Identifying the hidden multiplex architecture of complex systems

Lucas Lacasa[1], Inés P. Mariño[2], Joaquin Miguez[1,3], Vincenzo Nicosia[1], Jesús Gómez-Gardeñes[4,5],

[1]*School of Mathematical Sciences, Queen Mary University of London. Mile End Rd, E14NS London, United Kingdom*
[2]*Department of Biology, Geology, Physics & Inorganic Chemistry,*
Universidad Rey Juan Carlos, 28933 Móstoles, Madrid, Spain
[3]*Department of Signal Theory & Communications,*
Universidad Carlos III de Madrid. 28911 Leganés, Madrid, Spain
[4]*Department of Condensed Matter Physics, University of Zaragoza, 50009 Zaragoza (Spain)*
[5]*Institute for Biocomputation and Physics of Complex Systems (BIFI),*
University of Zaragoza, E-50018 Zaragoza, Spain
(Dated: May 15, 2017)

The architecture of many complex systems is well described by multiplex interaction networks, and their dynamics is often the result of several intertwined processes taking place at different levels. However only in a few cases can such multi-layered architecture be empirically observed, as one usually only has experimental access to such structure from an aggregated projection. A fundamental question is thus to determine whether the hidden underlying architecture of complex systems is better modelled as a single interaction layer or results from the aggregation and interplay of multiple layers. Here we show that, by only using local information provided by a random walker navigating the aggregated network, it is possible to decide in a robust way if the underlying structure is a multiplex and, in the latter case, to determine the most probable number of layers. The proposed methodology detects and estimates the optimal architecture capable of reproducing observable non-Markovian dynamics taking place on networks, with applications ranging from human or animal mobility to electronic transport or molecular motors. Furthermore, the mathematical theory extends above and beyond detection of physical layers in networked complex systems, as it provides a general solution for the optimal decomposition of complex dynamics in a Markov switching combination of simple (diffusive) dynamics.

Network theory has emerged as a powerful unifying framework for studying the emergence of collective phenomena in real complex systems from different domains [1, 2], and has allowed us to increase the accuracy and predictive power of our models of complex dynamics, including epidemic spreading [3], synchronisation [5], and social dynamics [4] to cite just a few. One of the most fascinating challenges faced in the last few years by network science is the need to incorporate and couple several network structures in order to correctly capture the inherently multidimensional nature of interaction patterns in real-world systems. As a result, much effort has been recently devoted to the definition and study of multilayer and multiplex networks [6–8]. The ubiquity of such structures in social, biological and technological systems has required the revision of the several canonical dynamical models that were previously studied only on isolated complex networks, including percolation [12–16], diffusion dynamics [17, 18], navigation [19–21], epidemics [22–25], evolutionary games [26–28], synchronization [29], or opinion dynamics [30, 31], among several others. These studies have frequently found that the collective behaviour of systems qualitatively changes when passing from isolated networks to coupled ones [32], highlighting the importance of the multilayer architecture of real-world systems.

There are two fundamental and indeed dual problems to face when one introduces multilayer network models of real world systems. The first one is the necessity to assess in a systematic way whether a multilayer network model is adequate to represent the system, and when such model gives redundant information. This

FIG. 1: A multiplex network with $L = 2$ layers and $N = 7$ nodes. A random walker diffusing over this structure generates a two dimensional time series $\{[X(t), l(t)]\}_{t=1}^{N}$ where $X(t)$ and $l(t)$ are the vertex and layer locations of the walker at time t. In many real-world cases, such as for human mobility, the layer indicator $l(t)$ is hidden and one has access only to $\{X(t)\}_{t=1}^{N}$, i.e. to the series of positions of the walker on the projected network, shown in the bottom of the figure. For $L > 1$ the resulting trajectory is non-Markovian: we rely on this Markovianity-breaking phenomenon property to detect multiplexity and to provide an estimate of the number of layers in the system by using only $\{X(t)\}_{t=1}^{N}$.

2

problem was first addressed in Ref. [33], and constitutes nowadays an intense field of research. The dual problem aims at understanding whether an empirical network whose multilayer character is not directly observable is genuinely monolayer or is only an aggregated projection of a hidden multilayer network (see figure 1 for an illustration of a multiplex network with $L = 2$ layers and its aggregated projection). Such scenario has received much less attention despite being, for instance, central for networks arising in natural systems whose architecture is not directly observable, as in genetic networks or in brain functional networks where pairs of nodes modelling different brain areas can interact according to an a priori unknown range of different biological pathways [39].

Here we focus on the problem of identifying the hidden multi-layer structure of a complex system from coarse-grained measurements of its state. We will show that, by using only local information extracted from simple random walk statistics, it is possible to reveal whether the underlying structure of the system is actually a single-layer or a multi-layer network, and in the latter case, to estimate the number of interacting layers in the system. Note that methods to infer network topological properties via random walk statistics have been explored previously [34]. Here, the method used to check whether the system consists of one or more layer exploits the breaking of Markovianity occuring in a multi-layer random walk, while the estimation of the most probable number of layers is based on a maximum a-posteriori (MAP) probabilistic criterion which can be implemented via numerical integration methods, including conventional grid-based approximations [48] or more sophisticated Monte Carlo algorithms [47].

