Supplementary Materials for "Geometric correlations mitigate the extreme vulnerability of multiplex networks against targeted attacks"

Kaj-Kolja Kleineberg,[1,*] Lubos Buzna,[2] Fragkiskos Papadopoulos,[3] Marián Boguñá,[4] and M. Ángeles Serrano[5,4]

[1] *Computational Social Science, ETH Zurich, Clausiusstrasse 50, CH-8092 Zurich, Switzerland*
[2] *University of Zilina, Univerzitna 8215/1, SK-01026 Zilina, Slovakia*
[3] *Department of Electrical Engineering, Computer Engineering and Informatics, Cyprus University of Technology, 33 Saripolou Street, 3036 Limassol, Cyprus*
[4] *Departament de Física Fonamental, Universitat de Barcelona, Martí i Franquès 1, 08028 Barcelona, Spain*
[5] *Institució Catalana de Recerca i Estudis Avançats (ICREA), Passeig Lluís Companys 23, E-08010 Barcelona, Spain*

(Dated: April 3, 2017)

CONTENTS

I. DATASETS

As explained in the main text, the increased robustness of real multiplexes to targeted attacks can be explained by hidden geometric correlations interwoven in their layers [1]. These correlations are called "hidden" because they are not directly observable by looking at the topology of each individual network. Specifically, each single network layer can be mapped (or embedded) into a separate hyperbolic space, where each node i is represented by its polar coordinates, (r_i, θ_i) [2–4]. The radial coordinate r_i abstracts node popularity. The angular distance between two nodes, $\Delta\theta_{ij} = \pi - |\pi - |\theta_i - \theta_j||$, abstracts their similarity [5]. The hyperbolic distance, $x_{ij} = \cosh^{-1}(\cosh r_i \cosh r_j - \sinh r_i \sinh r_j \cos \Delta\theta_{ij})$, is then a metric combination of the two attractiveness attributes, popularity (radial) and similarity (angular), such that the smaller the hyperbolic distance between two nodes, the higher the probability that they are connected in the observable network [6]. The node coordinates of a given real network can be inferred using Maximum Likelihood Estimation techniques [2–4]. It has been shown that both the radial and angular coordinates of nodes in different layers of real multiplexes are significantly correlated [1].

* kkleineberg@ethz.ch

Dataset	Abbreviation	MCC	$\bar{c}_1$	$\bar{c}_2$	NMI	ΔN	ΔN_{rs}
arXiv Layers 1, 2	Arx12	790	0.83	0.81	0.58	25.2	1.0
arXiv Layers 4, 2	Arx42	916	0.83	0.81	0.49	43.8	2.1
arXiv Layers 4, 1	Arx41	564	0.83	0.83	0.46	16.6	1.0
arXiv Layers 2, 8	Arx28	521	0.81	0.85	0.42	15.8	1.0
Physicians Layers 1, 2	Phys12	104	0.28	0.28	0.41	6.0	1.0
arXiv Layers 5, 2	Arx52	182	0.89	0.81	0.41	16.8	1.0
arXiv Layers 1, 5	Arx15	100	0.83	0.89	0.39	9.0	1.0
arXiv Layers 2, 6	Arx26	83	0.81	0.88	0.38	10.0	2.0
Internet Layer 1, 2	Internet	4710	0.63	0.55	0.34	81.4	14.1
arXiv Layers 3, 4	Arx34	26	0.92	0.83	0.34	3.0	1.0
C. Elegans Layers 2, 3	CE23	257	0.21	0.29	0.34	14.0	1.1
Physicians Layers 1, 3	Phys13	99	0.28	0.24	0.32	1.0	1.0
Physicians Layers 2, 3	Phys23	106	0.28	0.24	0.31	2.0	1.0
SacchPomb Layers 1, 3	Sac13	510	0.18	0.42	0.31	18.3	1.0
SacchPomb Layers 3, 5	Sac35	54	0.42	0.64	0.31	1.0	1.0
SacchPomb Layers 2, 3	Sac23	27	0.28	0.42	0.29	2.0	1.0
SacchPomb Layers 1, 2	Sac12	32	0.18	0.28	0.29	1.0	1.0
Drosophila Layers 1, 2	Dro12	449	0.28	0.29	0.26	8.4	2.0
C. Elegans Layers 1, 3	CE13	247	0.24	0.29	0.21	8.2	2.3
SacchPomb Layers 1, 4	Sac14	289	0.18	0.26	0.19	3.5	1.0
SacchPomb Layers 2, 4	Sac24	25	0.28	0.26	0.19	2.0	1.0
Brain Layers 1, 2	Brain	74	0.49	0.40	0.19	7.0	1.0
Rattus Layers 1, 2	Rattus	158	0.26	0.05	0.18	4.0	1.0
C. Elegans Layers 1, 2	CE12	226	0.24	0.21	0.17	3.0	2.0
SacchPomb Layers 3, 4	Sac34	426	0.42	0.26	0.17	4.2	1.5
Air/Train Layers 1, 2	AirTrain	67	0.79	0.48	0.10	3.0	3.0

Table I. Overview of the considered real-world multiplex network data. For each layer pair, MCC denotes the initial size of its mutually connected component and NMI is the normalized mutual information as calculated in [1]. $\bar{c}_{1,2}$ denotes the mean local clustering coefficient in the first or second layer considered. ΔN denotes the critical number of nodes for the real layers while ΔN_{rs} is the same number for their reshuffled counterparts.

Supp. Tab. I gives an overview of the considered real-world multiplex network data. More details on the data and on the topological characteristics of each layer can be found in [1]. For each layer pair we report the initial size of its mutually connected component (MCC), the normalized mutual information NMI as calculated in [1], which gives a measure of the strength of angular correlations between its layers (see Sec. IX), and the critical number of nodes ΔN whose removal reduces the MCC from 40% its initial value to the square root of it. The table also shows the corresponding critical number of nodes for the reshuffled counterparts of the real systems, ΔN_{rs}. As explained in the main text, to create the reshuffled counterparts, we randomly reshuffle the trans-layer node-to-node mappings. Specifically, for each real multiplex we select one of its layers and we interchange the ID of each node of the layer with the ID of a randomly selected node from the same layer. Thus, if a node with ID i is node n_1 in layer 1 and node n_2 in layer 2 with correlated coordinates (r_{n_1}, θ_{n_1}), (r_{n_2}, θ_{n_2}), then, after reshuffling layer 2, the node will become some other node n_2' in this layer, with coordinates $(r_{n_2'}, \theta_{n_2'})$ that will not be correlated with (r_{n_1}, θ_{n_1}). The reshuffled counterparts serve as a null model for what one would expect if there were no geometric correlations among the layers.

II. MUTUAL CONNECTIVITY TRANSITION FOR THE CONSIDERED REAL SYSTEMS

Supp. Fig. 1 shows the evolution of the MCC for all real systems of Supp. Tab. I and for their reshuffled counterparts, under the targeted attacks process described in the main text.

