Discrete-time distributed consensus on multiplex networks

I Trpevski\textsuperscript{1}, A Stanoev\textsuperscript{1,2}, A Koseska\textsuperscript{2} and L Kocarev\textsuperscript{1,3,4}

\textsuperscript{1} Research Centre for Computer Science and Information Technologies, Macedonian Academy of Sciences and Arts, Skopje, Macedonia
\textsuperscript{2} Department of Systemic Cell Biology, Max Planck Institute for Molecular Physiology, Germany
\textsuperscript{3} Faculty of Computer Sciences and Engineering, Skopje, Macedonia
\textsuperscript{4} BioCircuits Institute, University of California, San Diego, CA, USA

E-mail: aneta.koseska@mpi-dortmund.mpg.de

Received 15 August 2014, revised 29 September 2014
Accepted for publication 10 November 2014
Published 26 November 2014

New Journal of Physics 16 (2014) 113063
doi:10.1088/1367-2630/16/11/113063

Abstract

We introduce a discrete-time distributed consensus process on multi-layered complex networks represented by multiplex graphs. The proposed consensus process can be characterized with a multiplex Markov chain (MC) composed of a mixture of interlayer and intralayer MCs. The interlayer MC is characterized by a parameter which represents the probability that the chain will switch between different layers of the multiplex network. Surprisingly, for large regions of the parameter space the convergence speed of the multiplex MC is determined only by the convergence speed of the interlayer MC. Moreover, as the number of layers in the network is increased, these regions increase until they encompass the whole parameter space. We also show that for some regions of the parameter space, the multiplex MC has a faster convergence speed than the MCs on top of the individual layers, thus complementing previous results on diffusion on multiplex networks.

Online supplementary data available from stacks.iop.org/njp/16/113063/mmedia

Keywords: consensus algorithm, multiplex networks, Markov chains
1. Introduction

Reaching an agreement or a consensus regarding a certain quantity of interest is a general problem in various disciplines, ranging from social and management sciences [1], control theory [2–4] to various physics disciplines [5]. Formally, given a set of autonomous agents, the distributed consensus problem asks for a distributed rule that the agents can use to agree on an ‘opinion’ represented by a scalar or a vector value, starting from different initial options (states of the agents). For example, nodes in sensor networks may be tasked with calculating the average measurement value of all the sensors, or, as in the case of flocking of multi-agent systems, all agents communicate with each other to match their speed and direction.

The connection between consensus problems and physics is both fundamental and deeply profound. The prototypical example of a continuous consensus process studied by computer scientists and control theorists [2] is one where the state of an agent in the network changes according to a linear differential equation which is mathematically equivalent to the process of diffusion through a network known to physicists [6]. The vector equation in both cases is $\dot{x} = -cLx$, where $L$ is the graph Laplacian matrix. The continuous consensus process is further related to the theory of synchronization in coupled oscillator networks [5]. In fact, the linearization of the generalized Kuramoto model of coupled network oscillators has the same form as the previous linear equation. As another example, the classical model for emergence of coordinated movement of self-driven particles by Vicsek et al [7] in statistical physics has been used as a foundation for a rigorous theory on flocking phenomena by control theorists [8, 9].

However, real-world networks do not function in isolation. In contrast, multiple layers are usually interconnected to achieve a coordinated behavior [10–12]. Let us consider as an example a model opinion formation in a network of $n$ people discussing matters on different topics. In order to reach a consensus on a given question, the agents exchange information exclusively among available connections (each person listens to, or trusts, some people in the network more than others), thus forming an average opinion. This opinion however, can be influenced by the opinion formation on another question, represented by a different network layer with characteristic topology (different interactions between the same $n$ agents). Mathematically this can be represented as a multiplex graph, which provides formal ground to analyse systems in which each entity is present in multiple layers, but with a different connectivity pattern [13]. The emergent physics of dynamical processes on top of multiplex networks so far incorporated diffusion processes to show an enhanced-diffusive behavior [14], or random walk processes to address the problem of navigability of multiplex network under random failures [15]. However, how consensus can be reached in such a complex network constellation remains unclear. Additionally, it is significantly important to address the question regarding the amount of communication required and thereby the convergence speed of the consensus process, since the consensus is calculated in an iterative and distributed manner.

Here, we propose a discrete-time consensus process that can be analyzed through its dual process, a Markov chain (MC) on top of a multiplex network. In fact, this multiplex MC can be thought of as a combination of the individual MCs on top of each layer and a MC that represents the transition rules between different layers. The interlayer MC is characterized by a parameter $p$ which expresses the probability that the MC will stay in the same layer. Strikingly, for large regions of the parameter space the asymptotic convergence rate of the distributed consensus process is determined only by the convergence rate of the interlayer MC. Moreover,
we show that this region increases until it comprises the whole parameter space, as the number of layers in the network increases.

