

Epidemic graph diagrams as analytics for epidemic control in the data-rich era

Received: 18 January 2023

Accepted: 22 November 2023

Published online: 20 December 2023

 Check for updatesEugenio Valdano¹, Davide Colombi², Chiara Poletto³ & Vittoria Colizza¹ ✉

COVID-19 highlighted modeling as a cornerstone of pandemic response. But it also revealed that current models may not fully exploit the high-resolution data on disease progression, epidemic surveillance and host behavior, now available. Take the epidemic threshold, which quantifies the spreading risk throughout epidemic emergence, mitigation, and control. Its use requires oversimplifying either disease or host contact dynamics. We introduce the epidemic graph diagrams to overcome this by computing the epidemic threshold directly from arbitrarily complex data on contacts, disease and interventions. A grammar of diagram operations allows to decompose, compare, simplify models with computational efficiency, extracting theoretical understanding. We use the diagrams to explain the emergence of resistant influenza variants in the 2007–2008 season, and demonstrate that neglecting non-infectious prodromic stages of sexually transmitted infections biases the predicted epidemic risk, compromising control. The diagrams are general, and improve our capacity to respond to present and future public health challenges.

Public health response to infectious disease epidemics faces two major challenges. First, to predict if an emerging pathogen will cause a large-scale epidemic. Second, to control epidemics and prevent recurring outbreaks of known pathogens. SARS-CoV-2 is evidence of both: a new pathogen which caught the world by surprise, wreaked havoc, and tested our control capabilities with recurrent waves¹. Global crises like the COVID-19 pandemic may become more frequent, as climate change increases the threat of new viral species jumping to humans from animal reservoirs². Among them, avian strains of influenza A (e.g. H5N1, H7N9) are closely monitored for their ability to develop into major pandemic strains³. With a universal vaccine still unavailable, preparing for the next influenza pandemic requires large-scale access to antiviral treatment⁴ and optimal drugs administration^{5,6}.

The fight against epidemics of known pathogens is equally hard, and exceptionally relevant in the context of sexually transmitted infections (STI). The progress towards elimination of HIV/AIDS is faltering⁷. Chlamydia, gonorrhoea and syphilis cause more than 200 million new cases worldwide among adults each year⁸, and widespread antimicrobial resistance is compromising our ability to fight them⁹. Yet

elimination remains a priority, to reduce not only their burden, but also their effect on HIV acquisition and transmission⁸. Finally, mpox virus has spread to non-endemic areas, overwhelmingly through sexual transmission and in communities already burdened by other STIs¹⁰.

The two epidemic contexts—pandemics and STI—have little in common. But the goals of preventing large-scale emerging epidemics and eliminating endemic diseases find a common theoretical framework and application tool in the epidemic threshold^{11,12}. The epidemic threshold defines the critical value of disease transmissibility above which the infection can establish itself in a host population. Interventions that raise the epidemic threshold make the population more resilient to the pathogen. Opposite changes increase the epidemic risk.

The epidemic threshold has historically contributed to inform core public health activities, for its potential to provide context awareness and gauge the effort needed for prevention or eradication^{13–15}. In the current COVID-19 pandemic, scientists, public health officials, authorities, and even the general public have continuously evaluated control policies by their ability to bring epidemic

¹Sorbonne Université, INSERM, Institut Pierre Louis d'Epidémiologie et de Santé Publique, F75012 Paris, France. ²aizoOn Technology Consulting, 10146 Turin, Italy. ³Department of Molecular Medicine, University of Padova, 35121 Padova, Italy. ✉e-mail: vittoria.colizza@inserm.fr

waves below the epidemic threshold, i.e., reproductive number $R_t < 1^{16-19}$.

Current theories, however, cannot handle the complexity of infectious disease dynamics in real populations, and resort to two oversimplified schemes. One consists in allowing for detailed disease progression but limited population structure, and stemmed from mathematical epidemiology²⁰⁻²². This clashes, however, with the evidence of highly complex time-varying contact patterns measured or estimated in different contexts^{15,23-25}. To overcome this, the physics and network science community has pushed the development towards increasing realism in population structure^{12,26-32}. This has happened, however, at the expense of disease description³³⁻³⁷, which has opened new problems: Simplifying assumptions on disease natural history bias results³⁸, and variations in intervention protocols cause radically different epidemic outcomes⁶. These simplifications limit our knowledge of disease dynamics, and cripple the ability of models to serve public health.

The theoretical formalism presented here will eschew both. It will give estimates of the epidemic threshold that are both robust, to provide rapid and generalizable understanding of complex dynamical processes driving disease spread, and accurate, to turn high-precision data³⁹⁻⁴² into targeted and viable recommendations. Applied to syphilis transmission, it shows that the resolution of sexual contact data requires matching resolution in modeling disease progression. Namely, its prodromic stage affects the epidemic threshold even if non-infectious. Applied to influenza, our framework predicts whether antiviral-resistant variants will become dominant, in agreement with the observed emergence and fixation of the oseltamivir-resistant A(H1N1) variant in 2007.

Results and discussion

Epidemic graph diagrams

We introduce here Epidemic Graph Diagrams (EGDs) to represent the coupled dynamics of arbitrarily complex disease spread, and arbitrarily complex contact patterns among hosts. Following a tradition in physics^{43,44}, our diagrams are more than a representation of the diffusion equations on the network substrate: they replace them. We define rules to build the diagrams, and operations to manipulate them, by exploiting their topological properties. Then, we show that diagrams lead to a general, analytical derivation of the epidemic threshold. Also, diagram operations simplify its numerical computation, making EGDs a practicable analytics for public health.

We consider the classic compartmental description of the natural history of an infectious disease, through a progression of successive disease states regulated by transition rates. We consider population structure through an explicit time-varying contact network with adjacency matrix $\mathbf{A}(t)$, where nodes represent hosts and the entry $A^{ij}(t)$ encodes transmission-relevant connectivity between host i and host j at time t ^{12,15,23,24,45}. Contacts can thus change in time.

We use the Markov chain formulation in the quenched mean-field approximation, whereby the network appears in its explicit contact structure represented by the adjacency matrix $\mathbf{A}(t)$, and the dynamical states of the nodes are independent of each others and follow a system of deterministic differential equations^{26,28,30}. For a generic compartmental model these equations can be written as:

$$\dot{x}_b^i = \sum_c \sum_j x_c^j \left[(\gamma_{bc} - \mu_c \delta_{bc}) \delta^{ij} + \lambda_{bc} A^{ij}(t) \right] + f(x), \quad (1)$$

where $x_b^i(t)$ is the probability that node i is in compartment b at time t , and b runs over all compartments, except the susceptible S . μ_c , γ_{bc} , λ_{bc} are the rates of the transitions between compartments depicted in Fig. 1a: μ correspond to spontaneous transitions to S (e.g. recovery with no immunity); γ correspond to spontaneous transitions between any two compartments except S (e.g. exposed hosts becoming

infectious); λ correspond to transmission events causing a susceptible node to enter another compartment (e.g. a susceptible becoming infectious). Transmissions are generated by infectious hosts of any kind (e.g. asymptomatic or symptomatic infectious individuals) when they are in contact with susceptible nodes. The component $f(x)$ in Eq. (1) contains all terms that are nonlinear in \mathbf{x} (quadratic or higher). It may be complex, but its derivatives vanish in the disease-free state ($\partial f / \partial x_b^i |_{x=0} = 0$), therefore f does not contribute to the epidemic threshold and can be dropped. Eq. (1) can then be conveniently rewritten in operator form:

$$\dot{\mathbf{x}}(t) = \mathbf{x}(t)[\boldsymbol{\gamma} - \boldsymbol{\mu} + \boldsymbol{\lambda}\mathbf{A}(t)] = \mathbf{x}(t)\mathbf{J}(t). \quad (2)$$

In this formalism, \mathbf{x} is a vector on $\mathbb{R}^N \otimes \mathbb{R}^{N_c}$, where N is the number of hosts, and N_c is the number of compartments (excluding S). $\boldsymbol{\mu}$, $\boldsymbol{\gamma}$, $\boldsymbol{\lambda}$ are operators on \mathbb{R}^{N_c} , and A is an operator on \mathbb{R}^N . The term $\mathbf{J}(t) = \boldsymbol{\gamma} - \boldsymbol{\mu} + \boldsymbol{\lambda}\mathbf{A}(t)$ is the Jacobian of the system, itself an operator on $\mathbb{R}^N \otimes \mathbb{R}^{N_c}$. The component $\boldsymbol{\lambda}\mathbf{A}(t)$ generates the interaction between the contact network structure and the structure of the compartmental model, so that nodes that are neighbors at a given time t affect each other's probabilities to find themselves in given compartments.

