Pathogen Mutation Modeled by Competition Between Site and Bond Percolation

Laurent Hébert-Dufresne, Oscar Patterson-Lomba, Georg M. Goerg, and Benjamin M. Althouse

Pathogen mutation modeled by competition between site and bond percolation

Laurent Hébert-Dufresne,[1] Oscar Patterson-Lomba,[2] Georg M. Goerg,[3] and Benjamin M. Althouse[4]

[1]*Département de Physique, de Génie Physique, et d'Optique,*
Université Laval, Québec (Québec), Canada G1V 0A6
[2]*Mathematical, Computational, and Modeling Sciences Center,*
School of Human Evolution and Social Change, Arizona State University, Tempe, AZ, 85287
[3]*Department of Statistics, Carnegie Mellon University, Pittsburgh, PA, 15213*
[4]*Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, 21205*

While disease propagation is a main focus of network science, its coevolution with treatment has yet to be studied in this framework. We present mean-field and stochastic analysis of an epidemic model with antiviral administration and resistance development. We show how this model maps to a *coevolutive competition between site and bond percolation* featuring hysteresis and both second and first-order phase transitions. The latter, whose existence on networks is a long-standing question, imply that a *microscopic* change in infection rate can lead to *macroscopic* jumps in expected epidemic size.

With the recent focus of public health policies on planning the control of the next influenza pandemic [1], more complex models have been introduced in epidemiology [2, 3]. We extend on one of these studies [2] where treatment of influenza, as a selection pressure, favors the emergence and spread of pathogen strains with a drug-resistant phenotype. However, very similar adaptation dynamics could also be considered in the interactions of pathogens through ecological mechanisms [4], or of adaptive computer viruses [5, 6] and for behavioral changes in a population [7, 8] or ecosystem [9]. While we study a mutation dynamics, the terms adaptation and coevolution are not used as biological concepts, but simply in reference to dynamics where two variables influence one another.

Our model consists of a contact network where each individuals can be in one of five states: susceptible (S), infectious and untreated (I_u), infectious and treated (I_t), infectious with a resistant strain (I_r), or recovered (R). The dynamics then obey the following rules:

- a link from I_x to S leads to an infection at rate β_x ($x \in \{u, t, r\}$);

- a wild strain infection (through I_u or I_t) is untreated ($S \to I_u$) with probability $1 - \rho$, or treated with probability ρ;

- treatment is effective ($S \to I_t$) with probability $1 - c$, or leads to mutation ($S \to I_r$) with probability c;

- a resistant strain infection (through I_r) can only transmit this strain ($S \to I_r$);

- infectious individuals of type I_x recover at rate γ_x.

Once all infectious individuals have recovered, the final epidemic size is calculated.

Mean-field analysis. One of the benefits of network modeling resides in the possibility to account for heterogeneity in the contact structure of a population. Hence, we consider both delta and fat-tailed distributions of links per node (or degree distribution), to create homogeneous and heterogeneous networks. The distributions are detailed in the Supplemental Material (SM). However, to accurately follow such heterogeneity in a mean-field analysis, one must distinguish nodes not only by their states, but also by their degree [6]. For instance, the mean fraction of susceptible nodes of degree k at time t, $S_k(t)$ can be written as:

$$\dot{S}_k = -k(\beta_u \langle I_u \rangle + \beta_t \langle I_t \rangle + \beta_r \langle I_r \rangle) S_k \tag{1}$$

where $\langle I_x \rangle$ is the probability that a randomly chosen link of a susceptible node leads to an infectious individual of type x. Note that all time dependency are implicit. Similarly for other node states, we can deduce:

$$\dot{I}_{u,k} = k(\beta_u \langle I_u \rangle + \beta_t \langle I_t \rangle)(1 - \rho) S_k - \gamma_u I_{u,k} \tag{2}$$

$$\dot{I}_{t,k} = k(\beta_u \langle I_u \rangle + \beta_t \langle I_t \rangle)\rho(1 - c) S_k - \gamma_t I_{t,k} \tag{3}$$

$$\dot{I}_{r,k} = k(\beta_u \langle I_u \rangle + \beta_t \langle I_t \rangle)\rho c S_k + k\beta_r \langle I_r \rangle S_k - \gamma_r I_{r,k} \tag{4}$$

$$\dot{R} = \sum_k \gamma_u I_{u,k} + \gamma_t I_{t,k} + \gamma_r I_{r,k} . \tag{5}$$

