

Selective pinning control of the average disease transmissibility in an HIV contact network

E.F. du Toit and I.K. Craig*

*Department of Electrical, Electronic and Computer Engineering,
University of Pretoria, 0002 Pretoria, South Africa*

(Dated: June 1, 2015)

Medication is applied to the HIV-infected nodes of high-risk contact networks with the aim of controlling the spread of disease to a predetermined maximum level. This intervention, known as pinning control, is performed both selectively and randomly in the network. These strategies are applied to 300 independent realizations per reference level of incidence on connected undirectional networks without isolated components and varying in size from 100 to 10000 nodes per network. It is shown that a selective on/off pinning control strategy can control the networks studied with limited steady-state error and, comparing the medians of the doses from both strategies, uses 51.3% less medication than random pinning of all infected nodes. Selective pinning could possibly be used by public health specialists to identify the maximum level of HIV incidence in a population that can be achieved in a constrained funding environment.

PACS numbers: 87.19.lr, 89.75.Hc, 87.19.xd, 87.23.Ge

I. INTRODUCTION

HIV spread and incidence continues to be a challenging public health problem, and the development of new approaches to estimate future incidence attracts a great deal of attention. This is partly due to the difficulties that current approaches have in producing such estimates, such as the reliance on sparse household survey data prone to self-reporting bias, and the costly nature of laboratory-based methods. There are generally two approaches currently followed to obtain HIV incidence information, that of large cohort studies and that of mathematical models [1]. The latter is by far the cheapest and most popular.

One of the most important risk factors for the transmission of HIV is the number of sexual partners per individual [2]. If one represents the sexual partners of each person in a particular community as a contact network, intuitively the proportion of high-risk people with many partners will be small and the proportion with only one or a few partners comparatively large. Networks with such a degree distribution are typical in the real-world and can be modelled by the degree distribution proposed by [3], shown in Eq. (1).

$$p_k = \begin{cases} 0, & \text{for } k = 0 \\ \frac{k^{-\alpha} e^{-k/\kappa}}{Li_\alpha(e^{-1/\kappa})}, & \text{for } k \geq 1 \end{cases} \quad (1)$$

Note here that $Li_\alpha(x)$ represents the n th polylogarithm of x .

This distribution, with power-law exponent $\alpha = 2$ and degree cutoff at $\kappa = 5$ is applied to the networks in this work as shown in Fig. 1. The investigation of a scale-free network distribution has not been done in this work

due to the absence of an epidemic threshold in such networks [4]. This result implies that the techniques of bond percolation and compartmental disease modelling, which is Susceptible-Infected-Removed (SIR) and Susceptible-Infected-Susceptible (SIS) models, cannot be applied to a scale-free network. This is due to the proliferation of a disease on a scale-free topology regardless of its transmissibility, even if control is applied with either strategy in this work.

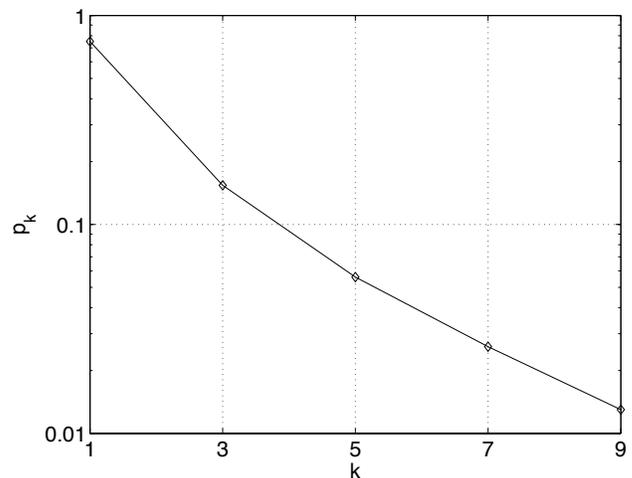


FIG. 1. Normalised degree distribution (p_k) of the simulated networks. The distribution is normalised to a fraction between zero and one. Parameters are fixed at $\alpha = 2$ and $\kappa = 5$. The average number of links (M) for each network of N nodes are: $N=100$ ($M=280$), $N=1000$ ($M=2792$), $N=5000$ ($M=13968$), $N=10000$ ($M=27938$). All networks have mean degree of 2.8.

Representing disease spread via proliferation through a contact network opens up the possibility to model and simulate the spread of HIV between nodes over time, especially given that the infection process is globally stable

* Corresponding author e-mail address: ian.craig@up.ac.za; Visit: <http://www.up.ac.za/ecece/article/1952234/prof-ik-craig-ian>

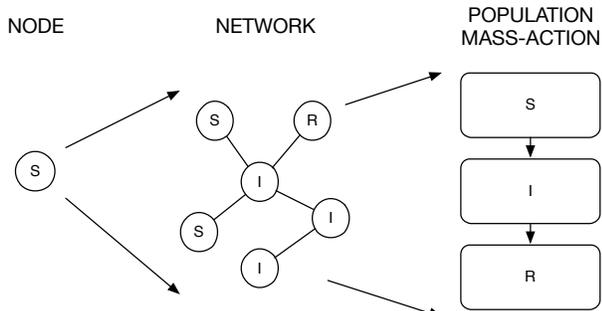


FIG. 2. Three tiers of models are simulated. A node-model is connected into a network of multiple nodes. The nodes of the network are then compartmentalised according to their susceptible, infected or removed (death) status, thus forming a population mass-action SIR model.

for all infection rates apart from the epidemic threshold T_c [5]. The influence of any intervention on a contact network can also be determined and optimal solutions can be sought.