RESULTS

Multiplex networks are the most ubiquitous class of multilayer networks. They are defined by a set of $L \geq 1$ interaction layers (networks), all of them having the same set of K nodes but different topology (edge set), with the peculiarity that each node has a replica in each layer (see fig. 1 for an illustration). This structure is thereby fully described by a set of adjacency matrices $\{\mathbf{A}^{(l)}\}_{l=1}^{L}$. Multiplex networks are a natural model for online social networks [9], where a given individual can communicate with others via different platforms (*e.g.* Facebook, Twitter, email, etc) or transportation networks [10, 11], where a set of locations can be connected in a multimodal way (e.g. bus, train, underground, etc). We consider a random walker navigating a multiplex [19] defined as follows: jumps between layers are governed by a Markov chain with $L \times L$ transition matrix $\mathbf{R}_L$ (R_{ij} is the probability to jump from layer i to layer

j) while the dynamics within each individual layer l is also Markovian and determined by a $K \times K$ transition matrix $\mathbf{T}^{(l)}$ (where $T_{ij}^{(l)}$ is the probability to walk from node i to node j at layer l). For the sake of clarity, from now on we only consider the diffusive dynamics where at each time step the walker at node i on layer l (i) remains in the same layer with probability $R_{ll} = 1 - r$ or instantaneously jumps with probability $R_{ll'} = r/(L-1)$ to a different layer l' and subsequently (ii) diffuses to one of the neighbours of node i in the chosen layer with uniform probability. When $r \ll 1$ (i.e. when walkers tend to remain in the same layer) this navigation model aims at mimicking human mobility in multilayered transportation networks [40, 41], where multimodality is minimized to avoid waiting times related to connections between different modes. Interestingly, this approach also goes beyond network science and is reminiscent of the so-called discrete flashing ratchet model [42, 43]. Ratchets are canonical models in statistical physics to describe the transport of Brownian particles embedded in periodic, asymmetric energy potentials, a paradigm originally proposed by Smoluchowski [44] and popularized by Feynman [45] as a possible thermal engine, and further shown to be a fruitful way of modelling molecular motors in biophysics. A Brownian particle subject to a periodic asymmetric potential that is switched ON and OFF stochastically is formally equivalent to our random walk navigating over a multiplex with $L = 2$ cycle graphs with different transition matrices.

The stochastic process defined above is fully described by an infinite two-dimensional time series $\{(X(t), l(t))\}_{t=1}^{\infty}$ where $X(t) \in \{1, \ldots, K\}$ and $l(t) \in \{1, \ldots, L\}$. As in real-world scenarios the multiplex nature of the system is not always empirically accessible, the layer indicator $l(t)$ is hidden and the only observable is the sequence of node locations $X(1), X(2), \ldots$ In other words, we only have experimental access to *partial* information of the process, described by a finite sequence of observations

$$\mathcal{O} = \{X(t)\}_{t=1}^{N} . \tag{1}$$

This is formally equivalent to observing a dynamical process on the aggregated (projected) network. Hence the question: is it possible to discern if the system is multiplex and in that case, to estimate the most probable number of layers having only access to $\mathcal{O}$?

Inferring multiplexity. Consider a Markov switching walker navigating a (multiplex) network from which we only have access to partial information given by eq. 1. Assuming $X(t)$ is Markov, we can estimate directly from $\mathcal{O}$ the (monoplex) transition matrix $\mathbf{Q}$ that would describe such a Markovian dynamics. Accordingly, we can define a Markov chain associated to $\mathbf{Q}$ and generate a sample series of N steps $\{Y(t)\}_{t=1}^{N}$, which ef-

FIG. 2: (Left panel) The normalised Kullback-Leibler divergence $D(m)$ between $X(t)$ and its Markovian surrogate $Y(t)$, in the case where $X(t)$ shows an induced current, for different values of the switching rate r. The series $X(t)$ records the position of a walker diffusing over two layers, with transition probabilities in the layers are $T_{i,i+1}^{(1)} = 1/3$ and $T_{i,i+1}^{(2)} = 1/2$ respectively, such that the transition probability in the Markovian surrogate series is $Q_{i,i+1} = 5/12$. We correctly find that $X(t)$ is non-Markovian even for large values of r as $D(m > 2) > 0$, what suggests an underlying multiplex structure. (Right panel) Finite size analysis where we compute $D(3)$ (showing the mean and standard deviation over 10 realizations) as a function of the series size N, for the multiplex considered in the left panel ($r = 0.1$, black dots) and a null model with equivalent monoplex dynamics (red squares) for which $D(3)$ should vanish asymptotically.