We must be careful in evaluating the mean-field quantities $\langle I_x \rangle$ as a susceptible is less likely to be connected to an infectious node than, for example, a recently infected node. To account for such correlations [10], we follow the density of each possible link attached to at least one susceptible node (denoted $[SX]$):

$$[\dot{SS}] = -2(\beta_u \langle I_u \rangle + \beta_t \langle I_t \rangle + \beta_r \langle I_r \rangle)\langle k'_s \rangle [SS] \tag{6}$$

$$[\dot{SI_u}] = -\left[(\beta_u \langle I_u \rangle + \beta_t \langle I_t \rangle + \beta_r \langle I_r \rangle)\langle k'_s \rangle + \beta_u + \gamma_u\right][SI_u] + 2(\beta_u \langle I_u \rangle + \beta_t \langle I_t \rangle)\langle k'_s \rangle (1 - \rho)[SS] \tag{7}$$

$$[\dot{SI_t}] = -\left[(\beta_u \langle I_u \rangle + \beta_t \langle I_t \rangle + \beta_r \langle I_r \rangle)\langle k'_s \rangle + \beta_t + \gamma_t\right][SI_t] + 2(\beta_u \langle I_u \rangle + \beta_t \langle I_t \rangle)\langle k'_s \rangle \rho(1 - c)[SS] \tag{8}$$

$$[\dot{SI_r}] = -\left[(\beta_u \langle I_u \rangle + \beta_t \langle I_t \rangle + \beta_r \langle I_r \rangle)\langle k'_s \rangle + \beta_r + \gamma_r\right][SI_r] + 2(\beta_u \langle I_u \rangle + \beta_t \langle I_t \rangle)\langle k'_s \rangle \rho c[SS] + 2\beta_r \langle I_r \rangle \langle k'_s \rangle [SS] \tag{9}$$

$$[\dot{SR}] = -(\beta_u \langle I_u \rangle + \beta_t \langle I_t \rangle + \beta_r \langle I_r \rangle)\langle k'_s \rangle [SR] + \gamma_u[SI_u] + \gamma_t[SI_t] + \gamma_r[SI_r] \tag{10}$$
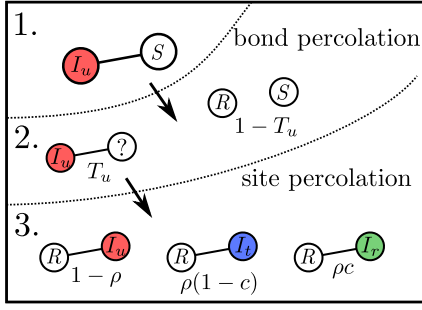
FIG. 1. (Color online) **Competitive co-evolution between site and bond percolation.** The percolative process of an $[SI_u]$ link is designed to be equivalent to the continuous time dynamics; 1. indicates the initial state; 2. bond percolation, formation of links (infection) with probability T_u; 3. three-state site percolation for treatment (to untreated, treated or mutated). Events involving I_t nodes use bond percolation with transmissibility T_t followed by site percolation as illustrated here, whereas events involving I_r use solely bond percolation with probability T_r (no possible treatment, hence no site percolation).

where $\langle k'_s \rangle$ is the average excess degree of susceptible nodes. Equations (1-10) represent the minimal set of equations required to describe the system at all time, in the sense that they are sufficient to calculate all the mean-field quantities on which they depend. A simple averaging procedure yields:

$$\langle k'_s \rangle = \frac{\sum_k k(k-1)S_k}{\sum_k kS_k} \tag{11}$$

$$\langle I_x \rangle = \frac{[SI_x]}{2[SS] + [SI_u] + [SI_t] + [SI_r] + [SR]} . \tag{12}$$

Integrating this system of equation provides mean-field predictions (i.e. in the infinite limit) for the final size of epidemics.