One such intervention related to networks is named pinning control. Pinning control is a control-theoretic technique by which the dynamics of network nodes are manipulated by adding a control term to the node models (for example giving medication to a person as described in this paper). The control system design at the individual level could be (and more often is) a feedback control that is added [6], but in the case of the nodes of the HIV disease network represented in this work, this node-level control is open-loop. The treatment is antiretroviral medicine applied to infected nodes. This manner of pinning nodes (or influencing the microscopic dynamics of a network), can drive the network towards a particular equilibrium [7]. With unsynchronised or oscillating networks, pinning control has the capability to synchronise the complete network towards a particular steady-state by pinning only a few nodes [8], or in some cases by pinning only a single node.

For disease networks, which are already synchronised by virtue of their steady-state characteristics, the focus is on the impact of the pinning control method on this steady-state rather than on synchronisation.

A. Models

Three tiers of linked models are used as shown in Fig. 2, each on a different scale in the population. The first tier is a node model representing an individual. Secondly, a contact network model links individual nodes with one another and thirdly an SIR population model incorporates the death of nodes in the network.

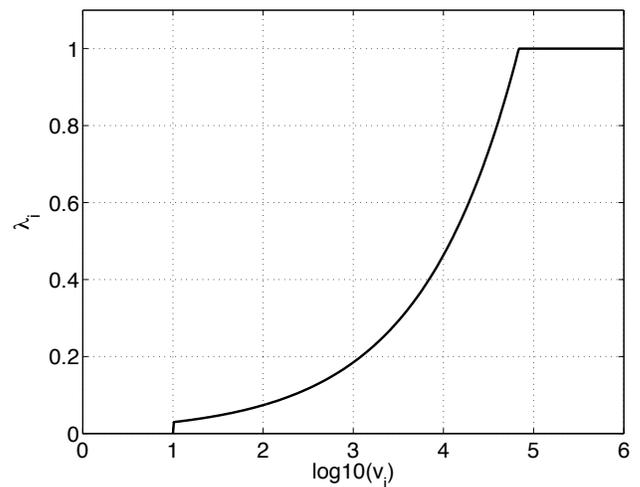


FIG. 3. Example transmission function (λ_i), where the added risk factor $R_f = 500$ indicates a very high risk HIV spread. The function is chosen to saturate just below a log-viral load of 5, for 100% transmission.

1. Node and network model

The node model incorporates the immune response of an individual to HIV infection. For this work, the node model is represented by the well-known 3D model of HIV *in-vivo* [9–11]. This model was the basis of the question for the control of HIV asked in [12], with parameters listed in Table I, as estimated in [13].

Each infected node has an inherent probability of transmitting disease to another node, influenced mainly by its viral load level. It is assumed that nodes receiving the infection are all equally susceptible and that nodes transmitting infection do so at the highest probability reachable given their risk profile. Additional risk is added via a factor R_f , and high risk is represented by $R_f > 100$. Additional risk factors that are included into the risk factor R_f include the presence of Sexually Transmitted Diseases (STDs), drug use and frequency of intercourse without protective measures such as condoms.

The highest risk groups are represented by the simulations in this work in order to suitably use an SIR model at the population level and to achieve disease spread above the epidemic threshold. Under normal risk circumstances HIV is not nearly as transmissible as represented here, but for the purpose of comparing control strategies this is a feasible configuration.

The per-sexual-act transmission probability of the i 'th node is λ_i , with increased risk added via the risk factor R_f as represented by Eq. (2) and shown in Fig. 3. The input to the function is node i 's viral load v_i [14].

$$\lambda_i = R_f * \left(1 - [1 - 9.36 \times 10^{-4}] e^{(-.92 * (\log_{10}(v_i) - 4))} \right),$$

with $0 \leq \lambda_i \leq 1$
and $\lambda_i = 0$, for $v_i = 0$ (2)

TABLE I. Summary of the node model parameters used in Eq. (3) [13, 15].