The dynamics of the epidemic close to the disease-free state is captured by the infection propagator \mathbf{P} (also an operator on $\mathbb{R}^N \otimes \mathbb{R}^{N_c}$), counting the possible transmission chains among hosts^{29,30,33}. Following³⁰, we compute \mathbf{P} as a function of \mathbf{J} using the theory of nonautonomous linear systems on Eq. (2):

$$\mathbf{P}(t) = \mathcal{T} \exp \left\{ \int_{t_0}^t ds \mathbf{J}(s) \right\} = \mathcal{T} \exp \left\{ \int_{t_0}^t ds [\boldsymbol{\gamma} - \boldsymbol{\mu} + \boldsymbol{\lambda}\mathbf{A}(s)] \right\}, \quad (3)$$

where t_0 is the initial time of the time window of analysis and \mathcal{T} is Dyson's time-ordering operator (see Methods). The largest eigenvalue of \mathbf{P} yields the epidemic threshold^{29,30,33}.

EGDs emerge as a graph-theoretical representation of the analytical treatment just presented. They are network representations that fully encode the infectious disease dynamics in real populations close to the critical spreading condition. They are composed by N_c nodes, each representing a disease compartment (except S), connected by single directed links (spontaneous transitions) or double directed links (transmission events) (Fig. 1b). Nodes can also have self-loops (single, for transitions to S ; double, for transmissions entering the same node). EGDs bypass Eq. (1), as they can be built directly from the classic compartmental model and the network adjacency matrix with simple rules (steps 1-2 of Fig. 1d). Most importantly, computing the weighted adjacency matrix of the EGD (plus probability conservation) yields exactly the operator $\mathbf{J}(t)$ (Fig. 1c,d) appearing in the expression of the infection propagator of Eq. (3) (see "Methods").

Under the simple representation of a weighted directed graph, EGDs hide a higher-order complexity. Unlike common graphs, link weights are not scalar but operators on \mathbb{R}^N . Considering the relation between tensors and multilayer networks⁴⁶, we can interpret EGDs as graphs of layers: each node in the EGD (compartment) is a replica layer of the contact network, and links in the EGD are coupling operators among layers.

Operationally, EGDs make it possible to simplify different disease models and interventions, compare them, build equivalence classes. This is done through three diagram operations: *CUT*, *ZIP*, and *SHRINK*. They decompose diagrams into simpler parts, and compress complex disease progressions. Figure 2 describes in detail the diagram operations, which we present hereafter through specific epidemic applications. The corresponding proofs are reported in Methods. Finally—and more pragmatically—EGDs are a powerful tool for computing the epidemic threshold.

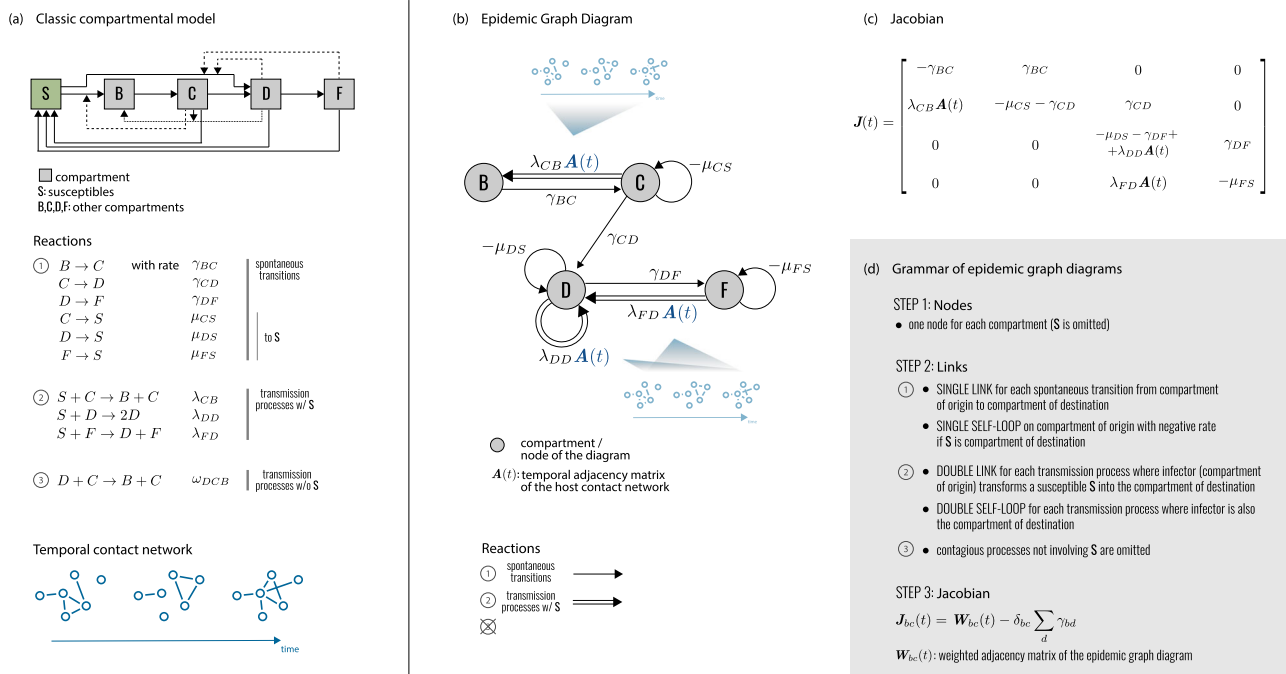


Fig. 1 | Epidemic graph diagrams. **a** Network epidemiology ingredients: compartmental model (top), and time-evolving network (bottom). The generic compartmental model appears in its standard representation: squares represent the different compartments, joined by transitions of three types. The susceptible compartment S (healthy individuals who can contract the disease from the infectious) is made explicit. We use generic letters for compartments to show the wide applicability of the approach without constraining to specific disease progressions. The transitions stem from the diagram visualization at the top of the panel. Type 1 transitions are shown as continuous lines and correspond to spontaneous processes; type 2 are shown as dashed lines, for transmission events involving susceptibles; type 3 are shown as dotted lines, for transmission events not involving susceptibles. **b** Epidemic Graph Diagram corresponding to the compartmental model of (a). Spontaneous transitions (type 1) become single links. Transmission

events (type 2) become double links. Links are weighted by the corresponding transition rates (see a) multiplied by an operator on \mathbb{R}^N . This operator is the identity matrix on single links (omitted), and the adjacency matrix on double links. All other transitions, e.g. transmissions infecting other compartments than S (type 3) can be neglected and do not appear in the EGD (see “Methods”). **c** Jacobian corresponding to the example in (a, b). **d** Rules of the EGD grammar. The EGD is built directly from the compartmental model (a) following steps 1 and 2. The Jacobian is then the weighted adjacency matrix of the EGD minus diagonal terms that enforce the probability conservation (step 3), encoded in the diagonal entries γ_{bb} . These are not free parameters but are fixed by probability conservation, since transitions among compartments not involving S do not change the sum $\sum_c x_i^c$ (Eq. (1)). From this, we derive $\gamma_{bb} = -\sum_{c \neq b} \gamma_{bc}$.

Syphilis

As a first application, we consider syphilis spreading on a network of sexual contacts. Syphilis is a bacterial STI that is still widespread in resource-constrained countries and resurgent in high-income countries in specific risk groups⁸. After infection, the disease progresses through a latency period lasting on average 3–4 weeks but with large fluctuations (10–90 days)⁴⁷. It is then followed by two infective stages—primary and secondary syphilis. If untreated, syphilis enters latency, potentially leading to severe complications in the tertiary stage. Most models greatly simplify this disease progression and consider early stages only, under the assumption of antibiotic treatment, through susceptible-infectious-susceptible⁴⁸ or susceptible-infectious-recovered-susceptible approaches⁴⁹. Latency period is neglected, because complexity is shifted from disease progression towards capturing the heterogeneity of human sexual behavior⁴⁸.

Here we want to keep such complexity in $A^i(t)$ while also preserving the progression of the early stages of syphilis infection. Introducing the latency period following infection is often the first step into building a multi-stage disease history in realistic epidemic contexts⁵⁰. Following the grammar of Fig. 1, we build the epidemic graph diagram of a susceptible-exposed-infectious-recovered-susceptible (SEIRS), where E corresponds to individuals exposed to the disease but not yet infectious (Fig. 3). Diagram operation *CUT* allows the isolation of two strongly connected components, one containing only R, and one containing E and I (Fig. 3b). The first has no transmission terms. It is thus always below threshold, and can be dumped.

The second leads to the simplified version of the Jacobian of the original diagram (Fig. 3c).