Mapping to percolation. Most SIR models feature an irreversible timeline. For our model, there are only four possible scenarios for each node: $S \rightarrow I_u \rightarrow R$, $S \rightarrow I_t \rightarrow R$, $S \rightarrow I_r \rightarrow R$, or S for all time, and thus none of these scenarios can be traveled in reverse. This implies that the considered continuous time model can be mapped unto a percolation process [11–13], or more precisely, a coevolutive competition arises between site and bond percolation. The bond percolation represents the propagation of the disease under certain assumptions [14, 15], see SM for details, while the site percolation represents both treatment and mutation (Figure 1). As we will see, these dynamics are both coevolutive (the disease mutates to adapt and resist treatment) and competitive (treatment aims to stop bond percolation, and the two strains can hinder each other's propagation). The details of this particular process are illustrated in Figure 1.

The different percolation probabilities involved can be easily evaluated. In fact, treatment and mutation are already modeled as site percolation in the original dynamics. Infection events on an $[SI_x]$ link are equivalent to a bond percolation

process where infection occurs with a given probability T_x, i.e. that the infection event precedes the recovery event (proof in SM):

$$T_x = \frac{\beta_x}{\beta_x + \gamma_x} . \tag{13}$$

Also note that while the infections map to classic boolean bond percolation, the treatment process maps to site percolation with three possible states (if $0 < \rho < 1$) akin to the 3-states Potts model [16].

Finally, while resistance will always emerge under mean-field assumption, one can account for this by approximating the probability of emergence of the resistant strain through the probability of treatment causing at least one mutation. The expected number of infections caused by a single infectious individual from a disease under its epidemic threshold, $\langle n \rangle$, is a well-known result of network epidemiology [11] and can be used to calculate the probability P of emergence of resistance. In $\langle n \rangle$ infections, resistance develops only if at least one leads to a failed treatment (probability ρc):

$$P = 1 - (1 - \rho c)^{\langle n \rangle} = 1 - (1 - \rho c)^{T\langle k \rangle / (1 - T\langle k' \rangle)} \tag{14}$$

where $\langle k \rangle$ and $\langle k' \rangle$ are the mean degree and excess degree of the network, respectively, and T is the effective transmissibility of the treated wild-strain. A more complete analysis is given in the SM.

Phase transition. Our model can lead to four possible final states: a disease-free state, epidemics caused by either the wild strain ($c = 0$), the resistant strain ($c > 0$), or a combination of both (if above their respective threshold). In standard epidemic and percolation models, the transition from the disease-free equilibrium to an epidemic is observed by keeping all parameters constant and progressively raising the transmissibility. Once the epidemic threshold is achieved, the disease is able to spread to an increasingly larger macroscopic fraction of the network [11].

To highlight certain features, we consider the case $\beta_r > \beta_u$ – corresponding biologically to the development of compensatory mutations in the pathogen in response to the fitness cost typically associated with treatment resistance [17, 18] – and set $\gamma_u = \gamma_t = \gamma_r = \gamma$ for simplicity. We note that while compensatory mutations are rare, the selective pressures exerted by treatment can still give a large advantage to uncompensated resistant strains. In fact, our results are qualitatively similar with or without these mutations as long as the resistant-strain epidemic undergoes a phase transition before the treated wild-type strain ($\beta_r > \beta_t$).