fectively is a Markovian surrogate of the original process. Now, if the underlying network was truly monoplex, then $X(t)$ would be actually Markov and $X(t)$ and $Y(t)$ would then have asymptotically equivalent statistics. For multiplex structures however, by construction $X(t)$ and $Y(t)$ would share the same joint distributions only up to blocks of size $m = 2$, but these would differ for $m > 2$. To quantify such difference, we make use of the (normalized) block Kullback-Leibler divergence of order m between blocks $\boldsymbol{X}_m = (X(t), \dots, X(t + m - 1))$ and $\boldsymbol{Y}_m = (Y(t), \dots, Y(t + m - 1))$:

$$D(m) := \frac{\mathrm{KLD}_m(X||Y)}{m} = \frac{1}{m} \sum_{\mathcal{B}(m)} P(\mathbf{X}_m) \log \frac{P(\mathbf{X}_m)}{P(\mathbf{Y}_m)} \tag{2}$$

where $\mathcal{B}(m)$ enumerates all the blocks of size m. $D(m)$ is semi-positive definite and vanishes only when the joint probabilities coincide [46]. Thus by construction $D(1) = D(2) = 0$. The Markovianity criterion finally establishes that $D(m) > 0$ for $m > 2$ implies multiplexity[1].

As a proof of concept, we initially consider the simple scenario where a random walker navigates over a two-layer multiplex ring (each layer is a cycle graph of K nodes), a model compatible with a discrete flashing ratchet as

---

[1] Even if $X(t)$ is Markov, the matrix $\mathbf{Q}$ cannot be computed exactly, but simply estimated from the observations. Similarly, one can only estimate the probabilities $P(\mathbf{X}_m)$. Therefore, in practical scenarios the decision rule requires to introduce an error threshold $\epsilon > 0$ such that $D(3) > \epsilon$ implies multiplexity. This detection process can be formally described as a statistical test and error bounds can be obtained under mild assumptions.

commented before. In the first layer, we define a Markov chain with transition probabilities $T_{i+1,i}^{(1)} = 2/3$; $T_{i,i+1}^{(1)} = 1/3$ and $T_{ij}^{(1)} = 0$ if $j \neq i + 1 \mod K$ or $j \neq i - 1 \mod K$. A random walker diffusing in this layer will have an induced current in the direction of decreasing node indices. In the second layer, we define a different Markov chain with transition probabilities $T_{i+1,i}^{(2)} = 1/2$; $T_{i,i+1}^{(2)} = 1/2$ and $T_{ij}^{(2)} = 0$ otherwise, i.e., a reversible Markov chain. While we can always estimate $\mathbf{Q}$ numerically from eq.1, in this simple case it is easy to derive it analytically: $Q_{ij} = T_{ij}^{(1)} W_1 + T_{ij}^{(2)} W_2$, where W_1 (resp., W_2) is the probability of finding the walker in layer $l = 1$ (resp., $l = 2$). Now, since in this case $R_{ij} = R_{ji} = r$, the system is symmetric with respect to the switching process and the walker spends on average the same amount of time in each of the two layers hence $W_1 = W_2 = 1/2$ and then, $Q_{ij} = T_{ij}^{(1)}/2 + T_{ij}^{(2)}/2$. For this specific example, we thus find $Q_{i+1,i} = 7/12$, $Q_{i,i+1} = 5/12$.

In the left panel of Fig.2 we plot $D(m)$ for different switching rates r. In order to deal with finite size effects (which increase exponentially with m), we systematically increase the size of the walker series under study as a function of m, taking series of size $N(m) = N_0 \cdot 2^m$ data extracted from the original system and from the corresponding Markovian surrogates (we used $N_0 = 10^5$ although smaller values yield qualitatively equivalent results). Results are averaged over 10 realisations and we also plot as error bars the standard deviation of the ensemble average. As expected, $D(1) = D(2) = 0$, meaning that $Y(t)$ is a faithful Markovian surrogate of $X(t)$. Furthermore $D(m > 2) > 0$, meaning that $X(t)$ is non-Markovian and hence the underlying network is correctly

4

recognized as a multiplex. This result is robust for a quite large range of values of r (note that $r = 0.5$ is a trivial exception), meaning that the method works even if the walker makes fast switches between layers. A similar scenario is found if we tune the transition probabilities such that no net induced current is found (see SI), pointing out that multiplexity can be unraveled even in that case. In the right panel of the same figure we plot $D(3)$ for different series sizes to showcase how finite size effects vanish as the series gets larger. Already with a few thousands data points we can already accurately detect multiplexity.

For completeness, we have considered a range of seven additional scenarios including (i) layers with increasingly different and disordered topologies controlled by both rewiring and edge addition, (ii) Erdos-Renyi graphs and (iii) similar scenarios on larger graphs. In every case we find similarly positive results: (i) a correct multiplexity detection and (ii) good scalability (see figures S2-S11 in section I the SI for details) thus concluding that the methodology is flexible and robust.

FIG. 3: (Outset panel) Posterior probability $P(L|\{X(t)\}_{t=1}^N)$ as a function of the number of layers L, computed from a trajectory of 2×10^4 time steps generated using a model with $L = 3$ layers and parameters $R_{ii} = 1 - r = 0.84$, $T_{i,i+1}^{(1)} = 0.16$, $T_{i,i+1}^{(2)} = 0.76$ and $T_{i,i+1}^{(3)} = 0.24$ (see the text and SI for details). The algorithm easily estimates the correct model $L = 3$. (Inset panel) A linear-log plot of the same graph. Note that the probability for $L = 1$ is zero up to the computer's accuracy, hence this point is not shown in logarithmic scales.