Parameter / Function	Description	Value / Range
T_i	Uninfected CD4+ T-cells	– copies/mm ³
T_i^*	Infected CD4+ T-cells	– copies/mm ³
v_i	Free virions	– copies/mm ³
T_{i0}	Initial uninfected CD4+ T-cells of first infected patient	1000 copies/mm ³
T_{i0}^*	Initial infected CD4+ T-cells of first infected patient	1 copy/mm ³
v_{i0}	Initial free virions of first infected patient	100 copies/mm ³
s	Source term for uninfected CD4+ T-cells	10.7 / (mm ³ day)
d	Death rate of uninfected CD4+ T-cells	0.015 / (mm ³ day)
θ	Infectivity rate of free virus particles	4.5×10^{-6} / (mm ³ day)
δ	Natural death of infected CD4+ T-cells	0.58 / day
k	Rate of virions produced per infected CD4+ T-cell	896.49 virions / (cell day)
μ	Death rate of virus	2.05 / day

Nodes are linked via the viral load state of the immune response model of each node, into a network represented by Eq. (3). All parameters used in Eq. (3) are obtained from literature. Here v_i is the viral load from the node model of the i 'th node, T_i^* is the infected CD4+ T-cells, μ is the natural death of virus in the body, a_{ij} is the connectivity matrix of the network and $\Gamma(x)$ is the activation function for transmission in the network. Given in Iverson Notation, where $[v_j > 0]$ evaluates to 1 if the condition in the bracket is *true*:

$$\begin{aligned} \dot{T}_i &= s - dT - \theta T v + \Delta_i u_i \\ \dot{T}_i^* &= \theta T v - \delta T - \Delta_i u_i \\ \dot{v}_i &= k T_i^* - \mu v_i + \Gamma(\lambda_i(t)) \sum_{j \neq i}^N a_{ij} [v_j > 0] \end{aligned} \quad (3)$$

Here, to control the node (in this paper to “pin” the node), the control u_i is activated by Δ_i as

$$\Delta_i u_i = \begin{cases} 0, & \text{for } \Delta_i = 0 \\ \epsilon_{RTI} \beta T v, & \text{for } \Delta_i = 1 \end{cases} \quad (4)$$

The function $\Gamma(x)$ is defined as follows:

$$\Gamma(x) = \begin{cases} 0, & \text{for } U([0, 1]) \leq x \\ 1, & \text{for } U([0, 1]) > x \end{cases} \quad (5)$$

where $U([0, 1])$ is a uniformly distributed random number between 0 and 1. To describe Eq. (3): At each time step of simulating the network model, transmission of a single virion from a neighbour j to node i is possible with probability λ_i , only if node j is already infected.

2. Population SIR model

An SIR model is used to be able to simulate an epidemic that reaches a particular maximum number of newly infected nodes (incidence) and then declines. This

maximum represents the steady-state in context of the simulations done here. The SIR model is given as

$$\begin{aligned} \frac{dS}{dt} &= -\beta IS \\ \frac{dI}{dt} &= \beta IS - \nu I \\ \frac{dR}{dt} &= \nu I \end{aligned} \quad (6)$$

Linking the network model in Eq. (3) with the SIR model in Eq. (6), and following the same notation as used in Eq. (3):

$$\begin{aligned} S &= \sum_{i=1}^N [v_i = 0] \\ I &= \sum_{i=1}^N [v_i > 0] \\ R &= N - S - I \end{aligned} \quad (7)$$

For the SIR model, N is the total number of nodes in the network. Also, β is the average transmissibility or rate of infection and ν is the rate of death or removal. The transmissibility of a specific node has already been defined in (2). The average transmissibility (β) of the network and the SIR model is taken as the sum of the individual transmissibilities from each of the infected nodes in the network as presented in Eq. (2), divided by the total number of infected nodes (I), as follows:

$$\beta = \frac{\sum_{i=1}^N \lambda_i}{I} \quad (8)$$

With the SIR model, the incidence I reaches a specific maximum number of newly infected nodes when $dI/dt = 0$ and then declines. The aim of this work is to control I to reach this point when $I = I_{ref}$.

The average transmissibility of a disease in a bond percolation model is also known as the bond occupation probability. For an outbreak of a disease to occur

TABLE II. Summary of the control system and network parameters used in this work.

Parameter / Function	Description	Value / Range
I_{ref}	Reference HIV incidence	0 – 100%
I_{max}	Maximum incidence of HIV for a network	0 – 100%
I_{sse}	Steady-state error of HIV incidence	0 – 100%
β	Average transmissibility of the network	0 – 1
$\hat{\beta}$	Bond-percolation estimate of average transmissibility	0 – 1
u	On/Off pinning controller output	0 – 100%
K	Controller gain	1
λ_i	Transmissibility of node i	0 – 1
R_f	Added risk factor for HIV transmission	500
N	Total number of network nodes	0 - 10000
S	Number of nodes susceptible to HIV infection	0 - N
I	Number of infected nodes	0 - N
R	Number of removed (by death) nodes	0 - N
ν	Mean probability of death for infected nodes	0.05

on a network, it has to spread from an initially infected node to other nodes in the network across the links between the nodes. This means that if one were to indicate that such transmission occurred across a particular set of links (or occupy such links in a bond percolation model, which happens with probability β), the ultimate size of the epidemic would be precisely the size of the cluster of nodes that can be reached from the initially-infected node by traversing only across infected (known as “occupied”) links.

B. Assumptions

The main assumptions made in this work are:

- Nodes in the network are homogenous and heterosexual, with identical HIV immune responses and consuming the same amount of medication.
- Sexual contact is assumed to be heterosexual, thus the network links represent a heterosexual network of sexual contacts.
- Reverse Transcriptase Inhibitors (RTIs) alone are effective in limiting the spread of HIV in the network. (Multiple drugs, as recommended by the World Health Organisation (WHO) standard treatment guidelines for maximal virus suppression, are not used in this work.)
- A single virion is transferred at infection.
- Sexual relationships are not broken during the time of the simulation.

