement between numerical (points) and analytic solutions (lines).

$\gamma + (1+p)/(1-p)$, and by definition $Q_y(1,1) = \gamma$. Substituting these results into $Q_x(1,1)+Q_y(1,1) = \langle k\rangle$ gives $\alpha = \beta = \gamma = [\langle k\rangle - (1+p)/(1-p)]/2$. At the fragmentation transition the active links vanish, thus $\gamma = 0$ and $p_c = (\langle k\rangle -1)/(\langle k\rangle +1)$. Because the longer-ranged correlations are not captured this result is only qualitatively correct. However, it performs better than simpler approximations [6], with the exception of [14] which is only applicable close to the fragmentation point.

CONCLUSIONS

In this paper we have proposed an approach for the investigation of network dynamics that combines heterogeneous expansions, generating functions, and the method of characteristics. Using this mathematical triple jump, analytic solutions for heterogeneous expansions can be obtained, which we demonstrated on the example of the adaptive voter model.

The triple jump approach does not involve any critical assumptions, other than that inherent in the heterogeneous expansion. For the adaptive voter model studied here, it is known that this provides only qualitative results [6]. However, heterogeneous expansions provide an excellent approximation in other models [7, 8].

Here, we focused on the adaptive voter model because its symmetry allows for a concise presentation of the triple jump methodology. When applying the approach to other models it will be necessary to write equations for multiple states. One generally has the choice between representing states as a variable ($B_{i,j}$ in addition to $A_{i,j}$, say), or as an extra index ($A_{i,j,k}$ instead of $A_{i,j}$). The latter choice may be advantageous as it leads to a single scalar PDE in a three-dimensional space, rather than two coupled PDEs in a two-dimensional space. Depending on the system the method of characteristics may not be appropriate, for the final step of the triple jump. However, in any case this step profits from well-developed PDE theory, such that analytic solutions should in many cases be possible.

[1] K. T. D. Eames and M. J. Keeling, PNAS **99**, 13330 (2002).
[2] C. T. Bauch, Math. Biosci. **198**, 237 (2005).
[3] T. Gross, C. Dommar D‘Lima, and B. Blasius, Phys. Rev. Lett. **96**, 208701 (2006).
[4] T. Rogers, J. Stat. Mech **1**, P05007 (2011).
[5] D. Kimura and Y. Hayakawa, Phys. Rev. E **78**, 016103 (2008).
[6] G. Demirel, F. Vazquez, G. A. Böhme, and T. Gross, arXiv:1211.0449 (2012).
[7] V. Marceau, P. A. Noël, L. Hébert-Dufresne, A. Allard, and L. J. Dubé, Phys. Rev. E **82**, 036116 (2010).
[8] J. P. Gleeson, Phys. Rev. Lett. **107**, 068701 (2011).
[9] M. E. J. Newman, *Networks: An Introduction* (Oxford University Press, 2010).
[10] S. Wieland, T. Aquino, and A. Nunes., Europhys. Lett. **97**, 18003 (2012).
[11] J. Ockendon, S. Howison, A. Lacey, and A. Movchan, *Applied Partial Differential Equations* (Oxford University Press, 2003).
[12] P. Holme and M. E. J. Newman, Phys. Rev. E **74**, 056108 (2006).
[13] F. Vazquez, V. M. Eguíluz, and M. San Miguel, Phys. Rev. Lett. **100**, 108702 (2008).
[14] G. A. Böhme and T. Gross, Phys. Rev. E **83**, 35101 (2011).
[15] G. A. Böhme and T. Gross, Phys. Rev. E **85**, 066117 (2012).
[16] C. Huepe, G. Zschaler, A. L. Do, and T. Gross, NJP **13**, 073022 (2011).
[17] R. A. Holley and T. M. Liggett, Ann. Prob. , 643 (1975).
[18] H. S. Wilf, *Generatingfunctionology* (A K Peters, 2006).