Quantifying multiplexity. Unfortunately, the Markovianity criterion described above is only a qualitative one, and does not provide a quantitative way of estimating the number of layers of the underlying multiplex. In order to bridge this gap, we now make use of statistical inference tools to define a model selection scheme [37]. We assume that two models are different if they have a different number of layers. Accordingly, the number of layers L is now modelled as a random variable with prior probability mass function (pmf) $P_0(L)$. Given the value of L, the motion of the random walker is deter-

mined by the Markov-switching of layers, with transition matrix $\mathbf{R}_L$, and Markov walks within each layer characterised by $\mathbf{T}_L = \{\mathbf{T}^{(l)}\}_{l=1}^L$. Assuming prior probability density functions (pdf's) for these parameters $p_{0,R}(\mathbf{R}_L)$ and a $p_{0,T}(\mathbf{T}_L)$, the likelihood of a given model with L layers conditional on the observed data $\{X(t)\}_{t=1}^N$ reads

$$P(\{X(t)\}_{t=1}^N|L) = \int P(\{X(t)\}_{t=1}^N|\mathbf{T}_L, \mathbf{R}_L) \times \qquad (3)$$
$$\times p_{0,T}(\mathbf{T}_L)p_{0,R}(\mathbf{R}_L)\mu(\mathrm{d}\mathbf{T}_L \times \mathrm{d}\mathbf{R}_L),$$

where μ is a suitable reference measure for $\mathbf{T}_L$ and $\mathbf{R}_L$ (see sections II and III in the SI for technical details). In general, this multidimensional integral cannot be computed exactly and needs to be approximated numerically. Then, the number of layers L in the system that generated the data $\{X(t)\}_{t=1}^N$ can be detected using a maximum a posteriori (MAP) criterion, i.e.,

$$\hat{L}^{\mathrm{MAP}} = \arg \max_{L \in \{1,2,...,L_+\}} P(\{X(t)\}_{t=1}^N|L)P_0(L), \quad (4)$$

where $\hat{L}^{\mathrm{MAP}}$ is the estimate of the actual number of layers, $L_+ < \infty$ is the (assumed) maximum possible value of L, and $P(L|\{X(t)\}_{t=1}^N) \propto P(\{X(t)\}_{t=1}^N|L)P_0(L)$ is the a posteriori pmf of L given the observations $\{X(t)\}_{t=1}^N$.

The practical computation of the MAP estimator in Eq. (4) can be addressed in different ways. The classical literature on hidden Markov models (HMMs) [57–59] suggests the use of the expectation-maximisation (EM) algorithm (in various forms) to compute approximate maximum likelihood (ML) estimators, $\hat{\mathbf{T}}_L$ and $\hat{\mathbf{R}}_L$, of the parameters and then assume that $P(\{X(t)\}_{t=1}^N|L) \approx P(\{X(t)\}_{t=1}^N|\hat{\mathbf{T}}_L, \hat{\mathbf{R}}_L)$ in order to compare the models. This approach relies on standard techniques but it has drawbacks. The equation of the *model* likelihood with the *parameter* likelihood easily breaks down when the parameter estimates are poor (e.g., because of overfitting), when the parameter likelihood is multimodal (EM algorithms converge locally) or when the parameter dimension varies significantly for different models. More sophisticated parametric schemes have been proposed (see, e.g., [60]) however they are still subject to these fundamental limitations. The integration in (3) has been favoured theoretically but criticised practically because of the computational cost of approximating $P(\{X(t)\}_{t=1}^N|L)$ numerically [59]. However, we have found that state-of-the-art variational Bayes [61] or adaptive importance sampling [62] methods can be applied effectively up to moderate values of L. See Sections II and III in the SI for further discussion, including examples of using both deterministic integration and the adaptive Monte Carlo sampler from [63].

To showcase the general method of Eq. (4), MAP model selection, we consider again the system reminiscent of

the discrete flashing ratchet where the true model is now formed by $L = 3$ ring-shaped layers with $K = 3$ (for other more sophisticated examples, see section III in the SI, here we consider a simple example for the sake of illustration). In this true model, the probability to stay in the same layer is $R_{ll} = 1 - r = 0.84$ and the probabilities for the walker to move from node $i \rightarrow i + 1$ are $T_{i,i+1}^{(1)} = 0.16$, $T_{i,i+1}^{(2)} = 0.76$ and $T_{i,i+1}^{(3)} = 0.24$ respectively ($T_{i+1,i}^{(l)} = 1 - T_{i,i+1}^{(l)}$ for every l, and $T_{ij}^{(l)} = 0$ for any other i and j).