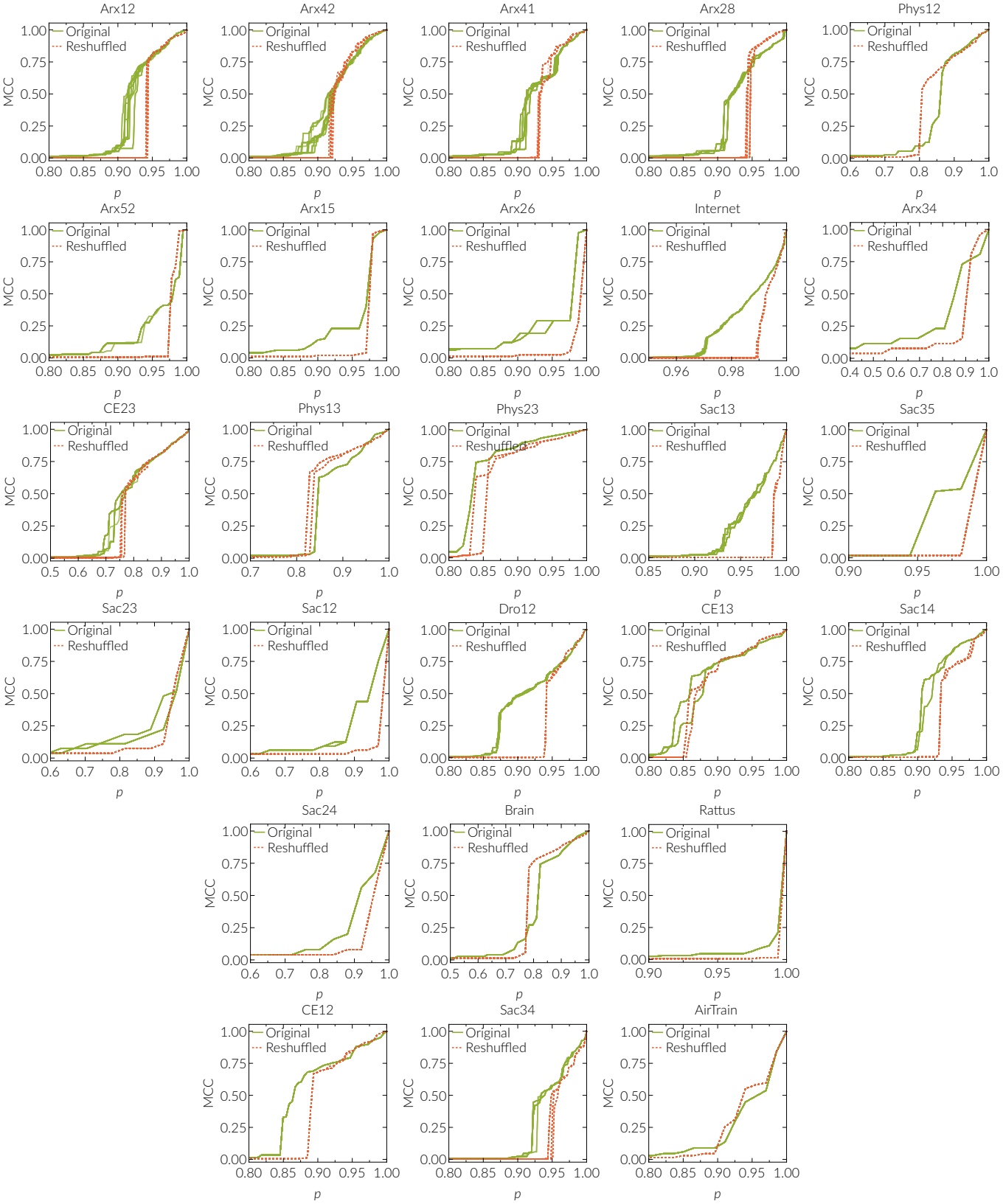

Figure 1. MCC for real systems (solid green lines) and for their reshuffled counterparts (dashed red lines). The plot titles are the abbreviations listed in Supp. Tab. I. The plots show the relative size of the MCC as a function of the fraction of nodes remaining in the system p. Different lines in each plot correspond to different realizations of the targeted attacks process.

III. MUTUAL CONNECTIVITY TRANSITION FOR MODEL NETWORKS

A. Description of the geometric multiplex model (GMM)

In this section we present the technical details of the geometric multiplex model (GMM), developed in [1][7]. The model constructs synthetic multiplex networks with geometric correlations. Specifically, it constructs single-layer topologies using the $\mathbb{H}^2$ model [8], and allows for radial and angular coordinate correlations across the different layers. Instead of working directly with the $\mathbb{H}^2$ model, the GMM model makes use of the $\mathbb{S}^1$ model [5] that is more convenient to work with, and which is isomorphic to the $\mathbb{H}^2$ model through a simple change of variables [8]. We first review the $\mathbb{S}^1$ model and its relation to the $\mathbb{H}^2$ model.

Instead of radial and angular coordinates r_i, θ_i, each node i in the $\mathbb{S}^1$ model has hidden variables κ_i, θ_i. The hidden variable κ_i is the node's expected degree in the resulting network, while θ_i is the angular (similarity) coordinate of the node on a circle of radius $N/2\pi$, where N is the total number of nodes. To construct a network with the $\mathbb{S}^1$ model that has size N, average node degree $\bar{k}$, power law degree distribution with exponent $\gamma > 2$, and temperature $T \in [0, 1)$ (which controls clustering in the network), we perform the following steps:

i. Sample the angular coordinates of nodes θ_i, $i = 1, 2, \ldots, N$, uniformly at random from $[0, 2\pi]$, and their hidden variables κ_i, $i = 1, 2, \ldots, N$, from the probability density function (PDF)

$$\begin{aligned} \rho(\kappa) &= (\gamma - 1)\kappa^{\min^{\gamma-1}}\kappa^{-\gamma}, \\ \kappa^{\min} &= \bar{k}\frac{\gamma - 2}{\gamma - 1}, \end{aligned} \tag{1}$$

where $\kappa^{\min}$ is the expected minimum node degree, which is a function of the average degree $\bar{k}$; [9]

ii. Connect every pair of nodes i, j with probability

$$\begin{aligned} r(\kappa_i, \theta_i; \kappa_j, \theta_j) &= \frac{1}{1 + \left[\frac{d(\theta_i, \theta_j)}{\mu\kappa_i\kappa_j}\right]^{\frac{1}{T}}}, \\ d(\theta_i, \theta_j) &= \frac{N}{2\pi}\Delta\theta_{ij},\ \Delta\theta_{ij} = |\pi - |\pi - |\theta_i - \theta_j|||, \\ \mu &= \frac{\sin T\pi}{2\bar{k}T\pi}, \end{aligned} \tag{2}$$

where $d(\theta_i, \theta_j)$ is the angular distance between nodes i, j on the circle.

The $\mathbb{S}^1$ model is equivalent to the $\mathbb{H}^2$ model after transforming the expected node degrees κ_i to radial coordinates r_i via

$$r_i = R - 2\ln\frac{\kappa_i}{\kappa^{\min}}, \tag{3}$$

where R is the radius of the hyperbolic disc in the $\mathbb{H}^2$ model where all nodes reside,

$$\begin{aligned} R &= 2\ln\frac{N}{c}, \\ c &= \bar{k}\frac{\sin T\pi}{2T}\left(\frac{\gamma - 2}{\gamma - 1}\right)^2. \end{aligned} \tag{4}$$

It is easy to see that after the above change of variables the connection probability in Eq. (2) becomes the Fermi-Dirac connection probability in the $\mathbb{H}^2$ model,

$$p(x_{ij}) = \frac{1}{1 + e^{\frac{1}{2T}(x_{ij} - R)}}, \tag{5}$$

where $x_{ij} \approx r_i + r_j + 2\ln\frac{\Delta\theta_{ij}}{2}$ is the hyperbolic distance between nodes i, j [8]. We note that without loss of generality, we use here a hyperbolic plane of curvature $K = -1$. See [8] for further details.

We now describe the framework for constructing a two-layer multiplex system with geometric correlations. Each single-layer (layer 1, layer 2) is constructed according to the $\mathbb{S}^1$ model, and we account for correlations among the hidden variables of nodes in the two layers, whose strength can be tuned. In a nutshell, the framework consists of the following steps:

i. Assignment of hidden variables $\kappa_{1,i}, \theta_{1,i}$ to each node i in layer 1 like in the $\mathbb{S}^1$ model (Eqs. (6), (8));

ii. Assignment of hidden variables $\kappa_{2,i}, \theta_{2,i}$ to each node i in layer 2, depending on the node's hidden variables in layer 1 (Eqs. (9), (13))—the assignment here is done such that the marginal (unconditional) distribution of $\kappa_{2,i}, \theta_{2,i}$ is still the one prescribed by the $\mathbb{S}^1$ model (Eqs. (11), (8));

iii. Creation of edges, by connecting node pairs in each layer with the corresponding $\mathbb{S}^1$ connection probability, which depends exclusively on the assigned hidden variables of nodes in each layer (Eqs. (15), (16));

iv. $\mathbb{S}^1$-to-$\mathbb{H}^2$ transformation, by mapping the hidden variables $\kappa_{1,i}, \kappa_{2,i}$ to radial coordinates $r_{1,i}, r_{2,i}$ (Eqs. (17), (18)).