The phase transition from the disease free state to the epidemic state, dominated mostly by the resistant strain, is demonstrated in Figure 2. The main feature is the explosive transition, where the observed maximal epidemic size jumps suddenly from zero to almost 10%. While the transition is technically of the second-order (i.e. continuous) the probability of resistance emergence falls to zero when β_u diminishes such that some epidemic sizes are practically impossible to
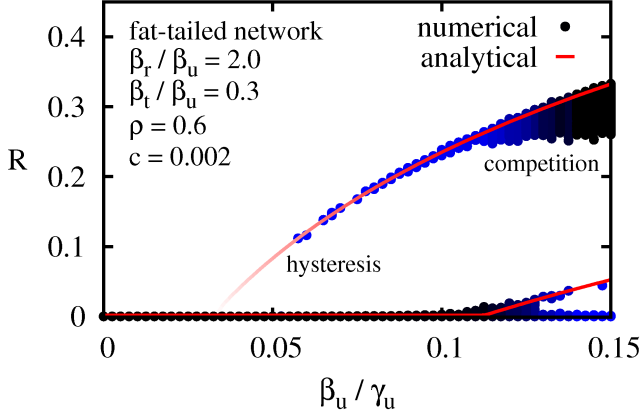
FIG. 2. (Color online) **Explosive phase transition.** Emergence of the giant component (i.e. of epidemics) as infection rates increase. The results for over 10^6 simulations of the percolation process on fat-tailed networks with 2.5×10^5 nodes are plotted (points). Point color represents the proportion of cases landing on either branch (lighter = less likely, darker = more likely). This coloring highlights the second transition, or invasive threshold, which marks the end of bistability. Analytical curves are obtained by integrating our equations with ($c > 0$) and without ($c = 0$) mutation for upper and lower branches respectively. The color code of the upper branch corresponds to $\log P$ (log of the probability of resistance emergence) and encodes the probability of reaching that branch: likely in color, unlikely in white.
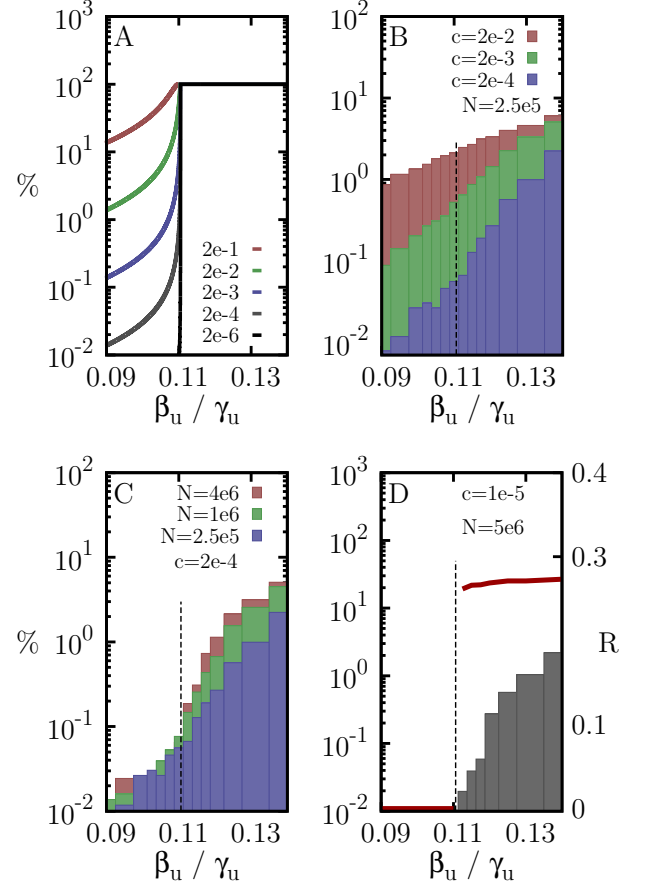

FIG. 3. (Color online) **From continuous to discontinuous epidemics.** (A) Eq. (14) for the probability of resistance emergence for various mutation probability c. As c goes to zero, the transition at threshold converges toward a step function leading to a discontinuity in possible epidemic size. (B) Probability of reaching a resistant strain epidemic as a function of infection rate β_u for various mutation probabilities c in simulations. Notice how Eq. (14) correctly predicts that probabilities below the treated-disease threshold (dotted line) are more affected by variations in c than for above the threshold. (C) Effect of population size N on the probability of reaching a resistant strain epidemic. Scenarios above threshold feature macroscopic epidemics (fractions of N) of the wild strain and are thus significantly more affected by population size than those below thresholds where epidemics are microscopic (independent of N); also as predicted by Eq. (14). (D) Combining these behaviors for $c \to 0$ and $N \to \infty$, we can expect that a very effective treatment in a very large populations will feature a discontinuity in observed or expected total epidemic size (as shown solid red line). Other simulation parameters are set to the value of Fig. 2.

observe. In fact, in the case of extremely rare mutations (i.e. $c \to 0$), even the infinite system will feature a discontinuous jump as the probability of mutation becomes a step function (see Fig. 3). The system thus features a first-order phase transition in the limit of rare mutations.