The diagram can however be further simplified if the network of contacts satisfies the weak-commutation condition³⁰ (see “Methods”). This applies to homogeneous mixing, static and annealed networks³⁰, and some temporally evolving network models such as the activity-driven model⁵¹. Under this condition, diagram operation *SHRINK* allows the removal of those nodes (compartments) having only one outgoing link in the EGD, as long as the link is single (Fig. 2c, e). Applied to syphilis, *SHRINK* eliminates E and reduces the original model to an SIS (Fig. 3c), whose threshold can be easily solved with the scalar version of the infection propagator approach^{29,30,33}. Under weak-commutation, the latency period therefore plays no role in the condition for endemic circulation, extending to temporal networks the result previously restricted to static and annealed networks⁵². For example, approaches neglecting syphilis latency within the homogeneous mixing approximation⁴⁹ will not be biased if temporal correlations in sexual activity are negligible. If instead such correlations exist, errors are to be expected. To evaluate such errors, we consider syphilis circulating on a sexual network between Brazilian sex workers and clients²³ and measure the relative variation of the epidemic threshold computed on the *CUT* and on the *CUT+SHRINK* diagrams of Fig. 3c. Within the variation of latency duration reported in syphilis infected individuals⁴⁷, we find errors in elimination predictions ranging from –5 to 40% (Fig. 3d, e), if the model neglects latency. The presence of one additional timescale in a non-infectious compartment induces an

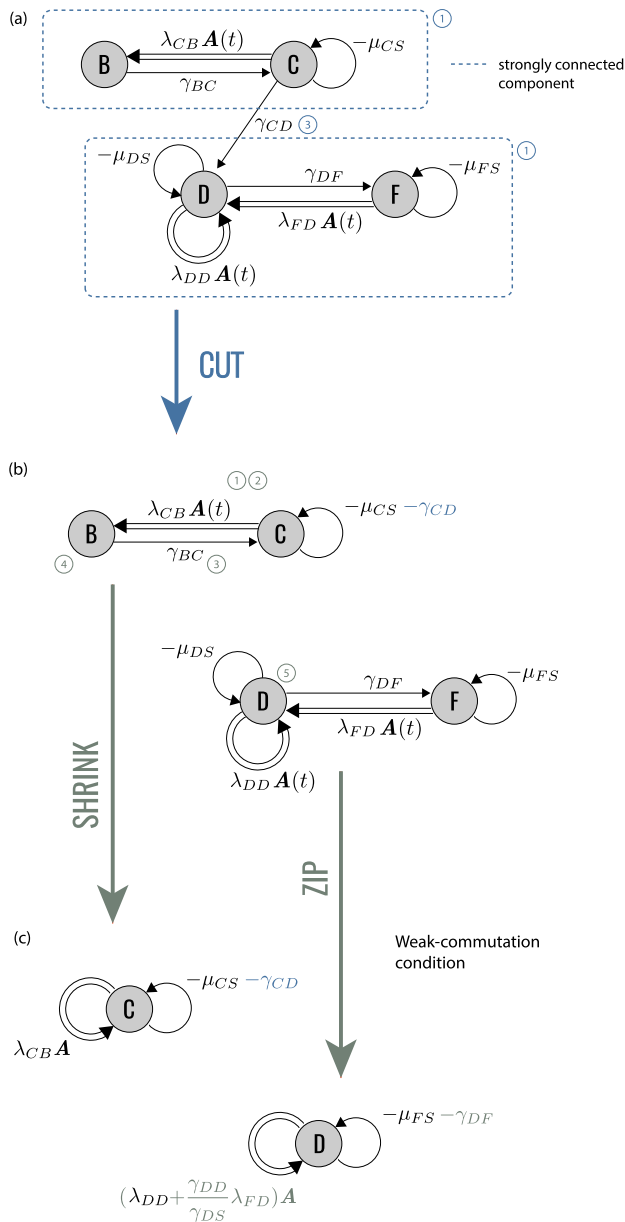


Fig. 2 | Diagram operations. **a** EGD of Fig. 1. The two strongly connected components are highlighted by blue rectangles. **b** Diagram operation *CUT* isolates the strongly connected components, which become disjoint subdiagrams. The epidemic threshold is then the smallest among the thresholds of each subdiagram. **c** Under the weak-commutation condition, each subdiagram can be further

- (d) Simplification rules - CUT**
- 1 identify all SCCs of the diagram, considering both link types as directed links (isolated nodes are SCCs)
 - 2 drop double links across different components
 - 3 convert simple links across different components into selfloops around their origin node and change sign of their weight (see $-\gamma_{CD}$)
 - 4 compute the threshold of each component
 - 5 the epidemic threshold of the original model is given by the smallest threshold of the components

- (e) Simplification rules under weak-commutation**
- SHRINK**
- 1 identify all pairs of nodes connected by opposite single and double links
 - 2 the source node of the simple link has no other outgoing link (simple or double) and no selfloops
 - 3 drop the simple link between the two nodes of the pair
 - 4 drop the node receiver of the double link of the pair and redirect all its incoming links to the other node (the double link of the pair becomes a double self-loop on the remaining node)
- ZIP**
- 5 verify that the component of M nodes has only one node receiver of all double links (including double self-loops)
 - 6 the component of M nodes is zipped into that node with simple and double self-loops with effective weights

reduced through (i) diagram operation *SHRINK* (top) and (ii) diagram operation *ZIP* (bottom), leading to two EGDs of SIS models. Their transition rates are renormalized by the diagram operations. **d, e** General rules of the three operations. Numbers refer to the steps to be performed for each operator; some of these steps are illustrated in (a–c) with the same numbers.

interplay with the timescales of the contacts and of the other disease stages, either increasing or decreasing the conditions for syphilis spread with respect to predictions based on the SIS model. In a wave analogy, latency acts in tuning the phase shift between disease and network, from in-phase (boosting) to counterphase (hampering)⁵³.

In the absence of treatment, or if treatment fails (e.g. for resistance emergence), later stages of syphilis infection need to be considered. The Supplementary Note reports the corresponding EGD and shows that even in that scenario the diagram can be simplified to the EGD of an SIS model (section S1).

The effect observed here for syphilis is relevant to other diseases, especially those for which natural history is poorly known, typically in an early phase of the outbreak. This is the case, for example, of the mpox outbreak: estimates from endemic areas were not directly applicable to the epidemic, which features different spreading routes,

symptomatology, and risk factors for acquisition¹⁰. Waiting for reliable estimates of disease time scales (latency, generation time, detection) and stages (asymptomatic transmission), the analysis of epidemic scenarios requires a flexible and agile theoretical framework, as the one developed here, to account for uncertainties.

Pandemic influenza

Progression from latent infection to active disease may be far more complex than what illustrated for syphilis. Fully describing the natural history of influenza in human hosts, for example, requires a progression from exposed (E) to pre-symptomatic (P, infectious without showing symptoms yet), asymptomatic (A) or symptomatic (I) infectious, and recovered (R) (Fig. S3)⁶. Next to the inclusion of the latency period, the distinction among infectious individuals is critical for public health interventions. Only symptomatic infectious (I) can be