Below, we describe these steps in detail. We assume that the two layers have the same number of nodes $N_1 = N_2 = N$.

i. Assignment of hidden variables in layer 1. For each node $i = 1, 2, \ldots, N$ in layer 1 we sample its hidden variable $\kappa_{1,i}$ from the PDF

$$\rho_1(\kappa_1) = (\gamma_1 - 1)\kappa_1^{\min\gamma_1 - 1}\kappa_1^{-\gamma_1}, \tag{6}$$

$$\kappa_1^{\min} = \bar{k}_1\frac{\gamma_1 - 2}{\gamma_1 - 1}, \tag{7}$$

where $\bar{k}_1$ and $\gamma_1 > 2$ are respectively the target average degree and power law degree distribution exponent in layer 1. The angular coordinate $\theta_{1,i}$ of each node $i = 1, 2, \ldots, N$, is sampled from the uniform PDF

$$f(\theta) = \frac{1}{2\pi},\ \theta \in [0, 2\pi]. \tag{8}$$

ii. Assignment of hidden variables in layer 2. We now want to assign to each node $i = 1, 2, \ldots, N$ its hidden variable $\kappa_{2,i}$ in layer 2, conditioned on the value of its hidden variable $\kappa_{1,i}$ in layer 1. In particular, we sample the hidden variable $\kappa_{2,i}$ of each node $i = 1, 2, \ldots, N$, from the conditional cumulative distribution function (CDF)

$$F_\nu(\kappa_2|\kappa_1, \{\gamma_1, \gamma_2, \kappa_1^{\min}, \kappa_2^{\min}\}) = e^{-(\varphi_1^{1/(1-\nu)} + \varphi_2^{1/(1-\nu)})^{1-\nu}}\left[\varphi_1^{1/(1-\nu)} + \varphi_2^{1/(1-\nu)}\right]^{-\nu}\frac{\varphi_1^{\nu/(1-\nu)}\kappa_1^{\min}\kappa_1^{\gamma_1}}{\kappa_1^{\min}\kappa_1^{\gamma_1} - \kappa_1^{\min\gamma_1}\kappa_1}, \tag{9}$$

$$\varphi_i = -\ln\left[1 - (\kappa_i^{\min}/\kappa_i)^{\gamma_i - 1}\right],\ \text{for } i = 1, 2, \tag{10}$$

where κ_1 is the value of the hidden variable of the node in layer 1, $\{\gamma_1, \gamma_2, \kappa_1^{\min}, \kappa_2^{\min}\}$ are the network parameters defined earlier, and $\nu \in [0, 1]$ is the correlation strength parameter. The higher the value of ν the stronger is the correlation between $\kappa_{2,i}$ and $\kappa_{1,i}$. Eq. (9) ensures that in the second layer $\kappa_{2,i}$'s satisfy the marginal (unconditional) PDF

$$\rho_2(\kappa_2) = (\gamma_2 - 1)\kappa_2^{\min\gamma_2 - 1}\kappa_2^{-\gamma_2}, \tag{11}$$

$$\kappa_2^{\min} = \bar{k}_2\frac{\gamma_2 - 2}{\gamma_2 - 1}, \tag{12}$$

where $\bar{k}_2$ and $\gamma_2 > 2$ are respectively the target average degree and power law degree distribution exponent in layer 2. See [1] for the derivation of Eq. (9).

The angular coordinate $\theta_{2,i}$ of each node $i = 1, 2, \ldots, N$ in layer 2 is obtained by

$$\theta_{2,i} = \mod\left[\theta_{1,i} + \frac{2\pi l_i}{N}, 2\pi\right], \tag{13}$$

where $\theta_{1,i}$ is the angular coordinate of the node in layer 1, and l_i is a directed arc length on the $\mathbb{S}^1$ circle of radius $R = N/2\pi$, which is sampled from the zero-mean truncated Gaussian PDF

$$f_\sigma(l) = \frac{\frac{1}{\sigma}\phi\left(\frac{l}{\sigma}\right)}{\Phi\left(\frac{N}{2\sigma}\right) - \Phi\left(-\frac{N}{2\sigma}\right)},\quad -\frac{N}{2} \le l \le \frac{N}{2}, \tag{14}$$

$$\sigma \equiv \sigma_0\left(\frac{1}{g} - 1\right),$$

where $\sigma_0 = \min[100, N/(4\pi)]$ denotes the variance for $g = 0.5$ and $\phi(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}$, $\Phi(x) = \int \mathrm{d}x\,\phi(x)$, $\sigma \in (0, \infty)$ is the variance of the PDF, and $g \in [0, 1]$ is the angular correlation strength parameter. The higher the value of g the

stronger is the correlation between $\theta_{2,i}$ and $\theta_{1,i}$. When $g \to 0$, $\sigma \to \infty$, $f_\sigma(l)$ becomes the uniform PDF, and $\theta_{2,i}$, $\theta_{1,i}$ are not correlated. When $g = 1$, $\sigma = 0$, and $l_i = 0$, meaning that the angles of each node i are identical in the two layers, $\theta_{2,i} = \theta_{1,i}$.

iii. Creation of edges. Once all node hidden variables are assigned, we connect each node pair i, j in layers 1 and 2 with the corresponding $\mathbb{S}^1$ connection probabilities given in Eqs. (15), (16) below,

$$
\begin{aligned}
r_1(\kappa_{1,i}, \theta_{1,i}; \kappa_{1,j}, \theta_{1,j}) &= \frac{1}{1 + \left[\frac{d_1(\theta_{1,i}, \theta_{1,j})}{\mu_1 \kappa_{1,i} \kappa_{1,j}}\right]^{\frac{1}{T_1}}}, \qquad (15)\\
d_1(\theta_{1,i}, \theta_{1,j}) &= \frac{N}{2\pi} \Delta\theta_{1,ij},\\
\Delta\theta_{1,ij} &= |\pi - |\pi - |\theta_{1,i} - \theta_{1,j}|||,\\
\mu_1 &= \frac{\sin T_1 \pi}{2\bar{k}_1 T \pi},\\
r_2(\kappa_{2,i}, \theta_{2,i}; \kappa_{2,j}, \theta_{2,j}) &= \frac{1}{1 + \left[\frac{d_2(\theta_{2,i}, \theta_{2,j})}{\mu_2 \kappa_{2,i} \kappa_{2,j}}\right]^{\frac{1}{T_2}}}, \qquad (16)\\
d_2(\theta_{2,i}, \theta_{2,j}) &= \frac{N}{2\pi} \Delta\theta_{2,ij},\\
\Delta\theta_{2,ij} &= |\pi - |\pi - |\theta_{2,i} - \theta_{2,j}|||,\\
\mu_2 &= \frac{\sin T_2 \pi}{2\bar{k}_2 T_2 \pi},
\end{aligned}
$$

where $T_1 \in [0, 1), T_2 \in [0, 1)$ are the temperatures, which control clustering in each layer. We recall that the average node clustering is maximized at temperature $T = 0$, and nearly linearly decreases to zero with $T \in [0, 1)$.

iv. $\mathbb{S}^1$*-to-*$\mathbb{H}^2$ *transformation.* Finally, we map the node hidden variables $\kappa_{1,i}, \kappa_{2,i}$ in layers 1, 2, to radial coordinates $r_{1,i}, r_{2,i}$ using the relations below,

$$
\begin{aligned}
r_{1,i} &= R_1 - 2 \ln \frac{\kappa_{1,i}}{\kappa_1^{\min}}, \; R_1 = 2 \ln \frac{N}{c_1}, \qquad (17)\\
c_1 &= \bar{k}_1 \frac{\sin T_1 \pi}{2T_1} \left(\frac{\gamma_1 - 2}{\gamma_1 - 1}\right)^2,\\
r_{2,i} &= R_2 - 2 \ln \frac{\kappa_{2,i}}{\kappa_2^{\min}}, \; R_2 = 2 \ln \frac{N}{c_2}, \qquad (18)\\
c_2 &= \bar{k}_2 \frac{\sin T_2 \pi}{2T_2} \left(\frac{\gamma_2 - 2}{\gamma_2 - 1}\right)^2,
\end{aligned}
$$

where $\kappa_1^{\min}, \kappa_2^{\min}$ are given in Eqs. (7), (12).

B. Results

Supp. Fig. 2 shows the evolution of the MCC for synthetic two-layer multiplexes generated by the geometric multiplex model (GMM) for different sizes and correlations strengths.

IV. DISTRIBUTION OF COMPONENT SIZES IN MODEL NETWORKS

In the case of no angular correlations ($g = 0$), where we observe a discontinuous percolation transition, we find that below a critical percentage of nodes remaining in the system, henceforth called a critical point, the layers are completely fragmented, and there are only very small mutually connected components of size at most 2. Above the critical point, we observe a macroscopic MCC and only nodes that do not belong to the MCC remain fragmented into small components, see Supp. Fig. 3.

The case where there are angular correlations is very different. As can be seen from Supp. Fig. 4, as p increases and approaches a certain value, which in the example of Supp. Fig. 4 is $p \approx 0.908$, the distribution of mutually component sizes becomes more and more heterogeneous, and at $p \approx 0.908$ it resembles a power law.