In light of the given assumptions, this work aims to show how a control-based strategy could possibly be used by public health specialists to quantify the level of HIV incidence in a population. The results achieved could be made more realistic by relaxing these assumptions in a systematic manner.

II. METHODS

A. Bond percolation estimates of epidemic size

To predict the maximum number of individuals that will become infected in a network, this work draws on percolation theory from statistical physics [16], and in particular on bond percolation applied to epidemic networks [3]. Bond percolation relies on the degree distribution of nodes in a contact network (reflected in Fig. 1). If it is applied to the spread of disease [17], the probability that an infected node transmits a pathogen to any of its connected nodes is calculated. If its neighbour nodes are not yet infected, and the probability for transmission is sufficient, the disease is percolated throughout the network.

This percolation happens at an average transmissibility β in the network, and if this average can be determined for the entire period that the infection process in a network is under scrutiny, the maximum number of infected individuals in the particular network can be determined. [18] states that in an undirected network, the probability of an epidemic and the expected fraction of the network infected during an epidemic, are equal.

Bond percolation is used to determine the average transmissibility $\hat{\beta}$ needed to reach a given a final epidemic size I_{ref} . The value $\hat{\beta}$ is then used in this work as the reference control signal for the network. To be able to do this, verification is done to determine whether the bond percolation estimates indeed produce final epidemic sizes in the SIR model as proposed, given a particular average transmissibility. The desired epidemic size is then translated back to the verified average transmissibility β and an attempt is made to control the network to this value. Considering what is measured in the network, the average transmissibility is the key parameter. When the output incidence I is used by itself in the feedback control system for a network without taking β into consideration, the network will not reach the required target incidence.

1. Probability generating functions

Probability generating functions in this section have been defined by [3]. Their purpose is for easier analysis of a network topology's disease percolation characteristics.

In general, the m 'th power of a generating function generates the sum of a property k of an object (in this case the object is a node and k is the node degree) over m independent realisations of that object (which are the m nodes of the network) and the first derivative of a generating function produces the mean of the probability distribution.

The following generating functions, further described in [19], are thus constructed.

First, $G_0(x)$ provides the degree distribution:

$$G_0(x) = \sum_{k=0}^{\infty} p_k x^k \quad (9)$$

with p_k the probability that a node has degree k .

Including disease into the network, propagating at average transmissibility β , the equation becomes:

$$G_0(x; \beta) = G_0(1 + (x - 1)\beta) \quad (10)$$

The following generating function constitutes the excess degree. This means that it generates a distribution equal to the degree of nodes minus the occupied or infected links to those nodes. The generating function $G_1(x)$ is given by:

$$G_1(x) = \sum_{k=0}^{\infty} q_k x^k \quad (11)$$

where q_k can also be calculated as:

$$q_k = \frac{(k+1)p_{k+1}}{z} \quad (12)$$

with $(k+1)$ signifying the actual degree of a node and k the number of links by which infection did not spread to the node, with the average degree z of nodes in the network given by:

$$z = \dot{G}_0(1) = \sum_k k p_k \quad (13)$$

Similarly as for $G_0(x)$, Eq. (11) is extended to include the propagation of disease on the network, given by:

$$G_1(x; \beta) = G_1(1 + (x - 1)\beta) \quad (14)$$

From [20], one can see that the basic reproductive number (\mathcal{R}_0) is defined as given in (15). This number signifies the average number of secondary infections resulting from a single node in the network:

$$\mathcal{R}_0 = \beta \dot{G}_0(1) \quad (15)$$

From this, the epidemic threshold transmissibility β_c is defined as the specific average transmissibility in the network at which the average number of secondary infections

resulting from a single node is greater than 1. For this, \mathcal{R}_0 is equal to 1, and:

$$\beta_c = \frac{1}{\dot{G}_0(1)} \quad (16)$$

Then the generating function $H_1(x)$ for the total number of nodes reachable by following a specific link, is given by:

$$H_1(x; \beta) = x G_1(H_1(x; \beta); \beta) \quad (17)$$

The generating function for the distribution of the total number of nodes reachable from a randomly chosen node is given by:

$$H_0(x; \beta) = x G_0(H_1(x; \beta); \beta) \quad (18)$$

When $\beta_c \leq 1$ (below the epidemic threshold \mathcal{R}_0), the network will experience outbreaks of average size $\langle s \rangle$ [3], where

$$\langle s \rangle = 1 + \frac{\beta \dot{G}_0(1)}{1 - \beta \dot{G}_1(1)} \quad (19)$$

Above the epidemic threshold, most nodes in the network will become infected. The proportion of the network infected can be estimated by:

$$I_{ref} = 1 - H_0(1; \beta) \quad (20)$$

In this paper the epidemic size is used as reference input I_{ref} to the pinning control system and the average transmissibility β needed to reach I_{ref} is calculated numerically. The bond percolation functions $H_0(x; \beta)$ and $H_1(x; \beta)$ are simulated numerically, using a recursive function, to converge to the required solution.