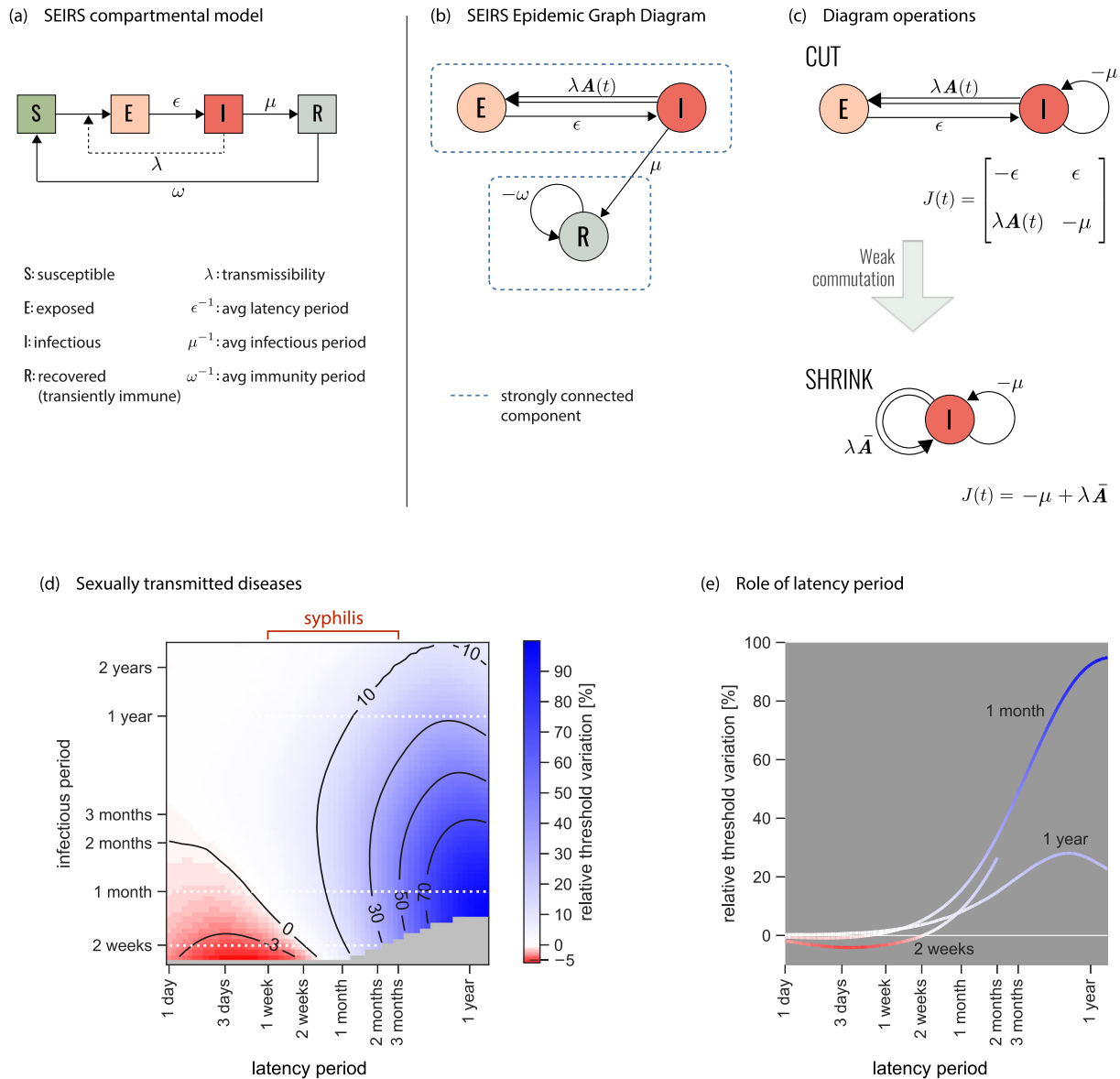


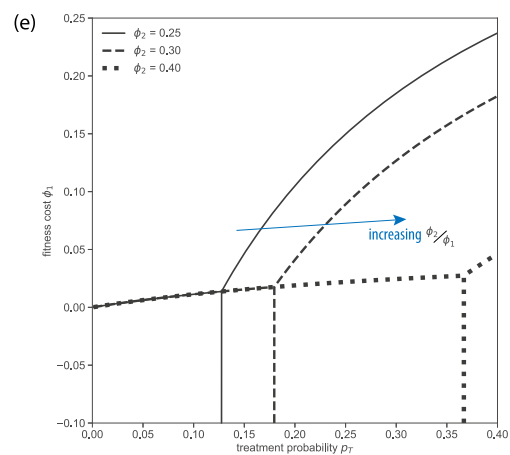
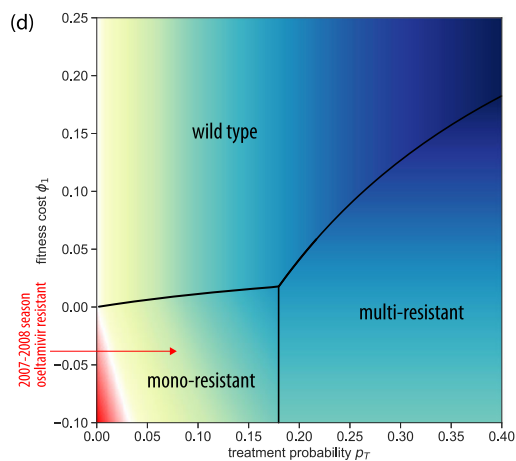
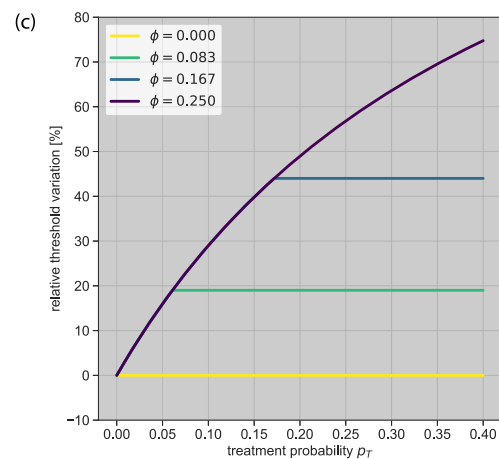
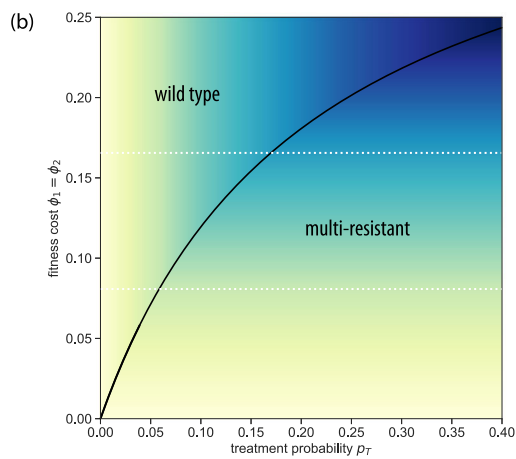
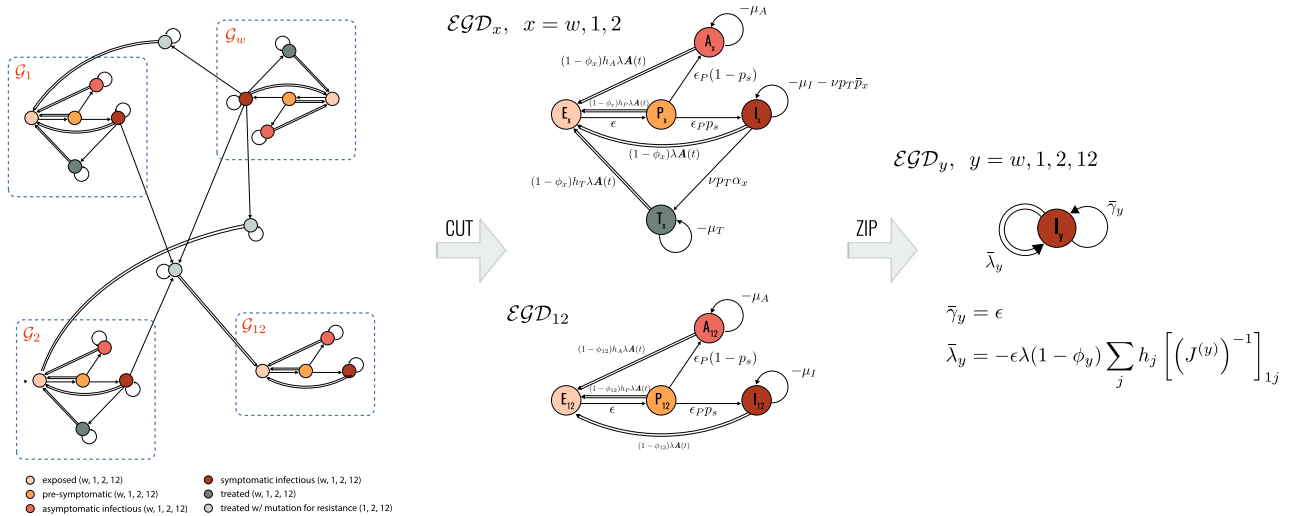
Fig. 3 | Epidemic threshold for syphilis. **a** Classic representation of the susceptible-exposed-infectious-recovered-susceptible (SEIRS) compartmental model here used to model the spread of a general STI. Parameterization for syphilis is provided in **(d)**. E represents the class of individuals exposed to the infection who are exposed to the disease, before becoming infectious (I). R corresponds to temporary immunity. Rates are also shown. **b** Associated EGD. Strongly connected components are highlighted by blue rectangles. **c** Simplification of the EGD through diagram operations. The reduced diagram after *CUT* is shown with its associated Jacobian (top; the R component does not contribute to the threshold and is discarded). Under the weak-commutation condition, *SHRINK* further reduces the diagram to an EGD of an SIS model (bottom). **d** Relative difference in the prediction of the epidemic threshold in the full model (that is equivalent to the top diagram of

panel c, after *CUT*) vs. the *CUT*+*SHRINK* diagram (bottom diagram of panel c) obtained under the weak-commutation condition (i.e. removing the role of latency). Results are obtained for an STI spreading on sexual network from real data²³, exploring different lengths of infectious period (i.e. time-to-treatment) and latency period. Parameterization for syphilis infection is highlighted. A negative relative threshold variation indicates the threshold is lower in the full model compared to the weak-commutation one, i.e. the more realistic model predicts a higher risk than the approximated one. Both negative and positive variations are observed in the region of parameters corresponding to syphilis infection. The gray region of the plot indicates that the system is always below threshold. **e** Relative threshold variation as a function of the latency period for the three infection durations (2 weeks, 1 month, 1 year) corresponding to the white dashed lines in **(d)**.

detected in the population and administered antivirals. P, A cannot. Antivirals treat severe forms of seasonal influenza, but most importantly represent the first line of defense against an emerging pandemic, as they mitigate morbidity and mortality at the population level^{5,6}. For this reason, neuraminidase inhibitors – the most common influenza antivirals—are stockpiled for pandemic preparedness⁴. However, large-scale antiviral administration may favor the emergence and spread of antiviral resistant influenza strains, compromising individual treatment and pandemic control⁵⁴. Combined administration of two antivirals (combination therapy) has been proposed to defuse this

threat, as opposed to monotherapy⁶. To study its effects in altering the pandemic potential of the circulating pandemic strain, we build the EGD of the full pandemic influenza disease progression model, adding antiviral combination therapy and emergence of drug resistance. Symptomatic infectious individuals (I) are treated with a certain probability p_T . Mutation can occur in treated individuals giving rise to mono-resistant (to either drug, 1 or 2) and multi-resistant variants, with fitness cost ϕ . In the context of current pandemic preparedness⁴, drugs 1 and 2 would correspond to oseltamivir Tamiflu and zanamivir Relenza. The full model is described in the Supplementary Note

(a) Epidemic Graph Diagram for pandemic influenza with combination therapy and development of resistance



(section S2). It includes 25 compartments, 19 of which describe infectious stages.