In order to evaluate the likelihood function $P(\{X(t)\}_{t=1}^{N}|L)$, we need to approximate the integral in (3) for each value $L = 1, \ldots, 4$ ($L_+ = 4$) as discussed previously. This is achieved in a first place via numerical integration over a deterministic grid of 19 points on the interval $(0, 1)$ for each unknown parameter. For the case $L = 1$ this reduces to a single unknown, $T_{i,i+1}^{(1)}$, and for $L > 1$ there are $L + 1$ unknowns: r and $T_{i,i+1}^{(l)}$ for $l = 1, \ldots, L$. Note that the particular choice of transition probabilities was taken to make the problem more challenging, as these values are not commensurate with the integration grid points (for other values we find similar results). The prior pdf $p_{0,R}(\mathbf{R}_L)$ reduces to a uniform pdf on $(0, 1)$ for the unknown parameter r, while the prior $p_{0,T}(\mathbf{T}_L)$ is used to penalise system configurations with two or more identical layers[2]. For this particular experiment we have chosen $p_{0,T}(\mathbf{T}_L) \propto \min_{(l,l')} |T_{i,i+1}^{(l)} - T_{i,i+1}^{(l')}|$, i.e., the prior pdf of a given configuration $\mathbf{T}_L$ is proportional to the minimum distance between any pair of matrices $\mathbf{T}^{(l)}$ and $\mathbf{T}^{(l')}$.

We have confirmed that, without penalisation (i.e., with uniform prior $p_{0,T}$), we obtain multiple equivalent solutions involving layers with identical values of the estimated parameters. Results are summarised in Fig. 3, that shows that the true model is easily detected (i.e., $\hat{L}^{\mathrm{MAP}} = 3$) with the proposed scheme, as the posterior probability emphatically peaks at $L = 3$. Additionally, we have verified that the maximum of the posterior probability density $P(\{X(t)\}_{t=1}^{N}|\mathbf{T}_L, r)p_{0,T}(\mathbf{T}_L)$ (i.e., the integrand of Eq. (3) with uniform prior for r) is indeed attained at the true value of $\mathbf{T}_3$ (see SI), meaning that our model selection scheme correctly estimates not only the number of layers but also the transition probabilities within each layer. We have also checked that the posterior probability density is smooth close to its maximum as perturbations $\tilde{\mathbf{T}}_3 = \mathbf{T}_3 + \delta\mathbf{T}$ systematically yield a smaller posterior pdf for sufficiently large N (see figure S12).

Finally, it is well known that direct, grid-based de-

―――――

[2] Note that a multiplex with $L = 2$ and $\mathbf{T}^{(1)} = \mathbf{T}^{(2)} = \mathbf{T}$ is equivalent to a monoplex with transition matrix $\mathbf{T}$.

terministic integration of the posterior probability is intuitive but computationally inefficient. Accordingly, we have further considered alternative approximations of the integral in (3) using a nonlinear population Monte Carlo algorithm [63], and effectively reduced the runtime by a factor close to 10^2 on the same computer (see SI for technical details) for the example of Fig. 3.

DISCUSSION

In this work we have introduced a method to detect and estimate multiplexity in the hidden underlying structure of a networked system by only having access to local and partial statistics of a random walker. Our working hypothesis (prior) is that there is a hidden multiplex where walkers diffuse, switching layers stochastically and diffusing over each layer. Under these circumstances, any random walker for which we only see a projection of such trajectory in the aggregated network is necessarily non-Markovian, if the number of layers is larger than one. Hence our algorithm for multiplexity detection is based on detecting such breaking of Markovianity. We have illustrated the validity and scalability of this first part by addressing several synthetic systems of varying complexity, as depicted in the main part of the manuscript and the SI. Here we have focused in the specific case of multiplex networks, where each layer has the same number of nodes. Actually, in a multiplex one can even have the same topology in each layer, where only transition weights differ. On the other hand, in a generic multi-layer network each layer will have in general different number of nodes and different topology. This latter situation can be reinterpreted as having a multiplex where in each layer we can have isolated nodes which are never reached by a walker, or forbidden transitions. Accordingly, we envisage that layers in a multiplex will be in general harder to distinguish via our method than layers in a generic multi-layer network, and as such we expect this method to be easily generalizable to the multi-layer case.

In a second step, we have introduced a probabilistic scheme to estimate the most probable number of layers composing the hidden multiplex. Note that probabilistic model selection is not new in network inference, for instance in [38] a similar concept is used to estimate the most probable combination of basic block models that accounts for a certain network topology (see also [35, 36]), whereas in [37] a probabilistic framework was developed to estimate the most probable number of communities in a single-layer network. In our case, the posterior probabilities quantify the likelihood of having a hidden multiplex with L layers. We were able to showcase the validity of this second part in synthetic networks up to $L = 10$ layers. The model selection protocol can be seen as a multidimensional Bayesian

inference problem, and these schemes –similarly to HMMs and other methods in statistical inference– suffer from poor scalability: essentially the computation of the model posterior probability explodes with the number of unknowns. This is a limitation of the method, and its optimization is therefore an open problem for future work. As a matter of fact, a simple and intuitive (although inefficient) way to estimate these posteriors is to use a (deterministic) grid integration scheme. In an effort to improve scalability and optimize such calculations we have proposed a nonlinear population Monte Carlo algorithm which reduces the computer runtime by a factor close to 10^2, without performance loss, for all the examples where we had previously used deterministic integration. This is described in detail in the SI.