The remainder of this paper is organized as follows: in section 2 we first present the theoretical background on consensus on single-layer networks and the definition of multiplex networks. The problem of discrete-time distributed consensus on multiplex graphs is introduced in section 3.1, containing as well the analytical results on the convergence of the consensus process 3.2. In the section 3.3 we provide results on the numerical analysis of the convergence rate of the consensus process. First, in the section 3.3.1 we provide the analysis on multiplex networks with two layers, followed by a generalization for an arbitrary number of layers in section 3.3.2. Finally, section 5 summarizes the main results and gives an outline of future directions and possible applications of the proposed process. The appendices provide formal proofs for some statements in the text.

2. Theoretical background

In this section, a brief overview of the theory on consensus and cooperation in networked multi-agent system, as well as the general definitions of multiplex graphs are given. We specifically underline here the distinction between a random walk and consensus on a multiplex graph, pointing their general properties.

2.1. Consensus on single layer networks

Let \( x_i \in \mathbb{R}, i = 1, \ldots, n \) be real scalar quantities (variables), and let their time evolution be described by a linear expression that can be written as:

\[
x_i(t + 1) = p_{i1}x_1(t) + p_{i2}x_2(t) + \ldots + p_{in}x_n(t) = \sum_{j=1}^{n} p_{ij}x_j(t), \forall i
\]

or in a more compact form as

\[
x(t) = P x(t - 1) = P' x(0),
\]

where \( x(t) = [x_1 \ldots x_n]^T \) describes the state of the system at time \( t \), \( P \) is a \( n \times n \) transition matrix describing the linear dynamics of the process, and \( x(0) \) is the vector of initial states. Let us order the set of \( n \) eigenvalues of \( P \) by magnitude such that \( |\lambda_1| \geq |\lambda_2| \geq \ldots \geq |\lambda_n| \) and the set of corresponding eigenvectors accordingly \( v_1, v_2, \ldots, v_n \). If the matrix \( P \) is diagonalizable (i.e. it has \( n \) linearly independent eigenvectors \( v_i \)), the initial vector can be expressed as a linear combination of the eigenvectors \( x_0 = \sum_i q_i v_i \), for some nonzero constants \( q_i \). For non-diagonalizable matrices it is necessary to start with the Jordan canonical form of the matrix \( P \) and proceed through a more involved set of similar arguments using the generalized eigenvectors of \( P \). Combining the expansion of the initial vector with (2), the following expression is obtained:

\[
x(t) = P'^t (q_1 v_1 + q_2 v_2 + \ldots + q_n v_n)
\]

\[
= q_1 \lambda_1^t v_1 + q_2 \lambda_2^t v_2 + \ldots + q_n \lambda_n^t v_n.
\]

This allows us to deduce the dynamical behavior of the system: (i) if all \( \lambda_i \) satisfy \( |\lambda_i| < 1 \), the solution of the system \( x^* = \lim_{t \to \infty} x(t) \) converges to \( 0 \); (ii) if at least one \( \lambda_i \) satisfies \( |\lambda_i| > 1 \), then \( x^* \) will have at least one component which grows exponentially; (iii) the only case when
the asymptotic behavior of (3) is different than trivial (zero or infinity) is when the largest eigenvalue $\lambda_1 = 1$. In addition, when $\lambda_1$ is strictly larger than all the other eigenvalues (i.e. $|\lambda_1| > |\lambda_i|$ for $i = 2, 3, \ldots, n$), we say that $\lambda_1$ is the dominant eigenvalue. Thus, rearranging the previous expression in the following form:

$$\mathbf{x}(t) = \lambda_1^t \left( q_1 \mathbf{v}_1 + q_2 \frac{\lambda_2}{\lambda_1} \mathbf{v}_2 + \ldots + q_n \frac{\lambda_n}{\lambda_1} \mathbf{v}_n \right)$$

(4)

allows one to note that the terms $(\lambda_i/\lambda_1)^t$ decay exponentially as $t \to \infty$, and $\mathbf{x}^* = q_1 \lambda_1^t \mathbf{v}_1 = q_1 \mathbf{v}_1$. Here, the rate of convergence is determined by the second largest eigenvalue, since $(\lambda_2/\lambda_1)^t$ decays more slowly than all other terms.

A large class of stochastic matrices describing MC processes on top of networks have an eigenvalue with magnitude 1. The well known Perron–Frobenius theorem [16] states that if the stochastic matrix is irreducible and primitive, i.e. the MC is aperiodic, the only eigenvalue on the unit circle is 1, hence it is the dominant eigenvalue.