Despite the model complexity, four strongly connected components can be identified in the corresponding EGD (Fig. 4a, neglecting the trivial component composed only of R, as before). They summarise the infection dynamics of the wild-type, the two mono-resistant, and the multi-resistant strains. Moreover, the diagrams of the first three are isomorphic, underlying the equivalence of the associated dynamics.

The CUT operation provides then the epidemic threshold as the lowest among the critical conditions of the different strains. Each condition can be computed from data on population structure, $A(t)$, and with empirically-informed strain-specific parameter values (section S2).

But the influenza model can be further simplified in most contexts. The short time scale of face-to-face proximity interactions along which influenza transmission can occur²⁴ makes annealed approximation, i.e. the weak-commutation assumption³⁰, the commonly

Fig. 4 | Epidemic threshold for pandemic influenza with antiviral combination therapy and emergence of resistance. **a** EGD (epidemic graph diagram) of the pandemic influenza model and associated simplifications. The model includes four strains—wild-type, two mono-resistant, and one multi-resistant. The recovered (R) compartment has already been *CUT* out for the sake of visualization (see also Fig. 3b). The four strongly connected components (blue rectangles) correspond to the dynamics of each strain, independently. They can be isolated using *CUT*. EGD indicates the representation of an epidemic graph diagram. Under the weak-commutation condition, the *ZIP* operation reduces each of the four subdiagrams to an SIS-like EGD with renormalized strain-specific parameters. Parameter definitions and values appear in Supplementary Tables S1–S3 and S5. **b** Relative threshold variation under treatment ($p_T > 0$) compared to no treatment ($p_T = 0$) as a function of treatment probability p_T and fitness cost, assumed to be the same for both antivirals ($\phi_1 = \phi_2$). Positive relative threshold variation indicates the epidemic threshold is higher when antiviral drugs are used for therapy, i.e. the risk for a

pandemic is reduced. The black line separates the two dominance regimes, for the wild-type strain and for the multi-resistant variant. **c** Relative threshold variation as a function of p_T along the values of fitness cost ϕ_1 indicated by the white dashed lines in **(b)**. Treatment increases the epidemic threshold. But after a critical p_T the multi-resistant strain becomes dominant and further increasing treatment has no additional effect. **d** As **(b)** when fitness costs are specific to the drug, i.e. $\phi_2 > \phi_1$. A phase in which a mono-resistant strain dominates appears, differently from the situation depicted in **(b)**. In addition, when $\phi_1 < 0$ (resistance increases transmissibility), a region in parameter space emerges where threshold variation is negative (red region), i.e. the pandemic risk is increased by the use of antiviral drugs, due to the emergence of resistance. The red arrow indicates the parameter values estimated for the oseltamivir-resistant H1N1 strain, globally dominant in 2007–2008⁵⁴. **e** Boundaries of the three dominance phases of panel d when varying ϕ_2 . We report their analytical derivation in S2.2.

adopted approximation. The diagram can then be fully compressed through the *ZIP* operation (Fig. 2c, e), because transmission events involve exclusively susceptible individuals entering the same compartment (here, exposed E). The difference between *ZIP* and *SHRINK* is that *ZIP* imposes global requirements on the diagram topology and compresses a full diagram, whereas *SHRINK* has only local requirements and merges two compartments at the time. The multi-resistant strain EGD and the three isomorphic EGDs (wild-type, and two mono-resistant strains), made of 4 and 5 compartments, respectively, are all *ZIP*ped into an EGD of an SIS model with renormalized transmission and recovery rates (Fig. 4a). The physical intuition is that multi-stage disease progression may change the speed of diffusion, but not whether the epidemic will break out or not, provided there are no dynamical interactions between the disease and the underlying contact network. Theoretically, we are able to disentangle the dynamics of the four different viral strains, analyze each of them separately, and show that they are equivalent (after appropriate parameter renormalization). Practically, we reduce a 25-compartment influenza model into a simple SIS model with substantial numerical gain: computing the largest eigenvalue of four N -dimensional matrices, instead of a $23N$ -dimensional matrix.

The factorization of the network component in the weak-commutation condition (section S2) allows us to make predictions on the dominant influenza variant by comparing the strain-specific thresholds against the scenario with no treatment ($p_T = 0$). Increasing treatment coverage helps controlling pandemic influenza, especially for high fitness costs of the resistant variants (Fig. 4b with $\phi_1 = \phi_2$). Above a certain value of p_T , however, the multi-resistant strain dominates and the likelihood of its establishment is not affected anymore by the treatment coverage. The saturation effect depends on the fitness cost of the resistant strains (Fig. 4c).

Allowing fitness costs to depend on the specific drug ($\phi_1 < \phi_2$) leads to the emergence of an additional phase where the mono-resistant strain dominates, for small enough treatment probabilities and fitness costs (Fig. 4d). Notably, there exists a range of ϕ_1 where increasing the intervention coverage leads to all possible phases—namely, either the wild-type dominates (small p_T), or the mono-resistant strain (intermediate p_T), or the multi-resistant variant (large p_T). The dominance of a single mono-resistant strain is favoured by an increase in transmissibility following mutation ($\phi_1 < 0$), as expected. Though rare, such condition was observed in the 2007–2008 influenza season when an oseltamivir-resistant H1N1 variant emerged and rapidly spread, becoming the dominant H1N1 strain globally⁵⁴. Dominance can happen also under reduced transmissibility of both mono-resistant strains ($\phi_1 > 0$), for a smaller region of parameter values. Large clusters of oseltamivir-resistant H1N1 variants were isolated in Japan in 2013–2014 influenza season that caused large community outbreaks but did not lead to large-scale dominance⁵⁴. Increasing the

difference between the two fitness costs magnifies the size of the mono-resistant phase (Fig. 4e).

The models presented so far are not age-structured. In the Supplementary Note (Section S3) we use COVID-19 to show that EGDs can accommodate age-structured populations, a crucial feature for diseases for which exposure, transmission, or morbidity are age-dependent.

The use of EGDs for pandemic influenza and syphilis show the versatility of the theoretical framework in solving the critical epidemic conditions while handling the full complexity of contact data, disease natural history, and interventions. Limited work so far has used a network representation to describe the dependency between different disease stages⁵⁵. The EGDs formalism makes this dependency and the role of the human contact network transparent. EGDs are agnostic of the data they are fed. As such, they may provide inaccurate or imprecise estimates if the input data are biased or feature high uncertainty, which may be the case for data feeds coming in real time or representing future estimates. EGDs, however, requiring minimal computational power, can accommodate any sensitivity analysis by varying the input data and observing the response in the EGD output. EGDs share the same approximation of the infectious propagator approach, i.e. the quenched mean-field assumption^{26,28}. Its validity has been numerically tested for the critical condition⁵⁶, and specifically within the infection propagator framework²⁹. All other results leading to the EGDs and their simplifications are obtained from the properties of the Jacobian, under no additional approximation. EGDs thus provide the analytics to predict public health risks at high granularity and to customise interventions, responding to the challenges of today's public health.