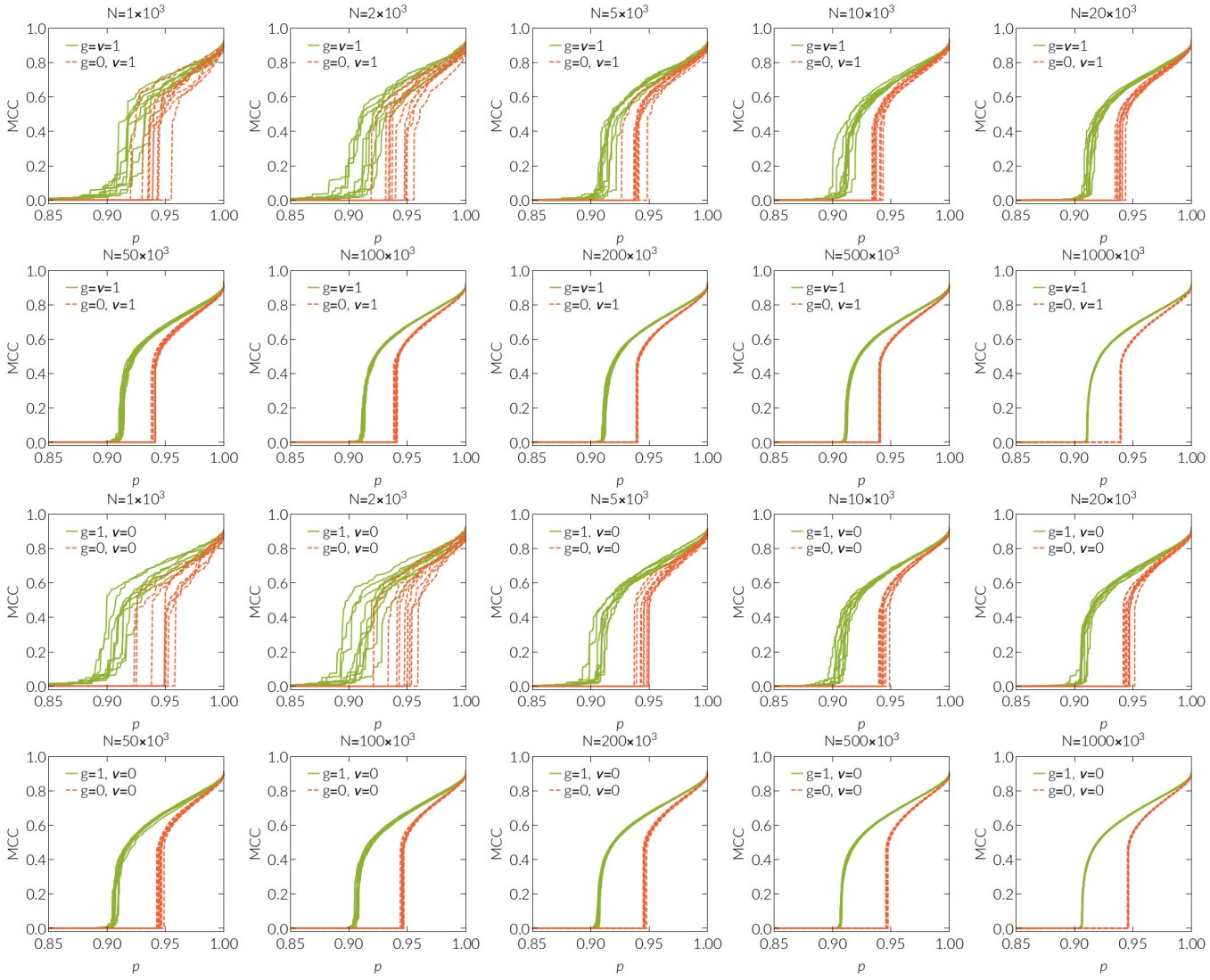

Figure 2. Evolution of the MCC for two-layer synthetic multiplexes of different sizes constructed with the GMM model. In each plot the layers have the same size N as indicated in the title, power law degree distribution with exponent $\gamma = 2.6$, average node degree $\langle k \rangle \approx 6$, and clustering $\bar{c} = 0.35$. Results are shown for different combinations of angular and radial correlation strengths g and ν, as indicated in the legends.

V. SCALING OF THE CRITICAL NUMBER OF NODES ΔN

Supp. Fig. 5 shows the critical number of nodes, ΔN, whose removal reduces the MCC from 40% of its initial value to the square root of it, as a function of the system size N. Each point in the plots corresponds to a two-layer synthetic multiplex constructed as in Fig. 2e of the main text, and is an average over 60 realizations of the multiplex. We consider different angular correlation strengths $g \in [0, 1]$ and either no radial correlations ($\nu = 0$) or maximal radial correlations ($\nu = 1$). The plot corresponding to $\nu = 1$ is Fig. 2h in the main text. We observe that the behavior is virtually the same irrespectively of the absence ($\nu = 0$) or presence ($\nu = 1$) of radial correlations. On the other hand, we observe a clear distinction depending on the absence ($g = 0$) or presence ($g > 0$) of angular correlations. Specifically, we see that in the presence of angular correlations ($g > 0$), ΔN scales with the system size. Furthermore, we observe that for sufficiently large sizes N the scaling of ΔN follows a somewhat similar trend for the considered values of $g > 0$. However, we also observe that when angular correlations are weak and the system size is relatively small, the behavior is close to the case of no angular correlations.

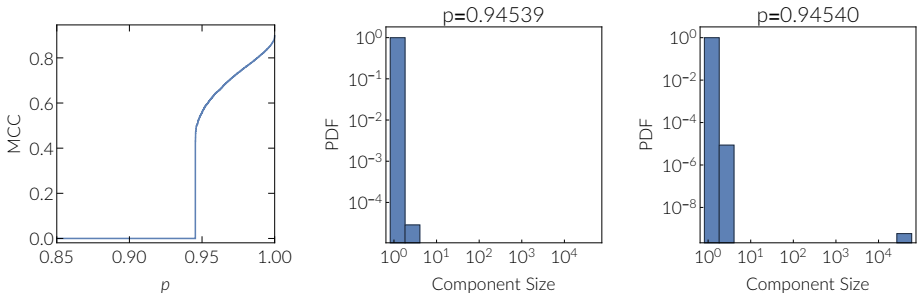

Figure 3. Evolution of the MCC for a two-layer synthetic multiplex constructed with the GMM model. Each layer has $N = 50000$ nodes, power law degree distribution with exponent $\gamma = 2.6$, average node degree $\langle k \rangle \approx 6$, and clustering $\bar{c} = 0.35$. There are no radial or angular correlations, $\nu = 0, g = 0$. Left: relative size of the MCC as a function of the fraction of nodes remaining in the system p. Center: distribution of mutually connected component sizes below the critical point. Right: distribution of mutually connected component sizes above the critical point.

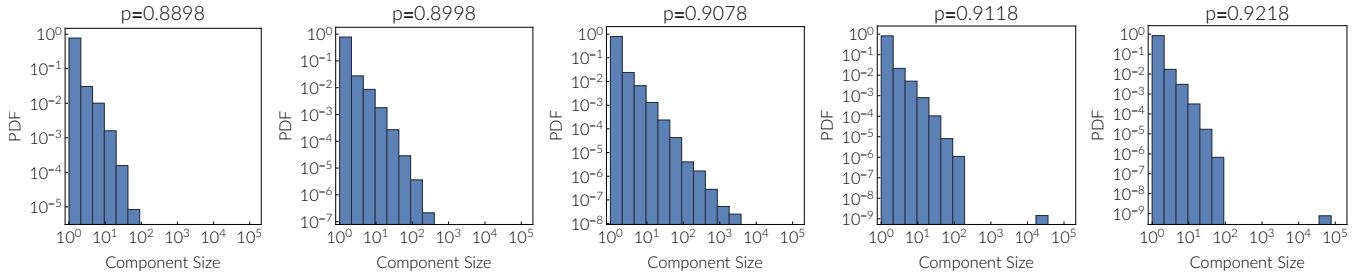

Figure 4. Distribution of component sizes during the evolution of the MCC for a two-layer synthetic multiplex constructed with the GMM model as in Supp. Fig. 3 but with maximal angular correlations, $g = 1$.