Let us consider now an undirected and connected network, described by its $n \times n$ symmetric adjacency matrix $A = [a_{ij}]$, where $n$ is the number of nodes. By definition, $a_{ij}$ is the topological weight of the edge $ij$, and $s_i = \sum_j a_{ij}$ is the strength of node $i$ representing the total weight of the links connected to it. If the network is unweighted, $s_i$ is simply the degree of node $i$. Two of the simplest dynamical processes that can be described by MCs on such networks are the random walk and its dual process, consensus. Both can be represented by equation (2). Here, the evolution of the quantity $x_i$ associated with node $i$ is driven by $P = [p_{ij}]$—a transition matrix related to the adjacency matrix $A$. The crucial difference between the two processes is that for a random walk $P$ is column stochastic, $p_{ij} = a_{ij}/s_j$, whereas for consensus it is row stochastic, $p_{ij} = a_{ij}/s_i$. Furthermore, for a random walk, the states $x_i(t)$ denote probabilities that a random walker can be found in node $i$ at time $t$ and thus sum up to 1. For the consensus problem on the other hand, these values can describe any scalar quantity. A goal of a consensus process is therefore to unify these quantities to an equal value $x_1 = x_2 = \ldots = x_n$. We use the following notation: $\pi$ denotes the left eigenvector of $P$ associated with the largest eigenvalue of $P$ (i.e. $P^T \pi = \lambda_1 \pi$), $\mathbf{1}_n$ is a length $n$ column vector of 1s, and $\otimes$ denotes the Kronecker product of matrices (or vectors). According to the Perron–Frobenius theorem (see appendix A), if the transition matrix is irreducible and primitive, i.e. the graph is connected and its periodicity equals 1, the random walk and the consensus process converge to their solutions [17]:

$$\mathbf{x}(t) = P^t \mathbf{x}(0)$$

$$\rightarrow \begin{cases} 
(\pi \otimes \mathbf{1}_n^T) \mathbf{x}(0) & P \text{ is column stochastic} \\
(\pi^T \otimes \mathbf{1}_n) \mathbf{x}(0) & P \text{ is row stochastic} 
\end{cases}$$

(5)

$$= \begin{cases} 
\pi \mathbf{1}_n^T \mathbf{x}(0) & P \text{ is column stochastic} \\
\mathbf{1}_n \pi^T \mathbf{x}(0) & P \text{ is row stochastic}, 
\end{cases}$$
that is

\[ x_i(t) \to \pi_i \sum_{j=1}^{n} x_j(0) \quad \text{column stochastic (random walk),} \tag{5} \]

\[ x_i(t) \to \sum_{j=1}^{n} \pi_j x_j(0) \quad \text{row stochastic (consensus).} \tag{6} \]

As determined from equation (5), the random walk solution which gives the fraction of time the random walkers spent in each node, depends only on the number of concurrent random walkers in the graph, i.e., \( \|x(0)\| = \sum x_i(0) \). On the other hand, equation (6) shows that the consensus solution depends on the initial vector \( x(0) \), and each node’s opinion converges to a mixture of the initial opinions in the network. This mixture is a non-trivial weighted average of the opinions, unlike the Laplacian-based consensus whose solution is always the average-consensus for undirected graphs. Also, the number of walkers in time \( t \) is usually 1, while in consensus processes, the aggregated opinion, i.e. the weighted average opinion in the network is constant, \( \pi^T x(t) = \text{const.} \). Notice, however, the dual nature of the processes. If the equation \( t P x(0) \) describes a consensus process, than the equation \( t P^T x(0) \) describes a random walk process (or some other MC). The convergence properties of both processes are directly related because they are described by the same transition matrix and the limit of the product \( \lim_{t \to \infty} P^t \).

In this work we focus mainly on consensus and aim to generalize the discrete consensus process for multiplex graphs, while preserving the form (2). We therefore define a novel supra-transition matrix that describes the multiplex consensus process. If this matrix is row stochastic, irreducible and primitive, the consensus process should converge to a single value, as shown below. Moreover, the convergence speed of the process can be tested through the second largest eigenvalue of the supra-transition matrix. Therefore, in the remaining part of this section we provide the general theory on multiplex networks.