Interpretation. The methodology proposed here actually extends above and beyond the reconstruction of (hidden) physical multiplex architectures using walkers with partial information. As a matter of fact, a similar approach can be considered even when the architecture is truly single-layered, however in this latter scenario one needs to switch the interpretation of the results. Suppose for instance that the observed series is *truly* the result of a complex (i.e., non-Markovian) dynamics running on a physical single-layered network. In that case, our multiplex reconstruction method would still provide the most probable multiplex model, with $L > 1$ due to lack of walker's Markovianity. The key difference is that now layers in the hidden multiplex would be *effective* instead of physical, and we would be providing the most probable multiplex with Markovian intralayer dynamics that would yield such complex dynamics. Note that this type of effective model is precisely what we have when we look for community structure in a single-layered network: finding the optimal number of *effective* groups of nodes which maximize a certain likelihood function [37].

In even more general terms, our methodology provides a mathematically sound solution for the optimal projection of any complex dynamics onto a base of simple dynamics modes. This sort of harmonic analysis interpretation by which non-Markovian processes (i.e. processes with infinite memory) can be projected into an appropriate Markov-switching *finite* basis of Markovian dynamics is remarkable as it is hugely reducing the complexity of the generative model, and clearly deserves further investigation.

Applications. As discussed above, our methodological framework is general and, as such, enjoys a broad range of potential applications in various fields of science and technology. Consider, to begin with, the general study of human and animal mobility [49–51]. This type of dynamical process has been explored almost

exclusively from the perspective of single-layer networks, and interestingly, has been found to display several degrees of non-Markovianity [52, 53]. According to the effective layer interpretation discussed above, we wonder whether such lack of Markovianity can also be interpreted as being the result of a Markovian dynamics taking place on a hidden multiplex network? Surely, in the case where mobility takes place across (hidden) multimodal transportation systems (as when we collect GPS traces of urban mobility), layers could be physical (underground, bus, car, etc). On the other hand, animal foraging dynamics are clearly different and alternate during day and night [54]. In a similar vein, human mobility patterns change when switching from work/leisure styles [55]: these would be cases where layers would be effective, rather than physical.

There are a large variety of problems involving the aforementioned scenarios which could be amenable to our approach. To guarantee computational efficiency, we would only require the network over which the agents move not to be too large, something that in the general case can be achieved by coarse-graining the network via community detection [56].

Another potentially interesting development is the more in-depth exploration of the relation between multiplex random walks and discrete flashing ratchets in discrete potential fields, which might constitute an unexpected contribution of network science to condensed matter physics and molecular biology. To be more concrete, it is well known that RNA polymerase (RNAP) are enzymes whose location spatially fluctuates inside the cell by switching (stochastically) between two different dynamical regimes: an active state (elongation) and a passive state (backtracking). While in elongation, RNAP's dynamics are closer to a ballistic motion or at least to a diffusion process with a net drift, whereas in backtracking regime RNAP is passively fluctuating due to several types of noise (thermal, chemical) and its movement is less biased. As RNAP traces inside the cell are rather noisy, it is a challenging task to identify the transition from elongation to backtracking and the general estimation of these dynamics from single particle trajectories. Under our approach, RNAP is akin to a molecular motor producing non-Markovian dynamics which can be effectively mapped into a Markov switching random walk dynamics in a multiplex with $L = 2$ layers (corresponding to elongation and backtracking respectively).

Finally, electronic transport in semiconductors can be modeled again as ratchets under the effect of electric potentials of several states. All these are just a few examples of realistic applications which can be addressed via our methodology, which will be at the core of further investigations.

Lisica and Vito Latora for fruitful discussions.

———————

[1] M. Newman. *Networks: An Introduction.* (Oxford University Press, New York, 2010).

[2] A. -L. Barabási. *Network Science.* (Cambridge University Press, Cambridge, 2016).

[3] R Pastor-Satorras, C Castellano, P Van Mieghem, and A Vespignani. Epidemic processes in complex networks *Rev. Mod. Phys.* **87**, 9259 (2015).

[4] C Castellano, S Fortunato, and V Loreto. Statistical physics of social dynamics. *Rev. Mod. Phys.* 81, 591 (2009).

[5] A Arenas, A Díaz-Guilera, J Kurths, Y Moreno, and C Zhou. Synchronization in complex networks. *Phys. Rep.* **469**, 93-153 (2008).

[6] S. Boccaletti, G. Bianconi, R. Criado, CI. Del Genio, J. Gómez-Gardeñes, M. Romance, I. Sendina-Nadal, Z. Wang, and M. Zanin. The structure and dynamics of multilayer networks. *Phys. Rep.* **544**, 1-122 (2014).

[7] M. Kivela, A. Arenas, M. Barthelemy, JP. Gleeson, Y. Moreno, and MA. Porter. Multilayer networks. *J. Complex Networks* **2**, 203-271 (2014).