Methods

Infection propagator and epidemic threshold

We expand here the expression of the infection propagator of Eq. (3), which was derived using the theory of nonautonomous linear systems⁵⁷. Dyson's time-ordering operator⁵⁸ is defined as follows: $\mathcal{T} \mathbf{A}(t_1) \mathbf{A}(t_2) = \theta(t_1 - t_2) \mathbf{A}(t_1) \mathbf{A}(t_2) + \theta(t_2 - t_1) \mathbf{A}(t_2) \mathbf{A}(t_1)$. θ is Heaviside's step function. The time-ordered exponential of Eq. (3) is then a common representation of the following series:

$$\mathbf{P}(t) = \sum_{h=0}^{\infty} \frac{1}{h!} \int_{t_0}^t dy_1 \cdots dy_h \mathcal{T} [\mathbf{y} - \boldsymbol{\mu} + \boldsymbol{\lambda} \mathbf{A}(y_1)] \cdots [\mathbf{y} - \boldsymbol{\mu} + \boldsymbol{\lambda} \mathbf{A}(y_h)]. \quad (4)$$

Time-ordering can be made explicit:

$$\mathbf{P}(t) = \sum_{h=0}^{\infty} \int_{t_0}^t dy_1 \int_{t_0}^{y_1} dy_2 \cdots \int_{t_0}^{y_{h-1}} dy_h [\mathbf{y} - \boldsymbol{\mu} + \boldsymbol{\lambda} \mathbf{A}(y_h)] \cdots [\mathbf{y} - \boldsymbol{\mu} + \boldsymbol{\lambda} \mathbf{A}(y_1)]. \quad (5)$$

The infection propagator encodes the epidemic dynamics close to the epidemic threshold: $P(t)_{bc}^j$ is the probability that j is in compartment c at time t , given that i is in compartment b at time $t=0$, under the quenched mean-field assumption. Equation (5) is also the most explicit formulation of the interaction between the temporal evolution of the epidemic and the temporal evolution of the contact network. We can now generalize what found in ref. 29 for the SIS model, and compute the epidemic threshold. The epidemic threshold is the critical parameter surface separating the disease-free state to the endemic phase. In the SIS model it is a one-dimensional curve relating the transmission rate from the recovery rate. Here instead it is a $(N_C - 1)$ -dimensional surface defined as follows:

$$\rho[\mathbf{P}(T)] = 1, \tag{6}$$

where T is the final observation time and ρ indicates the spectral radius, i.e., the largest eigenvalue. In practice, it is still possible to obtain an equation for one single parameter if all transmission rates λ_{bc} are expressed as functions of the baseline transmission rate, whose critical value gives the epidemic threshold. This is the case of the pandemic influenza model, for example, where transmissibility of each infectious compartment (e.g. P, A, treated symptomatic infectious individuals, mono-resistant or multi-resistant strains, etc.) is defined as a rescaling of the transmissibility of the symptomatic infectious individual (e.g. transmissibility of the mono-resistant strain is equal to the transmissibility of the wild-type, rescaled for the fitness cost, see Supplementary Note section S2).

To numerically compute the spectrum of \mathbf{P} , an operator on $\mathbb{R}^N \otimes \mathbb{R}^{N_C}$, we interpret the tensor products in \mathbf{J} as Kronecker products. The resulting matrix has dimension MN_C , and has a block structure, with each block of dimension N , as shown in Fig. 1c. Formally, we are exploiting the isomorphism $\mathbb{R}^N \otimes \mathbb{R}^{N_C} \simeq \mathbb{R}^{NN_C}$, and this is equivalent to the supra-adjacency representation of multilayer networks⁴⁶. When $\gamma = 0$, and $N_C = 1$ (scalar), we recover the infection propagator of the susceptible-infectious-susceptible model³⁰.

If the network obeys the *weak-commutation condition*^{30,59}, both the infection propagator (Eq. (3)) and the threshold (Eq. (6)) simplify. The weak-commutation condition is defined in ref. 30 as

$$\int_{t_0}^t dx [\mathbf{A}(x), \mathbf{A}(t)] = 0, \forall t, \tag{7}$$

where $[\mathbf{A}(x), \mathbf{A}(t)] = \mathbf{A}(x)\mathbf{A}(t) - \mathbf{A}(t)\mathbf{A}(x)$ is the standard matrix commutator. As ref. 30 proves, this is true in two cases: (i) temporal correlations in the evolution of $A(t)$ are absent (network annealing); (ii) the timescale of temporal correlations in the evolution of $A(t)$ is much shorter than the timescale of the spread of the disease (timescale separation). In both cases all the integrals in Eq. (4) commute and the time ordering can be dropped. As ref. 30 proves, this implies that it is possible to replace $\mathbf{A}(t)$ with the average adjacency matrix $\bar{\mathbf{A}} = \int_{t_0}^t dt \mathbf{A}(t) / (T - t_0)$, and, in the same way, to replace the time-evolving Jacobian $\mathbf{J}(t)$ with its average $\bar{\mathbf{J}} = \int_{t_0}^t dt \mathbf{J}(t) / (T - t_0)$ over the time window of analysis. This makes it possible to sum the series in Eq. (3). The infection propagator is then the exponential of the average Jacobian:

$$\mathbf{P}(t) = e^{\bar{\mathbf{J}}} = e^{t(\boldsymbol{\gamma} - \boldsymbol{\mu} + \lambda \bar{\mathbf{A}})}. \tag{8}$$

Its spectrum is the exponentiation of the spectrum of $\bar{\mathbf{J}}$. From this, the equation of epidemic threshold becomes

$$\rho[\boldsymbol{\gamma} - \boldsymbol{\mu} + \lambda \bar{\mathbf{A}}] = 0. \tag{9}$$

In case of an SIS model ($\gamma = 0$; μ, λ scalars), Eq. (9) becomes the well-known formula $\lambda = \mu / \rho[\bar{\mathbf{A}}]$ ^{26,28}.

Lastly, we comment on the term $f(x)$ in Eq. (1). As stated in the paper, this does not contribute to the epidemic threshold because its contribution to the Jacobian in the disease-free state vanishes. In the classical representation of a compartmental model showed in Fig. 1a, type 3 transitions (i.e. transmissions infecting other compartments than S) would be contained in f . They would correspond to quadratic terms of the form $x_b x_c$. Since f can be ignored, this means that type 3 transitions can be ignored too, and never appear in EGDs.

From the infection propagator to the EGD

We stated that epidemic graph diagram is a graphical representation of J , which can be obtained from the EGD by computing its weighted adjacency matrix (plus conservation condition), as shown in Fig. 1d:

$$J_{bc}(t) = W_{bc}(t) - \delta_{bc} \sum_d \gamma_{bd}. \tag{10}$$

To proof the equivalence between the EGD and the infection propagator approach to analytically compute the epidemic threshold, we need to demonstrate that the Jacobian of the above equation is exactly $\mathbf{J}(t) = \boldsymbol{\gamma} - \boldsymbol{\mu} + \lambda \mathbf{A}(t)$, i.e. the expression of the Jacobian obtained from the dynamical equations in the main paper. We note that the parameters γ_{bc} appearing in the general compartmental model (see Fig. 1a) are never diagonal: $b \neq c$. Therefore, the weighted adjacency matrix of the EGD that appears in Eq. (10) is

$$W_{bc}(t) = \gamma_{bc}(1 - \delta_{bc}) - \mu_b \delta_{bc} + \lambda_{bc} A(t), \tag{11}$$

where we made explicit the off-diagonal nature of γ_{bc} . The diagonal elements of operator $\boldsymbol{\gamma}$ in Eq. (3) are not parameters of the model, they are fixed by probability conservation, i.e. $\gamma_{bb} = -\sum_{d \neq b} \gamma_{bd}$. We therefore need to add this last term to the weighted adjacency matrix of the EGD to account for the diagonal entries of operator $\boldsymbol{\gamma}$ of Eq. (3), and this explains the nature of the second term in Eq. (10). This proves that the Jacobian built from the EGD is exactly the operator $\mathbf{J}(t)$ of the infection propagator.

Diagram operations

We prove here the diagram operations described in the paper, and in Fig. 2. They are *CUT*, *ZIP*, and *SHRINK*. *CUT* decomposes EGDs in smaller subdiagrams using network approaches and allows the computation of the threshold on each subdiagram separately. *ZIP* compresses complex diseases progressions into SIS-like diagrams. *SHRINK* merges pair of nodes. *ZIP* exploits and transforms the global topology of the diagram, while *SHRINK* acts locally on pairs of compartments.

Proof of CUT. An EGD is a directed graph, and it is possible to order its nodes so that its adjacency matrix is block-upper-triangular, with the blocks representing the strongly connected components. \mathbf{J} inherits the same property, if intended as the supra-adjacency matrix^{46,60} of the multilayer structure represented by the EGD, i.e., the adjacency matrix of the flattened graph exploiting $\mathbb{R}^N \otimes \mathbb{R}^{N_C} \simeq \mathbb{R}^{NN_C}$. And so does \mathbf{P} , because it is a convolution of $\mathbf{J}(t)$ at different times. It follows that it is possible to compute the spectrum of \mathbf{P} as the union of the spectra of the blocks in the main diagonal, which are the subdiagrams obtained via *CUT*.