This is of great interest for research in percolation processes as discontinuous phase transition in percolation models on networks have been claimed before [19, 20], but disproven [21]. We show here for the first time that these transitions can actually occur on a general network structure, as opposed to fractal networks [22]. While the mechanisms potentially leading to such transitions in percolation on networks are generally not well understood [20], the discontinuity in our biologically-inspired model can be explained by a classic phase transition concept. In short, a first-order (or explosive) phase transition is achieved because the resistant strain must wait for the wild strain to spread and then for treatment to allow resistance to spread throughout the system. While this is the most likely outcome above the threshold of the wild strain, this scenario is almost impossible for smaller infection rates. If more transmissible than the treated wild strain, the resistant strain then contains an "infection potential", conceptually equivalent to latent heat in classical phase transition theory, resulting in a discontinuity at transition.

Bistable and competitive regimes. For $c > 0$, there exists a regime of bistability where a given disease can either stay

in the disease-free state or reach the epidemic stable branch (hysteresis). Interestingly, when integrating our mean-field analysis with a finite precision, there exists a critical manifold (roughly $P \sim$ precision) marking a limit above which initial conditions escape the disease-free state towards the epidemic state. The analytical observation of the bistability was thus achieved by using different initial conditions (all $< 10^{-5}$). Though our finite simulations start with a single infectious in-
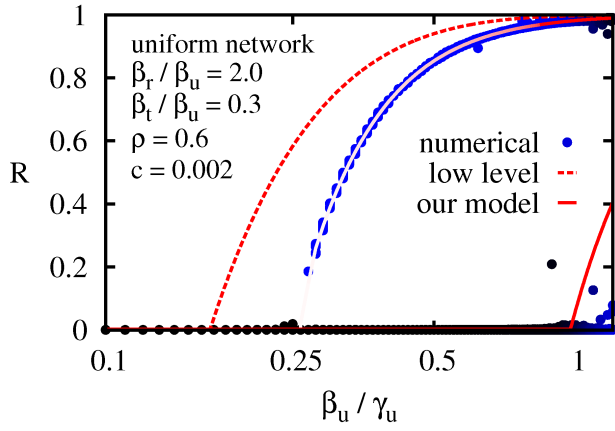
FIG. 4. (Color online) **Importance of state correlations and heterogeneity.** Using the uniform network (i.e. $p_k = \delta_{k,4}$), we see a very similar phase transition. For comparison, the dotted line corresponds to the prediction of classic epidemiological models, neglecting state correlations, as used in the original study of the present model [2]. Results of over 2×10^6 simulations on networks with 2.5×10^5 nodes are plotted, and both points and lines use the same color scheme as Fig. 2. The lower branch of our model, corresponding to epidemics of treated wild strain, is barely visited above its threshold (at $\beta_u \approx \gamma_u$) as the upper branch is by far the most likely outcome at this point.

dividual, they can stochastically tunnel through this manifold and reach the epidemic state. Figures 2 and 4 color the simulation results to illustrate the likelihood of such events for each transmissibility (points). As transmission rates increase, the system features a second phase transition corresponding to the epidemic threshold of the case without mutation ($c = 0$), after which all epidemics reach the highest branch (see Fig. 3).

This final regime also differs from regular percolation, as both strains end up competing for the potential infections. That is, if the wild strain spreads to high degree nodes early on, the system is less easily invaded by the resistant strain. To illustrate this competition for high degree nodes, consider the narrower spread of results on the uniform network of Fig. 4 as opposed to the large competitive regime observed on the heterogeneous network of Fig. 2. Within this competitive regime, the dynamics become highly sensitive to initial conditions. Although the different strains compete for the highest degree node even in the limit of an infinite population (i.e. the analytical system), this competition and sensitivity is always stronger in a finite system. In the limit of rare mutations and large infection rates, our model is akin to previous models of competing epidemics [23, 24].

Finally, note that these results are valid as long as the resistant strain propagates faster than the treated wild strain, i.e. $\beta_r > (1-\rho)\beta_u + \rho(1-c)\beta_t$ which is likely in practice according to realistic estimates [2]. Otherwise, the dynamics still feature competition, but lacks both bistability and the explosive phase transition as the disease never accumulates infection potential.