VI. SECOND LARGEST MUTUALLY CONNECTED COMPONENT AND DISTRIBUTION OF CASCADES IN MODEL NETWORKS

In Supp. Fig. 6a, we observe that different realizations of the multiplex converge to the same MCC line for large system sizes, as expected. Figs. 6b,c show the evolution of the second largest MCC and the scaling of its maximum with the system size, for maximal angular correlations ($g = 1$) and no radial correlations ($\nu = 0$). In Supp. Fig. 6c, we observe that the maximum of the second largest MCC scales with the system size approximatively $\propto N^{0.84}$ and so scales as a power law, which indicates infinite fluctuations as typically happens for the susceptibility in continuous transitions [10–12].

Furthermore, Supp. Fig. 7 shows the distribution of jumps in the presence ($g = 1$) and absence ($g = 0$) of angular correlations. We define a jump as the number of nodes that are disconnected from the MCC due to the removal of a single node in the targeted attack sequence. We note that in the absence of angular correlations the largest jump cascade is clearly distinct from the second largest one (see Supp. Fig. 7 for $g = 0$). Furthermore, as shown in Fig. 2f of the main text, the size of the largest jump cascade is a constant fraction of the system size in the absence of angular correlations, while the second largest jump cascade relative to the system size vanishes with the system size. On the contrary, as shown in Fig. 2f of the main text, if angular correlations are present, both the largest and second largest jump relative to the system size decrease, and are not clearly distinct in the distribution histograms (Supp. Fig. 7 for $g = 1$). This observation reflects the different nature of the transition depending on the presence or absence of angular correlations.

VII. TYPE OF TRANSITION AND MODEL PARAMETERS

As mentioned in the main text, we observe an apparently continuous transition in the thermodynamic limit only if the multiplex layers have sufficiently strong metric structure and do not loose the small-world property in the targeted attack process. The strength of the metric property in the layer topologies used in our model depends on

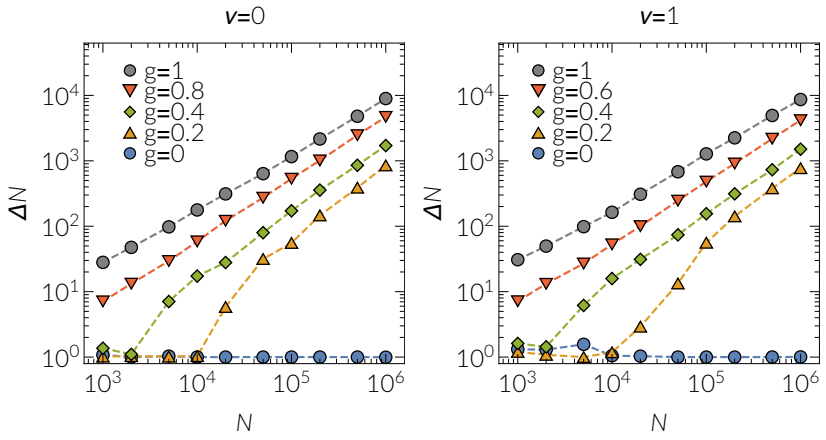

Figure 5. Scaling of the critical number of nodes ΔN for different values of the angular correlation strength parameter $g \in [0, 1]$. Left: no radial correlations, $\nu = 0$. Right: maximal radial correlations, $\nu = 1$.

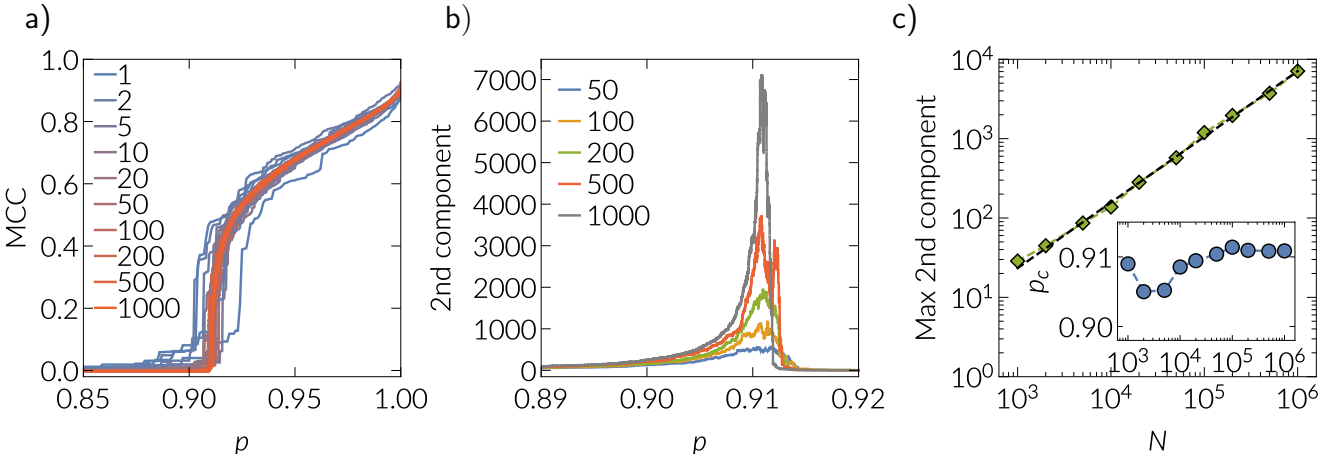

Figure 6. MCC and second largest MCC for a two-layer synthetic multiplex constructed with the GMM model. Each layer has the same number of nodes, power law degree distribution with exponent $\gamma = 2.6$, average node degree $\langle k \rangle \approx 6$, and clustering $\bar{c} = 0.35$. There are maximal angular correlations $g = 1$ and no radial correlations $\nu = 0$. Plot **(a)** shows the evolution of the MCC for different layer sizes N as indicated in the legend ($\times 10^3$); for each size, results are shown for three realizations of the multiplex. Plot **(b)** shows the size of the second largest MCC as a function of p for different layer sizes N as indicated in the legend ($\times 10^3$); for each size, the results are averages over 60 realizations of the multiplex. Plot **(c)** shows the scaling of the maximum of the second largest MCC; for each size N, the results are averages over 60 realizations of the multiplex. The black dashed line shows a fit $\propto N^{0.84}$, while the inset shows the value of $p \equiv p_c$ where the maximum is realized.

the temperature parameter $T \in (0, 1)$, which also controls the amount of clustering in the system (lower temperature means higher clustering and stronger metric properties, and for $T \to 1$ clustering vanishes in the thermodynamic limit) [1]. However, for $T < 0.5$ the resulting networks are only small-world if they are simultaneously scale-free [5], i.e., they have a power-law degree distribution exponent γ between 2 and 3. Hence, for $T < 0.5$, to fulfil the small-world properties the generated networks rely strongly on their hubs. The hubs, however, are removed in the targeted attacks process, and hence at some point the networks loose the small-world property (this will happen earlier for lower temperatures). As a consequence, the multiplex behaves effectively like a one-dimensional system, and the critical point of the transition is expected to go to $p_c = 1$. Indeed, Supp. Fig. 8b confirms that, as we reduce the temperature below 0.5, the critical point approaches 1.

On the other hand, if T is increased the strength of the metric structure is reduced. Eventually, for $T \to 1$, the metric structure breaks down and, in particular, the correlated case degenerates to the uncorrelated case in our model. As a consequence, it is not surprising that for high temperatures we obtain a discontinuous transition, even in the presence of geometric correlations. Indeed, Supp. Fig. 8c confirms that this is the case. The figure shows the scaling of the largest cascade in the system, like Fig. 2f in the main text, for different values of the temperature. We observe that only for $0.5 < T < 0.8$ the relative size of the largest cascade decreases. In Supp. Fig. 9 we show the evolution

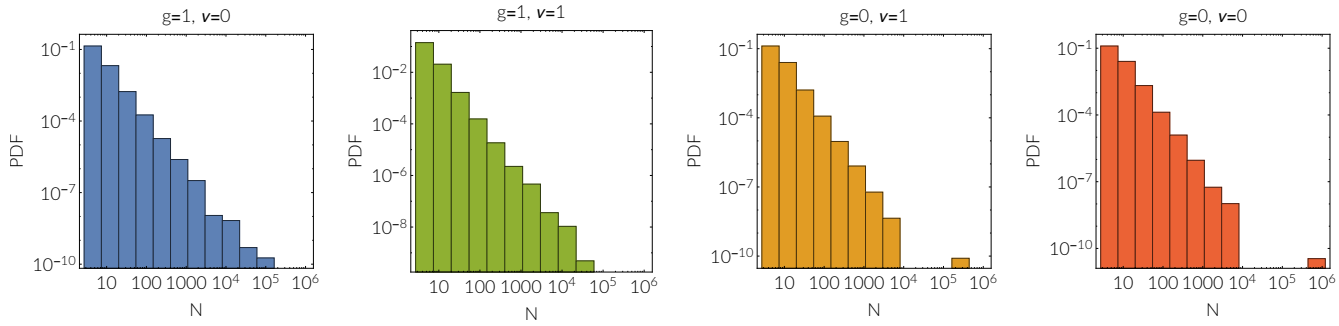

Figure 7. Distribution of cascade sizes for a two-layer synthetic multiplex constructed with the GMM model. Each layer has $N = 10^6$ nodes, power law degree distribution with exponent $\gamma = 2.6$, average node degree $\langle k \rangle \approx 6$, and clustering $\bar{c} = 0.35$. The angular and radial correlation strengths g and ν are indicated in the plot titles.