### 2.2. Definition of multiplex graphs

A multiplex graph is a multilayer network where the vertices belong to several different single-layer networks simultaneously, connected by means of a specific set of edges in each layer, as well as between them. Let \( \{1, \ldots, n\} \) denote the set of nodes, where the elements of this set belong to \( L \) different layers. Thus, each layer is a graph with \( n \) vertices. We denote with \( G^\alpha = (V^\alpha, E^\alpha) \) the graph which represents the layer \( \alpha \), where \( V^\alpha = \{1^\alpha, \ldots, n^\alpha\} \) and \( \alpha = 1, \ldots, L \). Let \( A^\alpha = [a^\alpha_{ij}] \) be the \( n \times n \) adjacency matrix of the layer \( \alpha \), that is, the graph \( G^\alpha \). In general, the layers are interconnected such that for each node \( i \), a graph \( G_i = (V_i, E_i) \) is defined, with \( V_i \) being the set of \( L \) vertices. Each vertex represents the state of the node in one of the layers, and \( E_i \) denotes the set of edges which describe how the layers are connected. Then, \( B_i = [b_i^{\alpha\beta}] \) represents the \( L \times L \) adjacency matrix of the graph \( G_\alpha \), defined as \( b_i^{\alpha\beta} = 1 \), if and only if the layers \( \alpha \) and \( \beta \) are connected, otherwise \( b_i^{\alpha\beta} = 0 \). Self loops in the graphs are also allowed. For simplicity, we assume that all graphs \( G_i \) and \( G^\alpha \) are undirected and connected. A multiplex graph \( G = (V, E) \) is then defined as a union of all graphs \( G_i \) and \( G^\alpha \): \( V = \bigcup_{\alpha=1}^{L} V^\alpha \) and \( E = \bigcup_{i=1}^{L} E_i \bigcup_{\alpha=1}^{L} E^\alpha \). An illustration of a multiplex graph consisting of two layers, with six nodes each, is given in figure 1.
3. Results

3.1. Discrete-time distributed consensus on multiplex networks

To introduce the problem of consensus on multiplex networks, we define formally the model of opinion formation in a network of \( n \) people discussing a given topic. Assume that each agent (node) \( i \) is actively involved in \( L \) different discussions (layers) \( G^{\alpha} \), \( \alpha = 1, \ldots, L \). We write \( x_i^{\alpha} \) for her/his opinion obtained on a given topic (specific layer) \( G^{\alpha} \). Since a person listens to (trusts) some people more than others, for the network \( G^{\alpha} \) we use an \( n \times n \) matrix \( D^{\alpha} \), where \( d_{ij}^{\alpha} \) denotes the relative influence person \( j \) has on person \( i \). Here, \( d_{ii}^{\alpha} \) can describe the stubbornness (resistance to other people’s opinion) of person \( i \). Consequently, the \( i \)th row in matrix \( D^{\alpha} = [d_{ij}^{\alpha}] \) contains the relative influences people have on person \( i \) in the network \( G^{\alpha} \), and thus it sums to 1. Therefore, the matrix \( D^{\alpha} = [d_{ij}^{\alpha}] \) is row stochastic.

Each person’s belief or viewpoint is an aggregation of opinions on \( L \) different questions (layers), and interactions can be more complex when the opinion in one layer can be influenced by the opinion formation in another layer. To incorporate this, for each person \( i \) we introduce an \( L \times L \) matrix \( C_i = [c_{ij}^{\alpha\beta}] \) which describes the dynamics of how opinions change after interaction. More specifically, we assume that the person opinion is a weighted average of the opinions in \( L \) networks, with the weights encoded in row \( \alpha \) of the matrix \( C_i = [c_{ij}^{\alpha\beta}] \). Therefore, the matrix \( C_i \) is again row stochastic. For simplicity, in the remainder of this section only models where the interlayer influence is the same for every node i.e. \( C_i = C \) for \( i = 1 \ldots N \) are used.

Thus, the distributed consensus problem for a multiplex graph is to find a set of rules which allow all nodes to compute a value \( x^* \) such that \( \lim_{t \to \infty} x_i^{\beta} (t) = x^* \), for all \( i \) and \( \beta \). In general, the limit \( \lim_{t \to \infty} x_i^{\beta} (t) \) may depend on \( i \) and \( \beta \). To address this problem, we use here the more general formalism of MCs, since the discrete time distributed consensus process can be considered as the dual process of an MC \([2, 18]\). Let \( C = [c_{ij}^{\alpha\beta}] \) be the transition matrix.
describing an MC (interlayer MC) on the graph $G_L$ that denotes the interlayer connectivity. The elements of the matrix give the transition probabilities from one layer to another. Then let $D^\alpha = [d^\alpha_{ij}]$, $\alpha = 1 \ldots L$ be the transition matrix that describes the MC on each individual layer $\{G^\alpha\}$, whose elements give the transition probabilities between the nodes in a single layer $\alpha$. This allows us to define the updating rules for the $i$th node in layer $\alpha$ as:

$$
\begin{align*}
    x_i^\alpha(t+1) &= \sum_{\beta=1}^{L} c^{\alpha\beta} \sum_{j=1}^{n} d^\beta_{ij} x_j^\beta(t) = c^{\alpha\alpha} d^\alpha_{ii} x_i^\alpha(t) + \sum_{j=1, j \neq i}^{n} c^{\alpha\alpha} d^\alpha_{ij} x_j^\alpha(t) \\
    &+ \sum_{\beta=1, \beta \neq \alpha}^{L} c^{\alpha\beta} d^\beta_{ii} x_i^\beta(t) + \sum_{\beta=1, \beta \neq \alpha}^{L} \sum_{j=1, j \neq i}^{n} c^{\alpha\beta} d^\beta_{ij} x_j^\beta(t).
\end{align*}
$$