Discussion. In light of recent studies in first-order transition in percolation on networks, our simple biologically-inspired model of coevolutive competition provides deep insight into how discontinuous transitions can emerge in such systems: due to the build up of potential connectivity (latent heat) from coevolution. Similar results had previously been observed on adaptive networks whose structure changes through time [25] and in jamming transitions for network with traffic awareness where routing protocols depends on the network's state [26]. This arguably hints at a new universality class corresponding to coevolutive dynamics on networks.

Our results also have important implications for the control of epidemics in finite structured populations. Due to the presence of bistability and hysteresis, treatment effectiveness depends highly on the initial conditions [35]. This is especially important given the relative ease of many pathogens to evolve resistance to treatment [27–31] and the potential morbidity and mortality associated with treatment failure (for example, neuraminidase inhibitors oseltamivir and zanamivir for the treatment of influenza) [2, 32–34]. In that optic, future work will study the implications of resistance development for the optimal targeting, timing and scale of treatment strategies. Finally and most importantly, the first-order phase transition indicates that a *microscopic* change in transmission rate can lead to a severe *macroscopic* jump in the expected epidemic size. It is thus primordial that future efforts focus not only on reducing mutation probability in treatment, but also on detecting and controlling the emergence of resistance.

[1] N. M. Ferguson, D. A. T. Cummings, C. Fraser, J. C. Cajka, P. C. Cooley, and D. S. Burke, "Strategies for mitigating an influenza pandemic," *Nature*, vol. 442, pp. 448–52, Jul 2006.

[2] M. Lipsitch, T. Cohen, M. Murray, and B. R. Levin, "Antiviral resistance and the control of pandemic influenza," *PLoS Med*, vol. 4, p. e15, Jan 2007.

[3] J. T. Wu, G. M. Leung, M. Lipsitch, B. S. Cooper, and S. Riley, "Hedging against antiviral resistance during the next influenza pandemic using small stockpiles of an alternative chemotherapy," *PLoS Med*, vol. 6, p. e1000085, May 2009.

[4] H. J. Wearing and P. Rohani, "Ecological and immunological determinants of dengue epidemics," *Proc Natl Acad Sci U S A*, vol. 103, pp. 11802–11807, 2006.

[5] C. Nachenberg, "Computer virus-coevolution," *Communications of the ACM*, vol. 40, pp. 46–51, 1997.

[6] R. Pastor-Satorras and A. Vespignani, "Epidemic spreading in scale-free networks," *Phys. Rev. Lett.*, vol. 86, pp. 3200–3203,

2001.

[7] C. Lindan, S. Allen, M. Carael, F. Nsengumuremyi, P. V. de Perre, A. Serufilira, J. Tice, D. Black, T. Coates, and S. Hulley, "Knowledge, attitudes, and perceived risk of AIDS among urban Rwandan women: relationship to HIV infection and behavior change," *AIDS*, vol. 5, pp. 993–1002, 1991.

[8] D. K. Gauthier and C. J. Forsyth, "Bareback sex, bug chasers, and the gift of death," *Deviant Behavior*, vol. 20, pp. 85–100, 1999.

[9] J. A. Dunne, R. J. Williams, and N. D. Martinez, "Network structure and biodiversity loss in food webs: robustness increases with connectance," *Ecology Letters*, vol. 5, pp. 558–567, 2002.

[10] J. C. Miller, A. C. Slim, E. M. Volz, "Edge-Based Compartmental Modeling for Infectious Disease Spread Part I: An Overview," *J. R. Soc. Interface*, vol. 9, p. 890–906, 2012.

[11] M. E. J. Newman, "Spread of epidemic disease on networks," *Phys Rev E*, vol. 66, p. 016128, Jul 2002.

[12] M. J. Keeling and K. T. D. Eames, "Networks and epidemic models," *J R Soc Interface*, vol. 2, pp. 295–307, 2005.

[13] L. A. Meyers, "Contact network epidemiology: bond percolation applied to infectious disease prediction and control," *Bull. Amer. Math. Soc.*, vol. 44, pp. 63–86, 2007.

[14] E. Kenah and J. M. Robins, "Second look at the spread of epidemics on networks," *Phys. Rev. E*, vol. 76, p. 036113, 2007.

[15] J. C. Miller, "Epidemic size and probability in populations with heterogeneous infectivity and susceptibility," *Phys. Rev. E*, vol. 76, p. 010101(R), 2007.