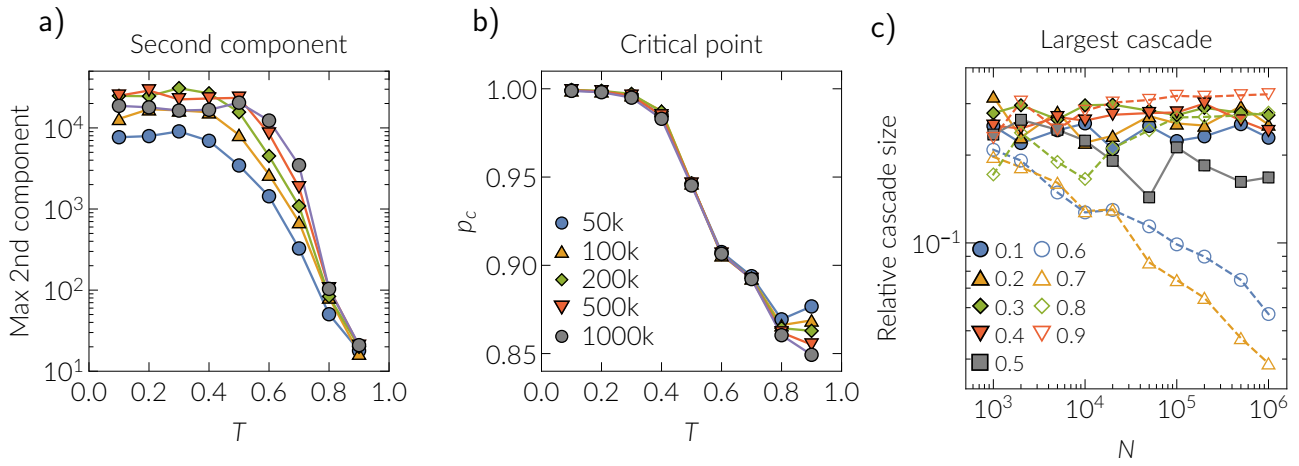

Figure 8. Second largest MCC and largest cascade size in a two-layer synthetic multiplex, as a function of the temperature parameter T. Each layer has the same number of nodes, power law degree distribution with exponent $\gamma = 2.6$, and average node degree $\langle k \rangle \approx 6$. Here, we have set $g = 1$ and $\nu = 0$. Plot **(a)** shows the maximum of the second largest MCC and plot **(b)** shows the value of $p \equiv p_c$ where this maximum is realized. The results correspond to different layer sizes as indicated in the legend of **(b)**. Plot **(c)** shows the size of the largest cascade as a function of the system size for different values of the temperature T as indicated by the colors in the legend.

of the MCC of the system for different values of T and different sizes.

VIII. EDGE OVERLAP IS NOT SUFFICIENT TO PRODUCE THE MITIGATION EFFECT

Geometric correlations induce, but are not equivalent to, edge overlap, i.e., a finite fraction of links that exist simultaneously in all layers of a multiplex. In this section, we show that edge overlap alone is not enough to mitigate the vulnerability of multiplex networks against targeted attacks.

To this end, we have designed the following null model that constructs synthetic two-layer multiplexes with the same single layer topologies and edge overlap as those with angular correlations considered in the main text. We first start with the trivial case of two identical network layers. That is, we create one layer with the GMM model, and then simple clone this layer. We denote this two-layer multiplex by *id*. Then, we perform a node id reshuffling procedure on this multiplex similar to the one described in the main text, but with the following difference. We select a layer, but instead of reshuffling all node-to-node mappings in the layer, we first randomly select a fraction q of nodes from the layer that we denote by $\mathcal{Q}$. Then, we perform a random reshuffling among the ids of nodes that belong to $\mathcal{Q}$. We denote this partially reshuffled counterpart by *rs*. Below, we present results for $q = 0.5$, which leads to the same edge overlap $\mathcal{O} \approx 0.2$ in *rs*, as in a two-layer multiplex constructed by the GMM model as in Fig. 2e of the main text with $g = 1$ (maximal angular correlations) and $\nu = 0$ (no radial correlations)—we denote the latter by *an*.

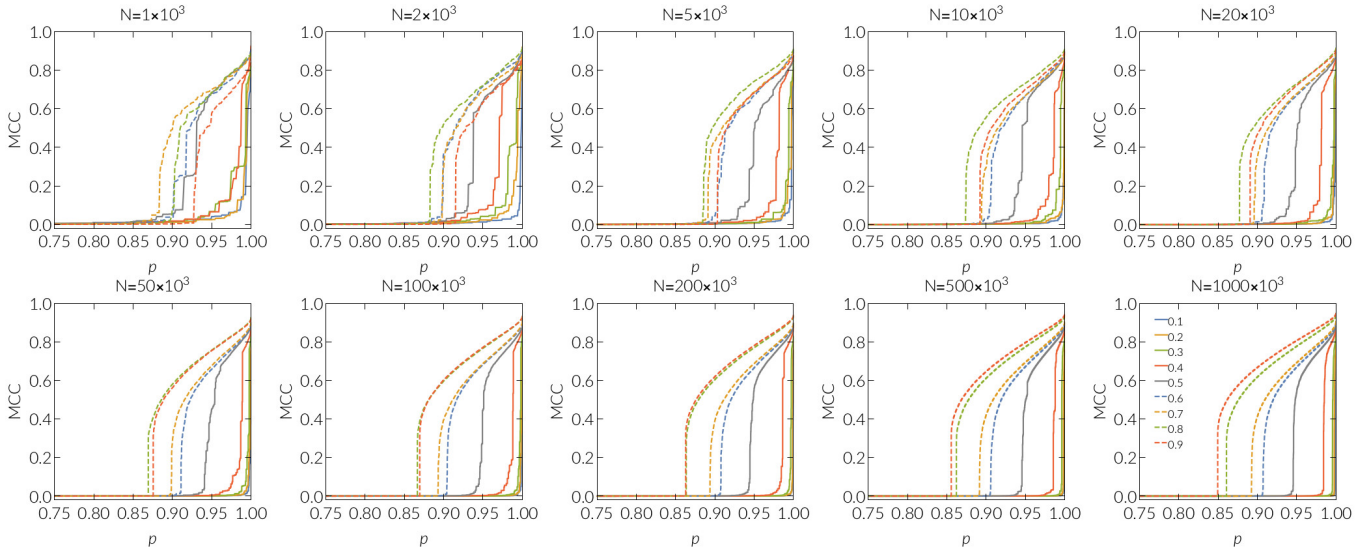

Figure 9. Evolution of the MCC for two-layer synthetic multiplexes of different layer sizes N as indicated in the plot titles and different temperatures as indicated in the legend of the last plot. Each layer has a power law degree distribution with exponent $\gamma = 2.6$, and average node degree $\langle k \rangle \approx 6$. Here, we have set $g = 1$ and $\nu = 0$.

In Supp. Fig. 10, we juxtapose the evolution of the MCC in *id*, *rs*, and *an*. In the figure, we also consider a two-layer multiplex generated by the GMM model as above but with no angular or radial correlations ($g = 0, \nu = 0$)—denoted by *un*. We note that the edge overlap in *un* is $\mathcal{O} \propto 1/N \approx 0$. From the figure, we observe that *rs* has a lower critical point p compared to the uncorrelated multiplex *un*, but has a steeper onset compared to the angular correlated multiplex *an*.