(7)

Where:

(i) $c^{\alpha\alpha} d^\alpha_{ii}$ is the term that defines the contribution of node $i$ in layer $\alpha$ to the state $x_i^\alpha(t+1)$, i.e. $x_i^\alpha(t) \to x_i^\alpha(t+1)$ (in terms of the MC this is the probability that the chain remains in node $i$ in layer $\alpha$);

(ii) $c^{\alpha\alpha} d^\alpha_{ij}$ is the term that defines the contribution of node $j$ in layer $\alpha$ to the state $x_i^\alpha(t+1)$, i.e. $x_j^\alpha(t) \to x_i^\alpha(t+1)$ (in terms of the MC this is the probability that the chain switches its state from node $i$ to node $j$ in layer $\alpha$);

(iii) $c^{\alpha\beta} d^\beta_{ii}$ defines the contribution $x_i^\beta(t) \to x_i^\alpha(t+1)$ (in terms of the MC this is the probability that the chain switches its state from node $i$ in layer $\alpha$ to node $i$ in layer $\beta$);

(iv) $c^{\alpha\beta} d^\beta_{ij}$ defines the contribution $x_j^\beta(t) \to x_i^\alpha(t+1)$ (in terms of the MC this is the probability that the chain switches its state from node $i$ in layer $\alpha$ to node $j$ in layer $\beta$).

In defining the product (iv), we use the fact that in multiplex networks [15] each node is present in multiple layers and we can assume that information transfer is faster between the same nodes in different layers than between different nodes in the same layer. For example, a person can adapt/change his opinion on a given question (discussed in layer $L = 1$), depending on the consensus opinion on a different question (layer $L = 2$). Thus, the information transfer between the layers is essentially faster than the information transfer within a layer, i.e. if the person changed/adapted his/her opinion based on the discussion with other people within the same layer. Moreover, the system (7) can be rewritten in vector form as

$$
y(t+1) = Py(t) = P^T y(0),
$$

(8)

where $y^T = [y_1, y_2, \ldots, y_n, y_{n+1}, \ldots y_{nL}]$ is a vector composed of the states of all nodes in the different layers of the multiplex network with $y_1 = x_1^1$, $y_2 = x_2^1$, $\ldots$, $y_{nL} = x_n^L$. $P$ is an $nL \times nL$ matrix that we term as a $\alpha$ supratransition matrix. With a slight abuse of notation, the supa-transition can be written in the following form, reminiscent to the classical Kronecker product of two matrices:

$$
P = C \otimes D^\alpha = \begin{pmatrix}
    c^{11} D^1 & c^{12} D^2 & \cdots & c^{1L} D^L \\
    c^{21} D^1 & c^{22} D^2 & \cdots & c^{2L} D^L \\
    \vdots & \vdots & \ddots & \vdots \\
    c^{L1} D^1 & c^{L2} D^2 & \cdots & c^{LL} D^L
\end{pmatrix}.
$$

(9)
Note that the matrix $P$, which describes a multiplex network, could also be treated as a matrix of a (large) single layer network with $nL$ nodes. In this case, the system (8) has the form (2), and therefore one can use well-known equations (5) and (6) to describe the asymptotic behavior of (8).

3.2. Convergence of the consensus process

The following theorem states the rules for convergence of the multiplex consensus process. The details of the proof are given in appendix A. In the following we let $1_{nL}$ be a vector of ones of length $nL$.

**Theorem 1.** Let $P$ be a row stochastic, irreducible, and primitive matrix describing the multiplex consensus process 7 and $y(0)$ be the vector representing any initial conditions of the system. Then, the product sequence $P^t$ in (8) converges to $1_{nL} \pi^T$, and at each node, the system (8) asymptotically converges to the weighted average $\sum_j \pi_j y_j(0)$, where $\pi$ is the left eigenvector of the matrix $P$ associated with $\lambda_1 = 1$. In addition, (ii) the average value $\sum_j y_j(0)/(nL)$ can also be computed at each node.

According to the theorem, the product sequence $P^t$ in (8) converges to $1_{nL} \pi^T$, if the matrix $P$ satisfies all three conditions. We can use this in equation (6) to see that the solution of the system (7) converges to the consensus value given by a weighted average of the initial vector $y(0)$ and the vector $\pi$:

$$y^* = (\pi^T \otimes 1_{nL}) y(0) = 1_{nL} \pi^T y(0) = 1_{nL} \sum_{j=1}^{nL} \pi_j y_j(0) = \sum_{i=1}^n \sum_{\alpha=1}^L \pi_i^\alpha x_i^\alpha(0).$$ (10)

Next, we provide two other results that follow from this theorem. Consider the dual process of the multiplex consensus given by the transpose of the supratransition matrix $P^T$. We show that by employing two parallel runs of the dual process with different initial values, the nodes can calculate the average of the initial conditions vector $y(0)$.