[8] Battiston, F., Nicosia, V., Latora, V. The new challenges of multiplex networks: measures and models, *EPJST* 226, 3 (2017), pp. 401-416.

[9] M. Szell, R. Lambiotte and S. Thurner. Multirelational organization of large-scale social networks in an online world. *Proc. Nat. Acad. Sci. USA* **107**, 13636–13641 (2010).

[10] M. Barthélemy. Spatial networks. *Phys. Rep.* **499**, 1–101 (2011).

[11] A. Cardillo, J. Gómez–Gardeñes, M. Zanin, M. Romance, D. Papo, F. del Pozo, S. Boccaletti Emergence of network features from multiplexity. *Sci. Rep.* **3**, 1344 (2013).

[12] S. V. Buldyrev, R. Parshani, G. Paul, H. E. Stanley and S. Havlin. Catastrophic Cascade of Failures in Interdependent Networks. *Nature* **464**, 1025–1028 (2010).

[13] J. Gao, S. V. Buldyrev, H. E. Stanley, S. Havlin. Networks Formed from Interdependent Networks. *Nat. Phys.* **8**, 40–48 (2012).

[14] S. -W. Son, G. Bizhani, C. Christensen, P. Grassberger and M. Paczuski. Percolation Theory on Interdependent Networks Based on Epidemic Spreading. *EPL* **97**, 16006 (2012).

[15] G. Bianconi, S. N. Dorogovtsev. Multiple percolation transitions in a configuration model of a network of networks. *Phys. Rev. E,* **89**, 062814 (2014).

[16] F. Radicchi and A. Arenas, Abrupt transition in the structural formation of interconnected networks F. Radicchi and A. Arenas, *Nat. Phys.* 9, 717 (2013).

[17] F. Radicchi, Driving interconnected networks to supercriticality. *Phys. Rev. X*, 4(2), 021014 (2014).

[18] S. Gómez, A. Díaz-Guilera, J. Gómez-Gardeñes, C.J. Pérez-Vicente, Y. Moreno and A. Arenas. Diffusion dynamics on multiplex networks. *Phys. Rev. Lett.* **110**, 028701 (2013).

[19] De Domenico, M., Sole-Ribalta, A., Gomez, S. and Arenas, A. Navigability of interconnectednetworks under random failures. *Proc. Nat. Acad. Sci. USA* 111, 8351?8356 (2014).

[20] Sole-Ribalta, Gomez, S. and Arenas, A. Congestion induced by the structure of multiplexnetworks. *Phys. Rev. Lett.* **116**, 108701 (2016).

[21] Battiston, F., Nicosia, V., Latora, V. Efficient exploration of multiplex networks, *New Journal of Physics* **18**(4), 043035 (2016).

[22] A. Saumell-Mendiola, M. Á. Serrano and M. Boguñá. Epidemic spreading on interconnected networks. *Phys. Rev. E* **86**, 026106 (2012).

[23] Granell, C., Gomez, S. and Arenas, A. Dynamical interplay between awareness and epidemicspreading in multiplex networks. *Phys. Rev. Lett.* 111, 128701 (2013).

[24] C. Buono , L.G. Alvarez–Zuzek, P.A. Macri, L.A. Braunstein. Epidemics in Partially Overlapped Multiplex Networks. *PLoS One* **9**, e92200 (2014).

[25] J. Sanz, Ch.–Y. Xia, S. Meloni, Y. Moreno. Dynamics of Interacting Diseases. *Phys. Rev. X* **4**, 041005 (2014).

[26] J Gómez-Gardeñes, I Reinares, A Arenas, LM Floría. Evolution of cooperation inmultiplex networks. *Sci. Rep.* 2, 620 (2012).

[27] Z. Wang, A. Szolnoki, M. Perc. Rewarding evolutionary fitness with links between populations promotes cooperation. *J. Theor. Biol.* **349**, 50 (2014).

[28] JT Matamalas, J Poncela–Casasnovas, S Gómez, A Arenas. Strategical incoherence regulates cooperation in social dilemmas on multiplex networks *Sci. Rep.* **5**, 9519 (2015).

[29] Ch Del Genio, J Gómez-Gardeñes, I Bonamassa, S Bocaletti. Synchronization in networks with multiple interaction layers. *Science Advances* (in press).

[30] Diakonova, M., San Miguel, M., Eguíluz, V.M., Absorbing and shattered fragmentation transitions in multilayer coevolution, *Phys. Rev. E* **89**, 062818 (2014).

[31] Diakonova, M., Nicosia, V., Latora, V., San Miguel, M., Irreducibility of multilayer network dynamics: the case of the voter model, *New Journal of Physics* **18**, 023010 (2016).

[32] M De Domenico, C Granell, M. Porter and A. Arenas. The physics of spreading processes in multilayer networks. *Nat. Phys.* (in press 2016).

[33] M. De Domenico, V. Nicosia, A. Arenas, V. Latora. Structural reducibility of multilayer networks, *Nat. Commun.* **6**, 6864 (2015).