2. Verification of the accuracy of epidemic estimates

For large networks ($N = 100000$), bond percolation estimates have been verified by [3]. Smaller networks ($N = 100$) are also verified in this paper for completeness. In Fig. 4, the solid line represents the bond percolation estimates of the epidemic size, given a particular average transmissibility. Each circle on the graph represents the average epidemic size from 300 networks simulated at a fixed average transmissibility.

It can be seen that percolation estimates slightly overestimate the epidemic size at lower transmissibilities and slightly underestimate the epidemic size at higher transmissibilities. This verification confirms, by inspection, that the predictions may be used as input to the control system proposed.

To obtain the estimated average transmissibility given the epidemic size (I_{ref}), requires the use of Fig. 4 and the numerical solution of β in Eq. (21).

$$G_0(H_1(1; \hat{\beta}); \hat{\beta}) = 1 - I_{ref} \quad (21)$$

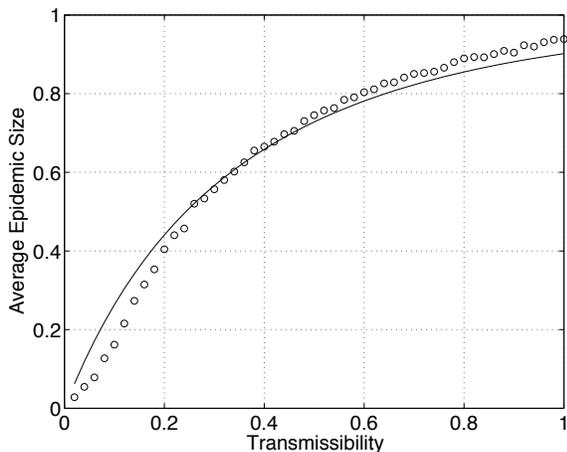


FIG. 4. Bond percolation estimates of the epidemic size (In this paper I_{ref}) plotted against the average transmissibility β . Estimates are represented by the solid line and each circle represents the simulated average of 300 independent networks of size $N=100$. Similar results were presented by [3], for networks of size $N=100000$, verifying the estimates for scale.

B. Feedback on/off pinning control of a network

The purpose of using feedback control is to be able to compare the ability of two strategies of medication to control the networks to reach a target epidemic size. The two strategies are random pinning and selective pinning. Random pinning provides medication to all infected nodes. Selective pinning medicates the proportion of infected nodes with the highest number of susceptible nodes connected to them in order of highest to lowest. A feedback control system of the form shown in Fig. 5 is applied to the HIV contact networks. The control system measures average transmissibility of HIV in the network and proportionally medicates more (or less) people based on the difference between the bond percolation predicted transmissibility and the actual transmissibility. This medication is given in an on-off pulsed manner to represent doses, where a pulse is made when $\beta \geq \hat{\beta}$ and no pulsing when $\beta < \hat{\beta}$.

A network node is controlled (pinned) by setting Δ_i from Eq. (4) equal to 1 (Pin the node) or equal to 0 (Don't pin the node).

A summary of all the control and performance parameters used in this section is given in Table II.

The flow of the feedback control system shown in Fig. 8 is described below:

a. Input The input to the feedback loop is the required average maximum network incidence I_{ref} . Note that this target can only be estimated by taking the average across several hundred networks, given the definition of bond percolation estimates from which I_{ref} is calculated.

b. Input translation The input, I_{ref} is translated to an average transmissibility $\hat{\beta}$ using by using Eq. (21).

c. Error calculation The difference between the estimated transmissibility $\hat{\beta}$ and the simulated average network transmissibility β is calculated and used as input to the controller.

d. Controller The controller selects the proportion of individuals in the network to receive medication. The i th node is pinned by setting its Δ_i parameter to 1 in Eq. (4). The output of the controller is limited to 1, given that this represents the control of all nodes in the network. If the random pinning strategy is used, all infected nodes are given medication. This strategy is random because the infection process is random. If selective pinning is used, the proportion of infected nodes to receive medication is calculated by Eq. (22). Nodes are then pinned in order of their highest number of connected susceptible individuals until the proportion u is reached. The control output u saturates at 1.

$$u = \begin{cases} 0, & \text{for } \beta < \hat{\beta} \\ K \times (\beta - \hat{\beta}), & \text{for } \beta \geq \hat{\beta} \\ 1, & \text{for } \beta \geq (2 \times \hat{\beta}) \end{cases} \quad \text{with } 0 \leq u \leq 1 \quad (22)$$

e. Gain The controller gain K is chosen to be 1 in this work.

f. Actuator The actuator is the application of medication (RTIs) to the immune response systems of the pinned nodes in the network. A node is pinned by setting Δ_i from Eq. (4) equal to 1. The drug used to simulate medication is a reverse transcriptase inhibitor named Tenofovir with an efficiency of 65% [15].

g. Process The process is the disease network and SIR model encapsulating it. The input to the process is the number of nodes to be pinned (controlled) and the output is the current incidence of HIV in the network.

h. Output The maximum of the output incidence from the process is calculated. This value is I_{max} .