The nodes run two algorithms $u(t+1) = P^t u(t)$ and $w(t+1) = P^t w(t)$ with initial conditions $u_i(0) = y_i(0)$ and $w_i(0) = 1$ for all $i$. Let $r_i^u = \lim_{t \to \infty} u_i(t)$ and $r_i^w = \lim_{t \to \infty} w_i(t)$. Since $P$ is row stochastic, its transpose is column stochastic (from equation (5)). Thus, at each node $i$ in the limit $t \to \infty$, one has:

$$\frac{r_i^u}{r_i^w} = \frac{\pi_i \sum_j y_j(0)}{\pi_i nL} = \frac{1}{nL} \sum_j y_j(0).$$ (11)

Therefore, we can state that both the weighted average $\sum_j \pi_j x_j(0)$ and the average $\sum_j x_j(0)/(nL)$ at each node can always be calculated, and the differences in the decisions made can be compared using these two averages.

Additionally, by running the algorithm with different initial conditions, using only local communication and computation on the graph, at each node one can compute the number of
nodes in the graph $nL$ (global graph characteristic). Indeed, setting $z_1(0) = 1$, and $z_2(0) = \ldots = z_n(0) = 0$ and $r_i^z = \lim_{t \to \infty} z_i(t)$ where $z(t + 1) = P^T z(t)$, we have $r_i^z = \pi_i$, and therefore, $r_i^w/r_i^z = nL$ at each node.

3.3. Speed of convergence of the consensus process

Apart from the question on the feasibility of reaching a consensus, the second most important question here is the speed of convergence of the consensus process. From equation (4) it follows that the convergence speed of any linear process is determined by the ratio $\lambda_1/|\lambda_2|$. The dominant eigenvalue for stochastic matrices is $\lambda_1 = 1$, thus, the convergence speed is completely determined by the magnitude of $\lambda_2$. As already noted, the convergence is faster for smaller values of $|\lambda_2|$.

To investigate this, we next define the matrix $C$ as a function of a parameter $p$ and explore the dependence of the eigenvalues and eigenvectors of the supratransition matrix $P$ on this parameter ($1 - p$ can be interpreted as the probability with which the MC switches between layers). We are mostly interested in the dependence of the modulus of the second largest eigenvalue of the supratransition matrix $\lambda_2$ on this parameter, since it determines the convergence speed of the consensus process [2].

3.3.1. The case of two layers. We first demonstrate results for the numerical analysis of the convergence rate of the consensus process on multiplex networks consisting of two layers. In this special case, the transition matrix of the interlayer MC is a $2 \times 2$ matrix:

$$C = \begin{pmatrix} p & 1 - p \\ 1 - p & p \end{pmatrix}$$

with eigenvalues $\lambda_1(C) = 1$ and $\lambda_2(C) = -1 + 2p$, whereas the eigenvectors are $v_1 = [1, 1]^T$ and $v_2 = [1, -1]^T$, respectively. The parameter $p$ varies in the range $[0, 1]$ and $1 - p$ can be interpreted as the probability that the MC process transitions from one layer to another. The two consensus processes defined on top of the two single layers are characterized by MCs with transition matrices $\{D^\alpha\}$, whose elements can be defined as:

$$d^\alpha_{i,j} = \begin{cases} 1/(s_{\text{max}}^\alpha + 1), & a^\alpha_{i,j} = 1, \ i \neq j, \\ 0, & a^\alpha_{i,j} = 0, \\ 1 - \sum_{l \neq j} d^\alpha_{l,j}, & i = j, \end{cases}$$

(12)

where $s_{\text{max}}^\alpha$ is the maximum strength of all nodes in layer $\alpha$, for $\alpha = 1, 2$. In figure 2(a), the modulus of the second largest eigenvalue $\lambda_2(P)$ of the transition matrix is given as a function of the parameter $p$ for a multiplex network consisting of two Erdos–Renyi networks with 200 nodes each, and a probability of edge $q = 0.1$. Large intervals of the parameter $p$ exist where $\lambda_2(P)$ is identical to the second largest eigenvalue (by modulus) of the inter-layer transition matrix $C$. Thus, this shows that the convergence rate of the whole transition matrix in these intervals is completely determined by the convergence rate of the inter-layer consensus.

Note that the modulus of the eigenvalue $\lambda_2 = 1 - 2p$ decreases in the first half of the range $[0, 1]$, and increases in the second half. In fact, half of the eigenspectrum of the supratransition matrix vanishes (goes to 0) at $p = 1/2$, since the first and the second $N$ rows of the matrix
become identical and its rank is halved. Also, the boundary conditions when \( p = 1 \) corresponds to two disconnected networks, whereas the case when \( p = 0 \) corresponds to an MC on a bipartite network.