Proof of ZIP. We use the notation \mathcal{EGD} for the mathematical representation of the diagram. Let us assume that weak-commutation condition holds ($\mathbf{A}(t) \equiv \bar{\mathbf{A}}$), and the epidemic graph diagram \mathcal{EGD} is made of a node D and a subdiagram \mathcal{EGD}_X . Moreover, let us assume that \mathcal{EGD}_X contains only single links (its Jacobian is $\boldsymbol{\gamma}_X - \boldsymbol{\mu}_X$). Single links may exist from D , to any node in \mathcal{EGD}_X : We call $\gamma_r^{(+)}$ the weight of the single link from D to the r -th node of \mathcal{EGD}_X . Single links may also exist from any node in \mathcal{EGD}_X to D : We call $\gamma_r^{(-)}$ the weight of the single link

from the r -th node of $\mathcal{E}GD_X$ to D . The index r runs on the $N_c - 1$ nodes of $\mathcal{E}GD_X$. D may have both single and double self loops, whose weights we call $-\mu_0$ and $\lambda_0 \bar{\mathbf{A}}$, respectively. Double links may exist from any node in $\mathcal{E}GD_X$ to D (λ_r). The resulting \mathbf{J} of the full EGD is

$$\mathbf{J} = \left(\begin{array}{ccc|ccc} & & & \gamma_1^{(-)} + \lambda_1 \bar{\mathbf{A}} & & \\ & & & \vdots & & \\ & \gamma_X - \mu_X - \mathbf{d} & & \gamma_{N_c-1}^{(-)} + \lambda_{N_c-1} \bar{\mathbf{A}} & & \\ \hline \gamma_1^{(+)} & \dots & \gamma_{N_c-1}^{(+)} & -\mu_0 - \sum_r \gamma_r^{(+)} + \lambda_0 \bar{\mathbf{A}} & & \end{array} \right), \quad (12)$$

with $\mathbf{d} = \text{diag}(\gamma_1^{(-)}, \dots, \gamma_{N_c-1}^{(-)})$. The threshold condition in timescale separation is $\det \mathbf{J} = 0$. Using block matrix determinant rules, this simplifies to a determinant in \mathbb{R}^M :

$$\det[-\mu_{eff} + \lambda_{eff} \bar{\mathbf{A}}] = 0, \quad (13)$$

with

$$\mu_{eff} = \mu_0 + \sum_r \gamma_r^{(+)} + \sum_{rs} \gamma_r^{(+)} \gamma_s^{(-)} \left[(\gamma_X - \mu_X - \mathbf{d})^{-1} \right]_{rs}; \quad (14)$$

$$\lambda_{eff} = \lambda_0 - \sum_{rs} \gamma_r^{(+)} \lambda_s \left[(\gamma_X - \mu_X - \mathbf{d})^{-1} \right]_{rs}. \quad (15)$$

$\gamma_X - \mu_X - \mathbf{d}$ is always invertible if we assume that the model is below the epidemic threshold in the absence of transmission. If it were not the case, the epidemic would be a trivial spontaneous generation of infected individuals. Or, in epidemiological terms, there would be only primary disease introductions. This means that $\mathbf{J}(\lambda_0 = 0, \lambda_r = 0)$ is negative definite. By virtue of Sylvester's criterion, this in turn implies that $\gamma_X - \mu_X - \mathbf{d}$ is also negative definite, because all the leading principal minors of the latter are also leading principal minors of \mathbf{J} . And since $\gamma_X - \mu_X$ does not depend on transmission, $\gamma_X - \mu_X - \mathbf{d}$ is always negative definite, and so invertible. We complete the proof of ZIP by noting that Eq. (13) is equivalent to an SIS model with renormalized recovery and transmission rates μ_{eff}, λ_{eff} .

Proof of SHRINK. The proof is conceptually similar to ZIP's and, as for ZIP, requires that the weak-commutation condition holds so that the time-evolving Jacobian is replaced by its average $\bar{\mathbf{J}}$. Let us assume that the epidemic graph diagram is composed of compartment B and a subgraph $\mathcal{E}GD_X$, with Jacobian $\gamma_X - \mu_X + \lambda_X \bar{\mathbf{A}}$. Moreover, let us assume that a single link γ_0 goes from B to a node in $\mathcal{E}GD_X$, which we label as the first ($r=1$). B may have no other outgoing link, and no self loops. Any node in $\mathcal{E}GD_X$ may have single and double links going to B . The Jacobian is

$$\mathbf{J} = \left(\begin{array}{ccc|ccc} & & & \gamma_1 + \lambda_1 \bar{\mathbf{A}} & & \\ & & & \vdots & & \\ & \gamma_X - \mu_X + \lambda_X \bar{\mathbf{A}} - \mathbf{d} & & \gamma_{N_c-1} + \lambda_{N_c-1} \bar{\mathbf{A}} & & \\ \hline \gamma_0 & 0 & \dots & 0 & -\gamma_0 & \end{array} \right), \quad (16)$$

with $\mathbf{d} = \text{diag}(\gamma_1, \dots, \gamma_{N_c-1})$. Using block matrix determinant rules, the condition $\det \mathbf{J} = 0$ is equivalent to the following $N_c - 1$ dimensional determinant

$$\det \left\{ \gamma_X - \mu_X + \lambda_X \bar{\mathbf{A}} - \mathbf{d} + \begin{pmatrix} \gamma_1 + \lambda_1 \bar{\mathbf{A}} & 0 & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ \gamma_{N_c-1} + \lambda_{N_c-1} \bar{\mathbf{A}} & 0 & \dots & 0 \end{pmatrix} \right\} = 0. \quad (17)$$

This is equivalent to a diagram in which B disappears, its outgoing link γ_0 disappears, and all incoming links of B get rerouted onto the ancient target of link γ_0 .

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The sexual contact network is publicly available as Supplementary Information of ref. 23.

Code availability

The code used in this study is public⁶¹ and available at <https://github.com/eugenio-valdano/egd>.

References

1. Koelle, K., Martin, M. A., Antia, R., Lopman, B. & Dean, N. E. The changing epidemiology of SARS-CoV-2. *Science* **375**, 1116–1121 (2022).
2. Carlson, C. J. et al. Climate change increases cross-species viral transmission risk. *Nature* **607**, 555–562 (2022).
3. Horby, P. Improving preparedness for the next flu pandemic. *Nat. Microbiol.* **3**, 848–850 (2018).
4. WHO - Ten Threats to Global Health in 2019. <https://www.who.int/news-room/spotlight/ten-threats-to-global-health-in-2019> visited on Sept 2022.
5. Colizza, V., Barrat, A., Barthélemy, M., Valleron, A.-J. & Vespignani, A. Modeling the worldwide spread of Pandemic Influenza: Baseline case and containment interventions. *PLoS Med.* **4**, e13 (2007).
6. Wu, J. T., Leung, G. M., Lipsitch, M., Cooper, B. S. & Riley, S. Hedging against antiviral resistance during the next influenza pandemic using small stockpiles of an alternative chemotherapy. *PLoS Med.* **6**, e1000085 (2009).
7. In Danger: UNAIDS Global AIDS Update 2022376 (2022).
8. WHO | global health sector strategies on, respectively, hiv, viral hepatitis and sexually transmitted infections for the period 2022-2030. <https://www.who.int/teams/global-hiv-hepatitis-and-stis-programmes/strategies/global-health-sector-strategies> visited on Sept 2022.
9. Tien, V., Punjabi, C. & Holubar, M. K. Antimicrobial resistance in sexually transmitted infections. *J. Travel Med.* **27**, taz101 (2020).
10. Thornhill, J. P. et al. Monkeypox Virus Infection in Humans across 16 Countries - April-June 2022. *N. Engl. J. Med.* **0**, null (2022).
11. Heesterbeek, Ja. P. & Dietz, K. The concept of Ro in epidemic theory. *Statistica Neerlandica* **50**, 89–110 (1996).
12. Pastor-Satorras, R., Castellano, C., Van Mieghem, P. & Vespignani, A. Epidemic processes in complex networks. *Rev. Mod. Phys.* **87**, 925–979 (2015).
13. Riley, S. et al. Transmission dynamics of the etiological agent of SARS in Hong Kong: Impact of public health interventions. *Science* **300**, 1961–1966 (2003).
14. Fraser, C. et al. Pandemic potential of a strain of influenza A (H1N1): early findings. *Science (New York, N.Y.)* **324**, 1557–1561 (2009).
15. Darbon, A. et al. Network-based assessment of the vulnerability of Italian regions to bovine brucellosis. *Prevent. Veterinary Med.* **158**, 25–34 (2018).
16. Flaxman, S. et al. Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. *Nature* **584**, 257–261 (2020).
17. Di Domenico, L., Pullano, G., Sabbatini, C. E., Boëlle, P.-Y. & Colizza, V. Impact of lockdown on COVID-19 epidemic in île-de-France and possible exit strategies. *BMC Med.* **18**, 240 (2020).
18. Badr, H. S. et al. Association between mobility patterns and COVID-19 transmission in the USA: a mathematical modelling study. *Lancet Infect. Dis.* **20**, 1247–1254 (2020).