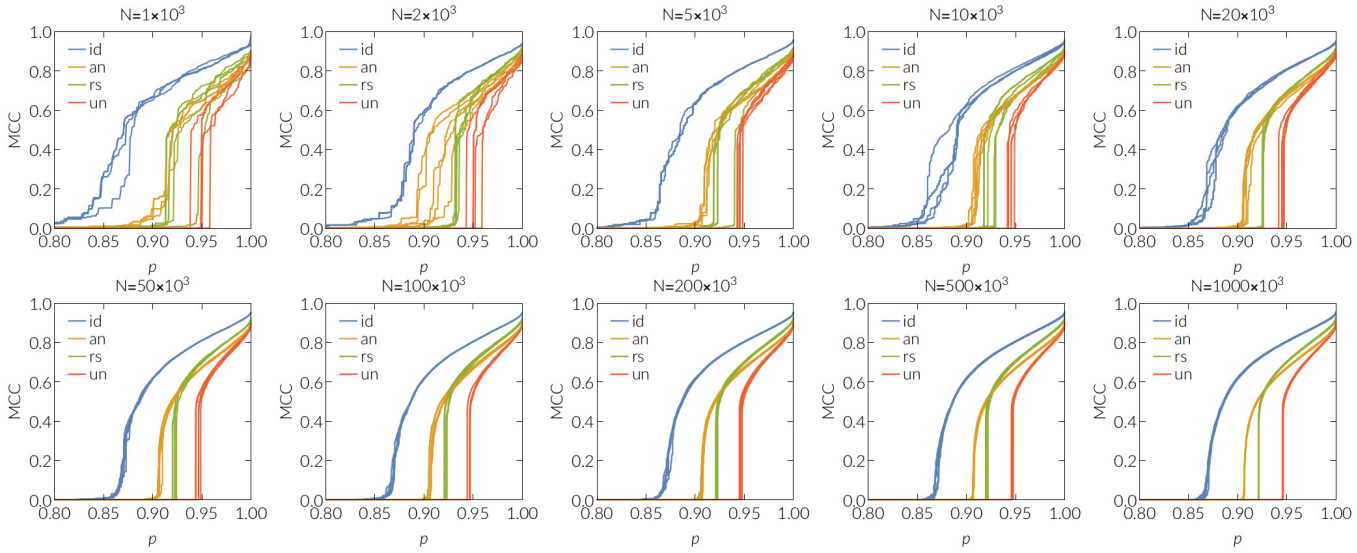

Figure 10. Evolution of the MCC for two-layer synthetic multiplexes constructed as described in the text: id (identical layers), rs (partially reshuffled, $q = 0.5$), an (angularly correlated layers), and un (uncorrelated layers). Each plot shows results for four realizations of each multiplex. In all cases the layers have the same size N as indicated in the plot titles.

Furthermore, Supp. Fig. 11a shows the scaling of the critical number of nodes ΔN with the system size. We observe that eventually *rs* behaves as the uncorrelated case *un*, i.e., $\Delta N = 1$ for sufficiently large system sizes. By contrast, ΔN increases similarly in *id* and *an* with the system size. For *rs*, we also find that the maximum of the second largest MCC increases with the system size, but slower compared to *id* and *an*, see Supp. Fig. 11b. Finally, regarding the cascade sizes, *rs* behaves similarly as the uncorrelated case *un*. Specifically, in Figs. 11c,d we observe that for *rs* and *un* the largest cascade remains constant with the system size, while the second largest decreases. By contrast, in *id*

and *an* both the largest and second largest cascades decrease with the system size.

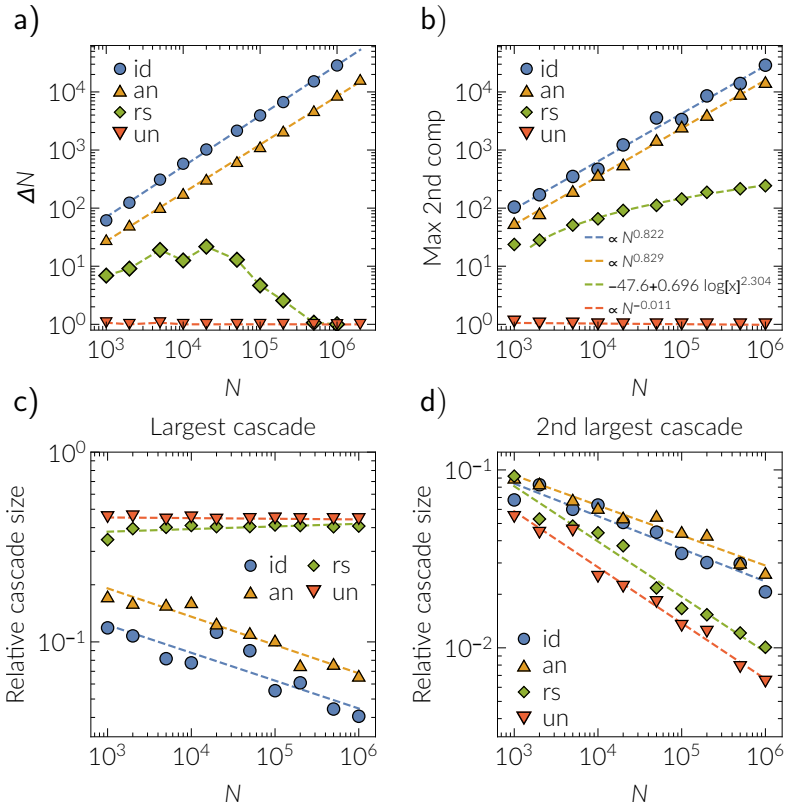

Figure 11. **(a)** Critical number of nodes ΔN, **(b)** maximum of the second largest MCC, **(c)** largest cascade size, and **(d)** second largest cascade size. The results correspond to two-layer synthetic multiplexes constructed as described in the text: id (identical layers), rs (partially reshuffled, $q = 0.5$), an (angularly correlated layers), and un (uncorrelated layers). In all cases the layers have the same size N indicated in the x-axes. Each point in the plots is an average over 30 realizations of the corresponding multiplex.

To conclude, *rs* exhibits qualitatively the same behavior as the uncorrelated multiplex *un*, while the behavior of *an* qualitatively resembles the one of *id*. Therefore, overlap alone is not sufficient to mitigate the extreme vulnerability of multiplex networks against targeted attacks.

IX. QUANTIFYING ANGULAR (SIMILARITY) CORRELATIONS VIA MUTUAL INFORMATION

Here, we provide an overview of the mutual information measure used to quantify the correlation strength between the inferred angular coordinates of nodes that exist in both layers of each considered layer pair in Supp. Tab. I. Formally, the mutual information between two random variables X, Y can be written as [13]

$$I(X;Y) = \int_Y \int_X p(x,y) \ln \left(\frac{p(x,y)}{p(x)p(y)} \right) \mathrm{d}x\mathrm{d}y, \tag{19}$$

where $p(x,y)$ is the joint probability density function of X, Y, and $p(x), p(y)$ are the marginal probability density functions of X and Y. The higher the mutual information $I(X;Y)$ the stronger is the correlation between X and Y. To be able to compare angular correlations across different layer pairs we compute the normalized mutual information,

$$\mathrm{NMI} = I(X;Y) / \max\{I(X;X), I(Y;Y)\}, \tag{20}$$

which takes values in $[0, 1]$. The higher the NMI, the stronger are the angular correlations between layers. To compute the mutual information between layer pairs we have used the k-nearest-neighbor estimator of Kraskov et al. [13], whose implementation can be found at [14]. See [1] for further details.

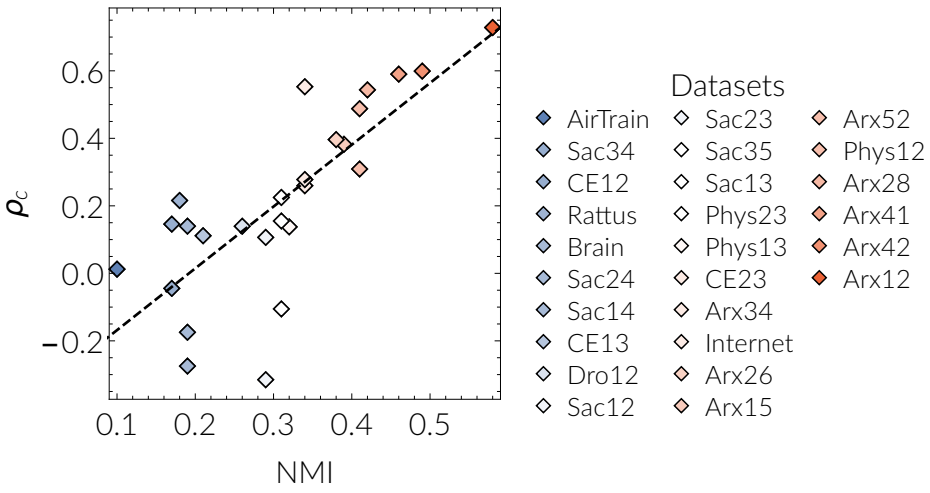

Figure 12. Clustering correlation coefficient ρ_c vs. NMI in layer pairs of real multiplexes.