At the point where \( \lambda_2(C) \) becomes less than the second largest eigenvalue of \( P \), it crosses another eigenvalue of \( P \) which is a monotonically increasing function of the parameter \( p \). At this crossing, there is an abrupt transition in the corresponding eigenvector similar to that reported in [19], as seen from figures 2(b) and (c). Indeed, before the crossing, the eigenvector corresponding to \( \lambda_2(P) \) is \([1_n, -1_n]\) (up to a multiplicative constant) where \( 1_n \) is a vector of ones with length \( n \). The eigenvector \([1_n, -1_n]\) is the tensor product of the eigenvector corresponding to \( \lambda_2(C) \) and the eigenvector \( 1_n \) of the transition matrices \( \{D^\alpha\} \) (which is always an eigenvector for a row stochastic matrix since its rows sum to one). In fact, it can be proven that the eigenvalues of an arbitrary matrix \( C \) are also eigenvalues of the supratransition matrix \( P \) when the transition matrices \( \{D^\alpha\} \) have a common eigenvalue and eigenvector (see appendix B). Additionally, there is a second crossing which occurs for larger values of \( p \) where the second largest eigenvalue of \( P \) again becomes \( 1 - 2p \). As in the previous case there is also an abrupt transition in corresponding eigenvector.

Another important observation is that there is a big range of values of \( p \) for which the multiplex consensus process has a faster convergence speed than the consensus processes of the individual layers. This is similar to the results for diffusion dynamics on multiplex networks [14], where the multiplex structure gives rise to a faster super-diffusion process by speeding up the slowest diffusive process.

We investigated again the speed of convergence towards consensus, but in this case, the intralayer transition probabilities are defined as:

\[
d_{i,j}^\alpha = \begin{cases} 
1/(s_i^\alpha + 1), & a_{i,j}^\alpha = 1, \ i \neq j, \\
0, & a_{i,j}^\alpha = 0, \\
1 - \sum_{i \neq j} d_{i,j}^\alpha, & i = j,
\end{cases}
\]  

(13)

where \( s_i^\alpha \) is the strength of node \( i \) in the layer \( \alpha \). Again, similar results were obtained, as seen from figure 3. The range of values for \( p \) for which \( \lambda_2(C) \) is the second largest eigenvalue of \( P \) is
even larger than in the previous case. Moreover, this type of consensus process converges faster to its stationary distribution due to the lower $\lambda_2$ values.

### 3.3.2. The case of arbitrary number of layers

Consider next an arbitrary topology of layers represented by a graph $G_L$ whose adjacency matrix is again $B = [b_{\alpha,\beta}]$. We define an associated consensus process on top of the interlayer graph as a function of the parameter $p$ which can be interpreted as the probability that the MC will not switch its state to another layer during one time step. The compatible transition matrix denoted by $C(p) = [c_{\alpha,\beta}(p)]_{L \times L}$ is row stochastic and its transition probabilities are defined as:

$$c^{a,\beta}(p) = \begin{cases} 
(1 - p)/s_{a,\alpha}, & b_{a,\beta} = 1, \alpha \neq \beta, \\
p, & \alpha = \beta,
\end{cases}$$  \hspace{1cm} (14)

where $s_{a,\alpha}$ is the strength of the nodes in terms of the different layers, i.e. the sum of the weights of the links connecting one of the nodes to the corresponding nodes in different layers. We can thus think of the discrete-time distributed consensus process as a dual process to the MC composed of individual MCs on top of each layer that are mixed through an interlayer MC defined by equation (14). A nice property of equation (14) is that in the boundary case when $p = 1$, the multiplex consensus reduces to $L$ disconnected consensus processes, and when $p = 0$, to a consensus process on an $L$-partite network.

We have performed simulations studies for different numbers of layers with different interlayer and intra-layer topologies to investigate the speed of convergence towards consensus in this case as well. The most surprising result is that the range of values of $p$ for which the second largest eigenvalue of the matrix $P$ is the same as the second largest eigenvalue of the inter-layer transition matrix $C(p)$ increases with the number of layers in the multiplex network, regardless of their topology and interlayer connectivity. In other words, the convergence speed of the entire MC is generally determined by the convergence speed of the interlayer MC, as the number of layers in the network is increased.

In figure 4 we show the second largest eigenvalues of the matrices $C$ and $P$ for multiplex networks with different number of layers. The particular configuration of the inter-layer...
network is a circle graph in this case, while the processes defined on top of the individual layers are given by equation (12). The size of the layers is 200 nodes and half of them are ER networks, whereas the remaining half are Barabasi–Albert (BA) networks. It can be seen that as the number of layers in the multiplex network is increased, the convergence speed of the multiplex MC is dominantly determined by the convergence speed of the inter-layer MC. (See the supplementary material at stacks.iop.org/njp/16/113063/mmedia for more numerical results that support this claim on multiplex networks with different interlayer organization and for different network sizes.) Another interesting observation in this particular case is that the convergence rate decreases as the number of layers is increased, since the second largest eigenvalue of the matrices $C$ and $P$ is increasing. This is not always true, however and it is likely that it will depend on the particular interlayer topology.