C. Simulations

Simulations of all networks are done over a maximum of 2000 days. Network sizes of nodes $N=100$, $N=1000$, $N=5000$ and $N=10000$ are used. Random and selective pinning are applied to 300 independent realizations of connected networks without isolated components, for each reference incidence level from 20% to 60% in increments of 10%. For the control system shown in Fig. 5 to function correctly, the average I_{max} of the 300 realizations for each I_{ref} was taken.

III. RESULTS

A. Controlled β and Incidence

Having applied random and selective pinning to varying sizes of networks, the outcomes can be seen in Fig.

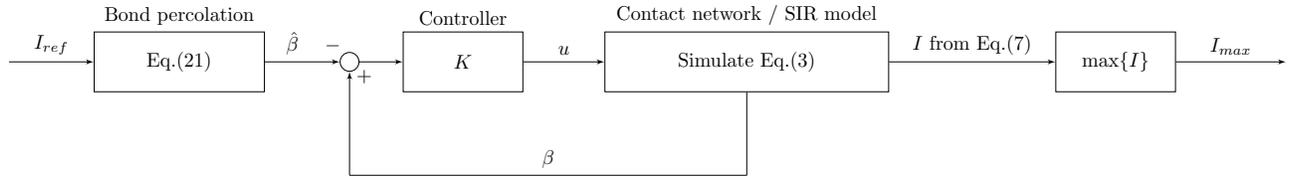


FIG. 5. Feedback control system.

6. For all investigated transmissibilities, the control performance (over- or undershooting of the target) for both strategies are similar. At lower references of around 20% the incidence is controlled to within 5% of the reference. For small networks of $N = 100$, the control of incidence always performs to within 5% for all strategies. As the network size increases, the less accurate the control to higher reference incidences (above 20%) becomes. The constant control gain value of $K = 1$, is the most likely explanation for this from a control perspective. From a network perspective, this means that larger networks are impacted differently by control than smaller networks. At higher transmissibilities and with larger networks ($N \geq 1000$), both control strategies are unable to reach the target incidence if the gain is kept constant.

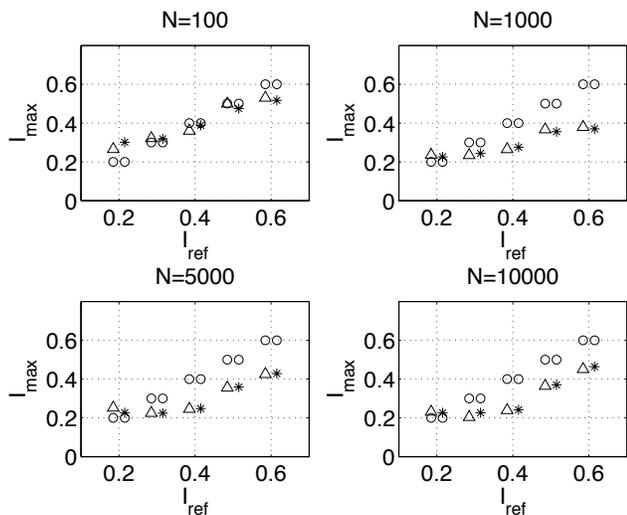


FIG. 6. The average maximum incidence for each reference signal I_{ref} . Each marker represents the average of 300 independently simulated networks. Asterisks (*) are randomly pinned networks (all nodes pinned) and triangles (Δ) represent selectively pinned networks. The reference points, I_{ref} , are represented by circles (\circ).

B. Average Steady-State Error

The average steady-state error simulation results can be seen in Fig. 7. The closer the markers (the mean value over 300 networks' data) are to zero, the more accurately the control system drove the network towards the desired reference. With both random and selective pinning, the lower references resulted in negative errors for all sizes of network. By inspection, there is almost no difference in steady-state error between selective and random pinning, for all sizes of network investigated.

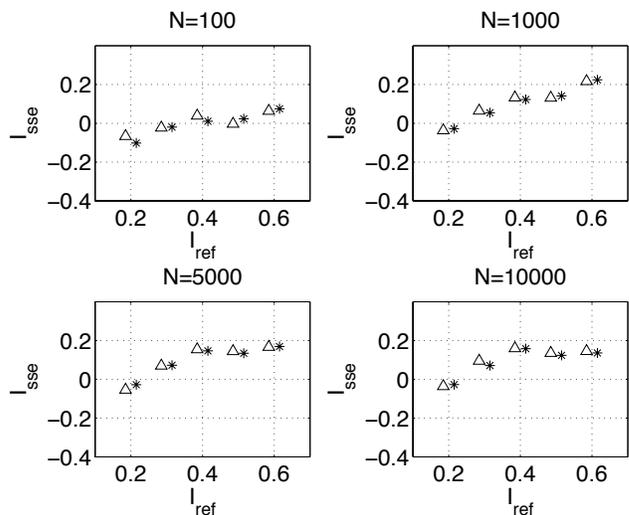


FIG. 7. The average steady-state error between the average maximum incidence and the reference maximum incidence. Each marker represents the average of 300 network simulations with asterisks (*) the random pinning strategy and triangles (Δ) the selective pinning strategy.