19. Nouvellet, P. et al. Reduction in mobility and COVID-19 transmission. *Nat. Commun.* **12**, 1090 (2021).
20. van den Driessche, P. & Watmough, J. Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. *Math. Biosci.* **180**, 29–48 (2002).
21. Diekmann, O., Heesterbeek, Ja. P. & Roberts, M. G. The construction of next-generation matrices for compartmental epidemic models. *J. R. Soc. Interface* **7**, 873–885 (2010).
22. Diekmann, O., Heesterbeek, H. & Britton, T. *Mathematical Tools for Understanding Infectious Disease Dynamics* (Princeton University Press, 2012). <https://doi.org/10.23943/princeton/9780691155395.001.0001>.
23. Rocha, L. E. C., Liljeros, F. & Holme, P. Simulated epidemics in an empirical spatiotemporal network of 50,185 sexual contacts. *PLoS Comput. Biol.* **7**, e1001109 (2011).
24. Stehlé, J. et al. Simulation of an SEIR infectious disease model on the dynamic contact network of conference attendees. *BMC Med.* **9** <http://www.biomedcentral.com/1741-7015/9/87> (2011).
25. Robineau, O., Velter, A., Barin, F. & Boelle, P.-Y. HIV transmission and pre-exposure prophylaxis in a high risk MSM population: A simulation study of location-based selection of sexual partners. *PLOS ONE* **12**, e0189002 (2017).
26. Wang, Y., Chakrabarti, D., Wang, C. & Faloutsos, C. Epidemic spreading in real networks: an eigenvalue viewpoint. In *Reliable Distributed Systems, 2003. Proceedings. 22nd International Symposium on*, 25–34 (2003).
27. Volz, E. & Meyers, L. A. Epidemic thresholds in dynamic contact networks. *J. R. Soc. Interface* **6**, 233–241 (2009).
28. Gómez, S., Arenas, A., Borge-Holthoefer, J., Meloni, S. & Moreno, Y. Discrete time Markov chain approach to contact-based disease spreading in complex networks. *Europhys. Lett.* **89**, 38009 (2010).
29. Valdano, E., Ferreri, L., Poletto, C. & Colizza, V. Analytical computation of the epidemic threshold on temporal networks. *Phys. Rev. X* **5**, 21005 (2015).
30. Valdano, E., Fiorentin, M. R., Poletto, C. & Colizza, V. Epidemic threshold in continuous-time evolving networks. *Phys. Rev. Lett.* **120**, 068302 (2018).
31. Koher, A., Lentz, H. H. K., Gleeson, J. P. & Hövel, P. Contact-based model for epidemic spreading on temporal networks. *Phys. Rev. X* **9**, 31017 (2019).
32. Miller, J. C., Slim, A. C. & Volz, E. M. Edge-based compartmental modelling for infectious disease spread. *J. R. Soc. Interface* **9**, 890–906 (2011).
33. Valdano, E., Poletto, C. & Colizza, V. Infection propagator approach to compute epidemic thresholds on temporal networks: impact of immunity and of limited temporal resolution. *Eur. Phys. J. B* **88**, 1–11 (2015).
34. Masuda, N. & Konno, N. Multi-state epidemic processes on complex networks. *J. Theor. Biol.* **243**, 64–75 (2006).
35. Prakash, B. A., Chakrabarti, D., Faloutsos, M., Valler, N. & Faloutsos, C. Threshold Conditions for Arbitrary Cascade Models on Arbitrary Networks. In *Proceedings of the 2011 IEEE 11th International Conference on Data Mining, ICDM '11*, 537–546 (IEEE Computer Society, Washington, DC, USA, 2011). <https://doi.org/10.1109/ICDM.2011.145>.
36. Guo, H., Li, M. & Shuai, Z. Global dynamics of a general class of multistage models for infectious diseases. *SIAM J. Appl. Math.* **72**, 261–279 (2012).
37. Miller, J. C. & Volz, E. M. Incorporating disease and population structure into models of SIR disease in contact networks. *PLOS ONE* **8**, e69162 (2013).
38. Menzies, N. A. et al. Progression from latent infection to active disease in dynamic tuberculosis transmission models: A systematic review of the validity of modelling assumptions. *Lancet Infect. Dis.* **18**, e228–e238 (2018).
39. Chang, S. et al. Mobility network models of COVID-19 explain inequities and inform reopening. *Nature* **589**, 82–87 (2021).
40. Pullano, G. et al. Underdetection of cases of COVID-19 in France threatens epidemic control. *Nature* **590**, 134–139 (2021).
41. Verelst, F. et al. SOCRATES-CoMix: a platform for timely and open-source contact mixing data during and in between COVID-19 surges and interventions in over 20 European countries. *BMC Med.* **19**, 254 (2021).
42. Wymant, C. et al. The epidemiological impact of the NHS COVID-19 app. *Nature* **594**, 408–412 (2021).
43. 't Hooft, G. & Veltman, M. J. G. *Diagrammar*. CERN Yellow Reports: Monographs (CERN, Geneva, 1973). <https://cds.cern.ch/record/186259>.
44. Do, A.-L., Boccaletti, S. & Gross, T. Graphical notation reveals topological stability criteria for collective dynamics in complex networks. *Phys. Rev. Lett.* **108** (2012).
45. Bansal, S., Grenfell, B. T. & Meyers, L. A. When individual behaviour matters: Homogeneous and network models in epidemiology. *J. R. Soc. Interface* **4**, 879–891 (2007).
46. De Domenico, M. et al. Mathematical formulation of multilayer networks. *Phys. Rev. X* **3**, 41022 (2013).
47. Garnett, G., Aral, S., Hoyle, D., Cates, W. & Anderson, R. The natural history of syphilis: implications for the transmission dynamics and control of infection. *Sexually Transmitted Dis.* **24**, 185–200 (1997).
48. Eames, K. T. D. & Keeling, M. J. Modeling dynamic and network heterogeneities in the spread of sexually transmitted diseases. *Proc. Natl. Acad. Sci.* **99**, 13330–13335 (2002).
49. Grassly, N. C., Fraser, C. & Garnett, G. P. Host immunity and synchronized epidemics of syphilis across the United States. *Nature* **433**, 417–421 (2005).
50. Keeling, M. J. & Rohani, P. *Modeling infectious diseases in humans and animals*. Princeton University Press (2007).
51. Perra, N., Gonçalves, B., Pastor-Satorras, R. & Vespignani, A. Activity driven modeling of time varying networks. *Sci. Rep.* **2** (2012). <https://doi.org/10.1038/srep00469>, <http://www.nature.com/srep/2012/120625/srep00469/abs/srep00469.html#supplementary-information>.
52. Newman, M. E. J. Spread of epidemic disease on networks. *Phys. Rev. E* **66**, 16128 (2002).
53. Colman, E., Spies, K. & Bansal, S. The reachability of contagion in temporal contact networks: How disease latency can exploit the rhythm of human behavior. *BMC Infect. Dis.* **18**, 219 (2018).
54. Leung, K., Lipsitch, M., Yuen, K. Y. & Wu, J. T. Monitoring the fitness of antiviral-resistant influenza strains during an epidemic: A mathematical modeling study. *Lancet Infect. Dis.* **17**, 339–347 (2017).
55. Gleeson, J. P. et al. Calibrating COVID-19 susceptible-exposed-infected-removed models with time-varying effective contact rates. *Philos. Trans. R. Soc. A: Math. Phys. Eng. Sci.* **380**, 20210120 (2021).
56. Gleeson, J. P., Melnik, S., Ward, J. A., Porter, M. A. & Mucha, P. J. Accuracy of mean-field theory for dynamics on real-world networks. *Phys. Rev. E* **85**, 026106 (2012).
57. Tian, J. P. & Wang, J. Some results in Floquet theory, with application to periodic epidemic models. *Appl. Anal.* 1–25. <http://www.tandfonline.com/doi/abs/10.1080/00036811.2014.918606> (2014).
58. Dyson, F. J. The Radiation theories of Tomonaga, Schwinger, and Feynman. *Phys. Rev.* **75**, 486–502 (1949).
59. Speidel, L., Klemm, K., Eguiluz, V. M. & Masuda, N. Temporal interactions facilitate endemicity in the susceptible-infected-susceptible epidemic model. *N. J. Phys.* **18**, 073013 (2016).
60. Cozzo, E., Baños, R. A., Meloni, S. & Moreno, Y. Contact-based social contagion in multiplex networks. *Phys. Rev. E* **88**, 50801 (2013).
61. Valdano, E. Epidemic graph diagrams as analytics for epidemic control in the data-rich era, source code. <https://doi.org/10.5281/zenodo.10012989>, <https://github.com/eugenio-valdano/egd> (2023).

Acknowledgements

This study was partially supported by: ANR grant DATAREDUX (ANR-19-CE46-0008-03) to V.C.; ANRS-MIE grant MPX-SPREAD (ANRS0292) to V.C.; EU Horizon 2020 grant MOOD (H2020-874850, paper 83) to E.V., C.P., V.C.; Horizon Europe grants ESCAPE (101095619) and VERDI (101045989) to V.C.. We thank Mason Porter and Alain Barrat for useful discussions.

Author contributions

E.V., C.P., V.C. conceived and designed the study. E.V. developed the theory, developed the code, and analyzed the results. E.V., D.C., C.P., V.C. interpreted the results. E.V. drafted the Article. All authors contributed to and approved the final version of the Article.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-023-43856-1>.

Correspondence and requests for materials should be addressed to Vittoria Colizza.

Peer review information : *Nature Communications* thanks the anonymous reviewers for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023