[34] M. Rosvall and C.T. Bergstrom, Maps of random walks on complex networks reveal community structure. *Proc. Natl. Acad. Sci. USA* 105(4) (2008) 1118-1123.

[35] T.P. Peixoto, Hierarchical block structures and high-resolution model selection in large networks,*Phys. Rev. X* **4**, 011047 (2014).

[36] T.P. Peixoto, Model selection and hypothesis testing for large-scale network models with overlapping groups,*Phys. Rev. X* **5**, 011033 (2015).

[37] M.E.J. Newman, and G. Reinert, Estimating the number of communities in a network. *Phys. Rev. Lett.* 117(7), 078301 (2016).

[38] T. Valles-Catala, FA Massucci , R. Guimera, and M. Sales-Pardo, *Phys. Rev. X* 6, 011036 (2016).

[39] M. Zanin, Can we neglect the multi-layer structure of functional networks?, *Physica A* 430, 184-192 (2015).

[40] R Gallotti, M Barthelemy. Anatomy and efficiency of urban multimodal mobility *Sci. Rep.* **4**, 6911 (2014).

[41] R Gallotti, MA Porter, M Barthelemy. Lost in transportation: Information measures and cognitive limits in multilayer navigation *Science Advances* **2**, e1500445 (2015).

[42] Drift induced by a spatially periodic potential of low symmetry-pulsed dielectrophoresis, A. Ajdari and J. Prost, *C.R. Acad. Sci. Ser. Gen. Ser. II* **315**, 13 (1992)

[43] F. Julicher, A. Ajdari, J. Prost, Modeling molecular motors, *Rev. Mod. Phys.* 69, 4 (1997)

[44] M.V. Smoluchowsky, *Phys. Z.* **13**, 1069 (1912).

[45] R.P. Feynman, R.B. Leighton and M. Sands, *The Feynman Lectures on Physics* (Addison-Wesley, Reading, MA, 1963) Vol 1, Sec. 46.1-46.9

[46] T.M. Cover, J.A. Thomas *Elements of Information Theory* (Wiley & Sons, New Jersey, 2006)

[47] N. Chopin, P. E. Jacob, and O. Papaspiliopoulos. SMC2: an efficient algorithm for sequential analysis of state space models. *J. Roy. Stat. Soc. B* 75, 3 (2013).

[48] P. j. Davis and P. Rabinowitz, *Methods of numerical integration*, Courier Corporation, 2007.

[49] M.C. Gonzalez, C.A. Hidalgo and A.L. Barabasi, Understanding individual human mobility patterns, *Nature* **453** (2008).

[50] C.E. Schneider, V. Belik, T. Couronne, Z. Smoreda and M.C. Gonzalez, Unravelling daily human mobility motifs, *J. R. Soc. Interface* **10** (2013).

[51] A. Lima, R. Stanojevic, D. Papagiannaki, P. Rodriguez, and M.C. Gonzalez, Understanding invidual routine behavior, *J. R. Soc. Interface* **13** (2016).

[52] Z-D Zhao, S-M Cai, Y Lu, Non-Markovian character in human mobility: online and offline, *Chaos* 25, 063106 (2015).

[53] D. Boyer, M.C. Crofoot, P.D. Walsh, Non-random walks in monkeys and humans, *J. Roy. Soc. Interface* doi:10.1098/rsif.2011.0582 (2011).

[54] JK Jansen, PL Boveng, JL Bengtson, Foraging modes of chinstrap penguins: contrasts between day and night, *Mar. Ecol. Prog. Ser.* 165 (1998) pp. 161-172.

[55] Martin Lanzendorf, Mobility Styles and Travel Behavior: Application of a Lifestyle Approach to Leisure Travel, *Transportation Research Record* 1807 (2002)

[56] M. Szell, R. Sinatra, G. Petri, S. Thurner, V. Latora, Understanding mobility in a social petri dish, *Sci. Rep.* 2, 457 (2012).

[57] L. R. Rabiner, A tutorial on hidden Markov models and selected applications in speech recognition. *Proceedings of the IEEE* **77**, 257–286 (1989).

[58] J. A. Fessler and A. O. Hero, *IEEE Transactions Signal Processing* **42**, 2664-2677 (1994).

[59] Z. Ghahramani, An introduction to hidden Markov models and Bayesian networks. *International Journal of Pattern Recognition and Artificial Intelligence* **15**, 9–42 (2001).

[60] S. M. Siddiqi, G. J. Geoffrey and A. W. Moore, Fast State Discovery for HMM Model Selection and Learning. *AISTATS*, 492–499 (2007).

[61] C.A. McGrory and D. M. Titterington, Variational Bayesian analysis for hidden Markov models. *Australian & New Zealand Journal of Statistics* **51**, 227–244 (2009).

[62] O. Cappé, R. Douc, A. Guillin, J. M. Marin and C. P. Robert, Adaptive Importance Sampling in General Mixture Classes. *Statistics and Computing* **18**, 447–459 (2008).

[63] E. Koblents and J. Míguez, A population Monte Carlo scheme with transformed weights and its application to stochastic kinetic models. *Statistics and Computing* **25**, 407–425 (2015).