C. Control effort

Control effort expenditure is reported as the comparison between the medians of doses used for each pinning control strategy. This is necessary because neither distribution of doses is Gaussian or symmetric. The measure of doses itself is obtained by taking the sum of all control action pulses over time, per network instance, and

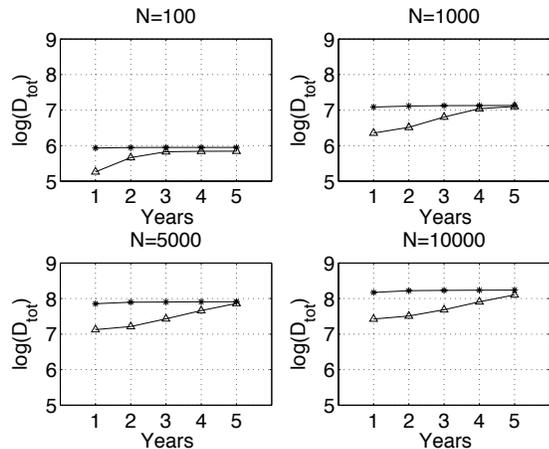


FIG. 8. The log of the cumulative total doses, $\log(D_{tot})$, for the two types of pinning control strategies over 5 years. Asterisks (*) represent random pinning and triangles (Δ) selective pinning. Each yearly data point is the cumulative control effort from 1500 independent networks spanning the studied reference levels I_{ref} .

multiplying this with the total number of nodes N in a network. Thereafter, the sum of the control action across all 300 network instances per data point is taken to obtain the total control action for the respective data point. After a period of 5 years, the calculated total doses per strategy appears to be much closer in value to each other than in preceding years. Each year of simulation presents the selective pinning control as using cumulatively more control action than the year before, whilst the random pinning strategy seems to cumulatively use all its control action within the first year of simulation. Thereafter, the cumulative random pinning does not vary considerably. This effect can be seen in Fig. 8.

A statistical analysis was performed on the cumulative doses after 5 years, both for networks stratified according to size and across network sizes. The results can be seen in Table III. Firstly, the normality of the data across all network sizes for selective pinning and random pinning was compared with respect to their distributions. The data was found not to be normally distributed, and the two strategies differed in their fundamental distributions. The Kolmogorov-Smirnov (KS) test for non-parametric data was used to confirm that the distributions of both dosing strategies are different, with p-value of less than 2×10^{-16} . The Inter-Quartile-Ranges (IQR) are also given. The ratio of the medians of the two sets of overall data suggests that, for the simulations performed, selective pinning control uses 51.3% less doses compared to random pinning after 5 years. The difference was statistically significant using the Mann-Whitney U-test, with a p-value less than 2×10^{-16} , similar to the KS-test. Stratifying for network size, the medians suggest that selective pinning uses 76% less doses for $N=100$, 48% less doses for $N=1000$, 24% less doses for $N=5000$ and 44% less doses

for $N=10000$ nodes.

TABLE III. Descriptive statistics for the doses applied under the two pinning control strategies of Selective (Δ) and Random (*) pinning across all values of I_{ref} , after 5 years.

N	Statistic	Selective (Δ)	Random (*)
100	Sample Size	1500	1500
	Median	145.87	600
	IQR	[73.55, 672.70]	[400, 800]
	Median Ratio (Δ :*)		0.24
1000	Sample Size	1500	1500
	Median	4641	9000
	IQR	[1052, 8436]	[5000, 12000]
	Median Ratio (Δ :*)		0.52
5000	Sample Size	1500	1500
	Median	38030	50000
	IQR	[5943, 80160]	[30000, 75000]
	Median Ratio (Δ :*)		0.76
10000	Sample Size	1500	1500
	Median	56070	100000
	IQR	[10840, 147200]	[60000, 170000]
	Median Ratio (Δ :*)		0.56
All	Sample Size	6000	6000
	Median	6823.0	14000
	IQR	[872.7, 43150]	[1000, 70000]
	Median Ratio (Δ :*)		0.48

IV. DISCUSSION

Two network-wide control strategies, random and selective pinning, were proposed. Both strategies were implemented with an On/Off type controller that medicates (with one dose each) a specific number of nodes when the actual incidence is above the reference incidence.

It has been shown that contact networks with a high transmissibility and removal of nodes can be controlled to a target incidence by On/Off proportional control, although a constant gain results in an increasing steady-state error as the network size increases.

The control schemes proposed here have an important impact on HIV incidence, in particular, the maximum incidence can be controlled to within 20% with either strategy and a constant gain.

Selective pinning results in a similar average steady-state error for the target incidence. This means that neither strategy outperforms the other on accuracy.

The statistical analysis performed on the total control effort applied across all simulations per strategy suggests, by comparing the medians of the dosing data, that *selective pinning uses 51.3% less medicine compared to random pinning*.

The main limitations of this work due to the assumptions in Section IB are:

- In real-world HIV networks, each individual's immune response is different. The node models would need to be updated for each node for more accurate

results. To obtain enough information to estimate the parameters of each individual is a very resource-intensive task.