X. GEOMETRIC CORRELATIONS VS. CLUSTERING CORRELATIONS

Real networks usually exhibit high clustering, i.e., a large number of triangular subgraphs, which can be seen as a natural reflection of the metric property of the underlying similarity space [5, 6]. Intuitively, if a node a is close in the angular similarity space (i.e., it is similar) to a node b, and b is close to a third node c, then a is also close to c because of the triangle inequality, forming triangle abc in the observed network topology. Indeed, most of the real layers we considered have an average clustering coefficient in the range 0.3-0.8 [1], see Supp. Tab. I.

Geometric correlations, that is, correlations among the similarity coordinates of nodes in different layers, are expected to lead to clustering correlations, that is, correlations between the clustering coefficients of the nodes. We verify this in Fig. 14, which shows that the Pearson correlation coefficient [15] between the clustering coefficients of nodes, ρ_c, is significant in the majority of the considered real layers, and well correlated with the NMI between the layers. However, the converse is not always true. That is, clustering correlations do not necessarily lead to geometric correlations. Below, we verify that clustering correlations alone cannot explain the increased robustness observed in real multiplexes.

To this end, we create counterparts of real systems by reassigning ids to nodes of the layers such that clustering correlations across layers are maximized. Specifically, we perform the following *clustering-correlation-maximization* procedure. Given two layers $1, 2$, we first create two corresponding lists l_1, l_2, each consisting of the nodes of each layer ranked according to the decreasing order of their clustering coefficient. That is, $l_1(i)$ and $l_2(i)$ contain the nodes with the i^{th} highest clustering coefficient in layers 1 and 2, respectively. For each position i, we then assign the same id to nodes that are at $l_1(i), l_2(i)$. We note that this procedure only reassigns node ids in the two layers and does not alter the layers' topologies. However, it breaks geometric correlations between the layers as two nodes with the same clustering coefficient can have very different coordinates in the similarity space of each layer after this procedure.

Supp. Tab. II juxtaposes the clustering correlation coefficients in the real systems and in their reshuffled counterparts obtained following the above procedure, as well as the corresponding critical number of nodes ΔN and ΔN^c_{rs}. We observe that even though ΔN^c_{rs} is higher in some of the cases compared to the number obtained after the random reshuffling procedure in Table I (ΔN_{rs}), it is still significantly lower compared to the one in the real systems, ΔN, for the majority of the cases. These results validate that clustering correlations alone cannot explain the increased robustness observed in real systems.

[1] Kaj-Kolja Kleineberg, Marián Boguñá, M. Ángeles Serrano, and Fragkiskos Papadopoulos, "Hidden geometric correlations in real multiplex networks," Nature Physics **12**, 1076–1081 (2016).
[2] Marián Boguñá, Fragkiskos Papadopoulos, and Dmitri Krioukov, "Sustaining the Internet with hyperbolic mapping." Nature communications **1**, 62 (2010).
[3] F. Papadopoulos, C. Psomas, and D. Krioukov, "Network mapping by replaying hyperbolic growth," Networking, IEEE/ACM Transactions on **23**, 198–211 (2015).

Dataset	Abbreviation	ρ_c	ρ_c^{rs}	ΔN	ΔN_{rs}^c
arXiv Layers 1, 2	Arx12	0.73	1.00	25.2	1.1
arXiv Layers 4, 2	Arx42	0.60	0.98	43.8	7.2
arXiv Layers 4, 1	Arx41	0.59	0.98	16.6	3.1
arXiv Layers 2, 8	Arx28	0.54	0.98	15.8	3.1
Physicians Layers 1, 2	Phys12	0.31	0.97	6.0	1.0
arXiv Layers 5, 2	Arx52	0.49	0.99	16.8	3.0
arXiv Layers 1, 5	Arx15	0.38	0.99	9.0	1.0
arXiv Layers 2, 6	Arx26	0.40	0.97	10.0	3.0
Internet Layer 1, 2	Internet	0.26	0.91	81.4	16.2
arXiv Layers 3, 4	Arx34	0.55	0.95	3.0	4.0
C. Elegans Layers 2, 3	CE23	0.28	0.98	14.0	4.6
Physicians Layers 1, 3	Phys13	0.14	0.96	1.0	2.0
Physicians Layers 2, 3	Phys23	0.16	0.93	2.0	1.0
SacchPomb Layers 1, 3	Sac13	0.22	0.88	18.3	1.0
SacchPomb Layers 3, 5	Sac35	-0.11	0.58	1.0	1.0
SacchPomb Layers 2, 3	Sac23	-0.32	0.90	2.0	1.0
SacchPomb Layers 1, 2	Sac12	0.11	0.95	1.0	2.0
Drosophila Layers 1, 2	Dro12	0.14	0.98	8.4	6.2
C. Elegans Layers 1, 3	CE13	0.11	0.91	8.2	1.0
SacchPomb Layers 1, 4	Sac14	0.14	0.94	3.5	1.5
SacchPomb Layers 2, 4	Sac24	-0.17	0.86	2.0	1.0
Brain Layers 1, 2	Brain	-0.27	0.96	7.0	1.0
Rattus Layers 1, 2	Rattus	0.22	0.53	4.0	1.0
C. Elegans Layers 1, 2	CE12	0.15	0.95	1.0	1.5
SacchPomb Layers 3, 4	Sac34	-0.04	0.98	4.2	10.7
Air/Train Layers 1, 2	AirTrain	0.01	0.83	3.0	3.0

Table II. ρ_c and ρ_c^{rs} denote respectively the clustering correlation coefficients in the real systems and in their reshuffled counterparts obtained by the clustering-correlation-maximization procedure. ΔN and ΔN_{rs}^c are the corresponding critical number of nodes.

[4] Fragkiskos Papadopoulos, Rodrigo Aldecoa, and Dmitri Krioukov, "Network geometry inference using common neighbors," Phys. Rev. E **92**, 022807 (2015).

[5] M. Serrano, Dmitri Krioukov, and Marián Boguñá, "Self-Similarity of Complex Networks and Hidden Metric Spaces," Phys Rev Lett **100**, 078701 (2008).

[6] F. Papadopoulos, M. Kitsak, M. Serrano, M. Boguñá, and D. Krioukov, "Popularity versus similarity in growing networks," Nature **489**, 537–540 (2012).

[7] This description is taken directly from the Supplementary Information of Ref. [1] and repeated here for the convenience of the reader.

[8] Dmitri Krioukov, Fragkiskos Papadopoulos, Maksim Kitsak, Amin Vahdat, and Marián Boguñá, "Hyperbolic geometry of complex networks," Physical Review E **82**, 036106 (2010).

[9] By sampling from a PDF $f(x)$ we mean that we first compute the CDF $F(x) = \int_{x_{\min}}^{x} \mathrm{d}x' \rho(x')$, where $x_{\min}$ is the minimum value of x, then generate a random number u_i uniformly at random from $[0, 1]$, and finally compute the value x_i such that $F(x_i) = u_i$. The value x_i is a sample from the PDF $\rho(x)$ (or the CDF $F(x)$).

[10] D. Stauffer and A. Aharony, *Introduction to Percolation Theory* (Oxford University Press, New York, 1971).

[11] H. E. Stanley, *Introduction to Phase Transitions and Critical Phenomena* (Oxford University Press, 1987).

[12] Didier Sornette, *Critical Phenomena in Natural Sciences (Chaos, Fractals, Selforganization and Disorder: Concepts and Tools)* (Springer Series in Synergetics, 2006).

[13] Alexander Kraskov, Harald Stögbauer, and Peter Grassberger, "Estimating mutual information," Phys. Rev. E **69**, 066138 (2004).

[14] "Mutual Information Least-dependent Component Analysis (MILCA)," `https://www.ucl.ac.uk/ion/departments/sobell/Research/RLemon/MILCA/MILCA`.

[15] Theodore Wilbur Anderson, *An introduction to multivariate statistical analysis*, Wiley series in probability and mathematical statistics (J. Wiley, 1984).