[16] F.-Y. Wu, "The Potts model," *Rev. Mod. Phys.++*, vol. 54, pp. 235–268, 1982.

[17] B. R. Levin, V. Perrot, and N. Walker, "Compensatory mutations, antibiotic resistance and the population genetics of adaptive evolution in bacteria," *Genetics*, vol. 154, pp. 985–97, Mar 2000.

[18] S. Maisnier-Patin and D. I. Andersson, "Adaptation to the deleterious effects of antimicrobial drug resistance mutations by compensatory evolution," *Res Microbiol*, vol. 155, pp. 360–9, Jun 2004.

[19] D. Achlioptas, R. M. D'Souza, and J. Spencer, "Explosive percolation in random networks," *Science*, vol. 323, pp. 1453–1455, 2009.

[20] J. Nagler and A. L. M. Timme, "Impact of single links in competitive percolation," *Nature Physics*, vol. 7, pp. 265–270, 2011.

[21] O. Riordan and L. Warnke, "Explosive percolation is continuous," *Science*, vol. 333, pp. 322–324, 2011.

[22] S. Boettcher, V. Singh, and R. M. Ziff, "Ordinary percolation with discontinuous transitions," *Nature Communications*, vol. 3, pp. 1–5, 2011.

[23] V. Marceau, P.-A. Noël, L. Hébert-Dufresne, A. Allard and L.J. Dubé, "Modeling the dynamical interaction between epidemics on overlay networks," *Phys. Rev. E*, vol. 84, p. 026105, 2011.

[24] B. Karrer and M. E. J. Newman, "Competing epidemics on complex networks," *Phys. Rev. E*, vol. 84, p.. 036106, 2011.

[25] T. Gross, C. J. D. D'Lima, and B. Blasius, "Epidemic dynamics on an adaptive network," *Phys. Rev. Lett.*, vol. 96, p. 208701, 2006.

[26] P. Echenique, J. Gómez-Garde ñes and Y. Moreno, "Dynamics of jamming transitions in complex networks," *Europhys. Lett.*, vol. 71, p. 325, 2005.

[27] D. M. Weinstock and G. Zuccotti, "The evolution of influenza resistance and treatment," *JAMA*, vol. 301, pp. 1066–9, Mar 2009.

[28] M. Lipsitch, J. B. Plotkin, L. Simonsen, and B. Bloom, "Evolution, safety, and highly pathogenic influenza viruses," *Science*, vol. 336, pp. 1529–31, Jun 2012.

[29] S. Herfst *et al.*, "Airborne transmission of influenza A/H5N1 virus between ferrets," *Science*, vol. 336, pp. 1534–41, Jun 2012.

[30] C. A. Russell *et al.*, "The potential for respiratory droplet-transmissible A/H5N1 influenza virus to evolve in a mammalian host," *Science*, vol. 336, pp. 1541–7, Jun 2012.

[31] M. Imai *et al.*, "Experimental adaptation of an influenza H5 HA confers respiratory droplet transmission to a reassortant H5 HA/H1N1 virus in ferrets," *Nature*, vol. 486, pp. 420–8, Jun 2012.

[32] "WHO rapid advice guidelines on pharmacological management of humans infected with avian influenza A (H5N1) virus," tech. rep., World Health Organization, 2006.

[33] B. M. Althouse, T. C. Bergstrom, and C. T. Bergstrom, "A public choice framework for controlling transmissible and evolving diseases," *Proc Natl Acad Sci U S A*, vol. 107 Suppl 1, pp. 1696–701, Jan 2010.

[34] A. E. Fiore, A. Fry, D. Shay, L. Gubareva, J. S. Bresee, T. M. Uyeki, and Centers for Disease Control and Prevention (CDC), "Antiviral agents for the treatment and chemoprophylaxis of influenza — recommendations of the advisory committee on immunization practices (acip)," *MMWR Recomm Rep*, vol. 60, pp. 1–24, Jan 2011.

[35] B. M. Althouse, O. Patterson-Lomba, G. M. Goerg, and L. Hébert-Dufresne, "The Timing and Targeting of Treatment in Influenza Pandemics Influences the Emergence of Resistance in Structured Populations," *PLoS Comput Biol*, in press.