- HIV rarely spreads with such high rates of transmissibility as used in this work. The results here therefore only applies to very high risk HIV contact networks.
- For simplicity, a single drug (an RTI named Tenofovir) has been used with an efficiency of 65% [15]. Multiple drugs are the recommended treatment scheme, hence the node models should be updated before comparing the results obtained to a real-world scenario.

V. CONCLUSIONS

This work quantified the impact of two different pinning control schemes on high-risk contact networks in which HIV is transmitted. The average transmissibil-

ity could be controlled to reach a particular target average maximum incidence in the simulated networks. It was found that selective pinning is equally as accurate as random pinning in reaching the reference incidences. A statistical analysis of the two dosing strategies suggested that selective pinning requires, according to the median of doses, 51.3% less medication (control action) compared to random pinning. These results were achieved with a constant gain controller.

ACKNOWLEDGMENTS

The authors would like to acknowledge the funding of the South African Centre for Epidemiological Modelling and Analysis (SACEMA) and the inputs provided by Prof. Martin Niewoudt from the University of Stellenbosch and Anna Grimsrud from UCT CIDER. This work is based on research supported in part by the National Research Foundation of South Africa (Grant Number 90533).

-
- [1] Thomas Rehle, Olive Shisana, and Victoria Pillay, "National HIV incidence measures new insights into the South African epidemic," *South African Medical Journal* **97**, 194–199 (2007).
- [2] Paul Arora, Nico J D Nagelkerke, and Prabhat Jha, "A systematic review and meta-analysis of risk factors for sexual transmission of HIV in India." *PloS One* **7**, e44094 (2012).
- [3] M. Newman, "Spread of epidemic disease on networks," *Physical Review E* **66**, 016128 (2002).
- [4] Romualdo Pastor-Satorras and Alessandro Vespignani, "Epidemic Spreading in Scale-Free Networks," *Physical Review Letters* **86**, 3200–3203 (2001).
- [5] Xinchu Fu, Michael Small, and Guanrong Chen, *Propagation dynamics on complex networks: Models, methods and stability analysis* (John Wiley & Sons, 2013).
- [6] A M Jeffrey, X Xia, and I K Craig, "When to initiate HIV therapy : A control theoretic approach," *IEEE Transactions on Biomedical Engineering* **50**, 1213–1220 (2003).
- [7] X. Li, X. Wang, and G. Chen, "Pinning a Complex Dynamical Network to Its Equilibrium," *IEEE Transactions on Circuits and Systems I: Regular Papers* **51**, 2074–2087 (2004).
- [8] L.Y. Xiang, Z.X. Liu, Z.Q. Chen, F. Chen, and Z.Z. Yuan, "Pinning control of complex dynamical networks with general topology," *Physica A: Statistical Mechanics and its Applications* **379**, 298–306 (2007).
- [9] A S Perelson, D E Kirschner, and R De Boer, "Dynamics of HIV infection of CD4+ T cells." *Mathematical Biosciences* **114**, 81–125 (1993).
- [10] MA Nowak and CRM Bangham, "Population dynamics of immune responses to persistent viruses," *Science* **272**, 74–79 (1996).
- [11] MA Nowak and R May, *Virus dynamics: mathematical principles of immunology and virology* (Oxford University Press, Oxford, 2001).
- [12] I K Craig and X Xia, "Can HIV/AIDS be controlled?" *IEEE Control Systems Magazine* **25**, 80–83 (2005).
- [13] Reuben A Filter, Xiaohua Xia, and Clive M Gray, "Dynamic HIV/AIDS parameter estimation with application to a vaccine readiness study in southern Africa." *IEEE Transactions on Biomedical Engineering* **52**, 784–91 (2005).
- [14] James P Hughes, Jared M Baeten, Jairam R Lingappa, Amalia S Magaret, Anna Wald, Guy de Bruyn, James Kiarie, Mubiana Inambao, William Kilembe, Carey Farquhar, and Connie Celum, "Determinants of per-coital-act HIV-1 infectivity among African HIV-1-serodiscordant couples." *The Journal of infectious diseases* **205**, 358–65 (2012).
- [15] Sulav Duwal, Christof Schütte, and Max von Kleist, "Pharmacokinetics and pharmacodynamics of the reverse transcriptase inhibitor tenofovir and prophylactic efficacy against HIV-1 infection." *PloS One* **7**, e40382 (2012).
- [16] P Grassberger, "On the critical behavior of the general epidemic process and dynamical percolation," *Mathematical Biosciences* **63**, 157–172 (1983).
- [17] Lauren Ancel Meyers, M E J Newman, and Babak Pourbohloul, "Predicting epidemics on directed contact networks," *Journal of Theoretical Biology* **240**, 400–418 (2006).
- [18] Lauren Ancel Meyers, "Contact network epidemiology: Bond percolation applied to infectious disease prediction and control," *Bulletin of the American Mathematical Society* **44**, 63–87 (2006).
- [19] M. E. J. Newman, "The structure and function of complex networks," *SIAM Review* **45**, 167–256 (2003).
- [20] Fred Brauer, "An introduction to networks in epidemic modeling," in *Mathematical epidemiology* (Springer Berlin Heidelberg, 2008) pp. 133–146.