4. Related work

To the best of our knowledge we do not know of any other consensus processes that have been extended to multiplex networks. The only related work that we are aware of is [15] where the authors used discrete random walks on multiplex networks to address the problem of navigability of a multiplex network under random failures. The multiplex consensus process
presented here is the dual process of the multiplex random walk processes which is investigated in [15]. However, instead of the supraterransition matrix $P$ as defined here, the authors in [15] use a normalized supra-Laplacian matrix $L$. We found that the two matrices are related as $P = I - L$. Additionally, the eigenvectors of the two matrices are identical and the corresponding eigenvalues are related as $\lambda_i = 1 - \mu_i$, where $\mu_i$ are the eigenvalues of the matrix $L$. To show this, let us assume that $v$ is an eigenvector of $P$, i.e., $Pv = \lambda v$. Then, using the relation $L = I - P$

$$Lv = (I - P)v = v - \lambda v = (1 - \lambda)v.$$ (15)

Thus, $v$ is also an eigenvector of $L$ for the eigenvalue $(1 - \lambda)$. Since both matrices and their eigenvalues are directly related, our study indicates that the obtained results on the convergence rate of the interlayer and the multiplex MC can be extended to the random walk processes presented in [15]. In particular, for random walks on multiplex graphs for which the interlayer matrix is given by equation (14), the second eigenvalue of the normalized supra-Laplacian matrix $L$ depends only on the second eigenvalue of the corresponding (normalized Laplacian) interlayer matrix. Therefore, using the results presented here, it is possible to design a navigation process in multiplex graphs for which the convergence rate depends only on interlayer matrix so that a complete control on the speed of the navigability through the network can be established.

5. Conclusions

In this paper we introduced a discrete distributed consensus process on multiplex graphs. Even in this general setting when the nodes belong to $L$ different layers represented by $n \times n$ graphs $G^\alpha$, $\alpha = 1, \ldots, L$, and interdependencies among layers are described with a $L \times L$ graph $G_L$, $i = 1, \ldots, n$, we have suggested a consensus process that reaches both a weighted average $\pi = \sum_{j=1}^{nL} \pi_j y_j(0)$ and the standard average $\bar{y} = \frac{1}{nL} \sum_{i=1}^{n} \sum_{\alpha=1}^{L} x_i^\alpha(0)$.

We have performed detailed numerical analysis on the asymptotic convergence rate of the consensus process defined on multiplex networks and we can make several notable conclusions. First, we observe a region for the parameter $p$, where the convergence speed is completely determined by the second largest eigenvalue of the interlayer transition matrix $C$. Second, as we increase the number of layers in the network we observe that this region increases until it finally represents the whole range of possible values for $p$ regardless of the topology of the different layers and of the interlayer topology. Finally, there is also a region in this parameter space where the consensus protocol on the whole multiplex network has a faster convergence rate than those of the individual layers which complements previous results on diffusion dynamics in multiplex networks.

Appendix A. Proof of convergence of the multiplex consensus process

The proof relies on the Perron–Frobenius theorem and several other theorems that have been proven elsewhere [16]. We first provide several definitions.

A matrix $P_{n \times n} = [p_{ij}]$ is compatible with the graph $G = (V, E)$, if (i) $p_{ij} \neq 0$ when $i \neq j$ and $(i, j) \in E$ and (ii) $p_{ii} \neq 0$. A transition matrix $P_{n \times n} = [p_{ij}]$ is a row stochastic matrix whose
entries are the probabilities for transition from node \( j \) to node \( i \). A matrix \( P_{n \times n} \) is non-negative if all its entries \( p_{ij} \) are larger or equal then 0. We denote a non-negative matrix with \( P \geq 0 \). A transition matrix is by definition a non-negative matrix. A transition matrix having nonzero probability for staying at the same node is a matrix that is compatible with its graph. An MC (and its transition matrix) is irreducible if it is possible to get from any state to any other state. A non-negative irreducible matrix \( P \) having only one maximum eigenvalue \( \lambda_1 \) on its spectral circle is said to be a primitive matrix.

**Theorem 2.** A transition matrix \( P_{n \times n} \) defined on top of a graph \( G \) is irreducible if the underlying graph is connected.

Proof in [16].

**Theorem 3.** A sufficient condition for an irreducible matrix to be primitive is to have one diagonal element which is strictly positive.

Proof in [16].

A corollary of this theorem is that an irreducible and compatible matrix is also a primitive matrix.

**Theorem 4.** [The Perron–Frobenius theorem.] If \( P_{n \times n} \geq 0 \) is a stochastic, irreducible and primitive matrix then:

(i) It has a unique dominant eigenvalue. \( \lambda_1 = 1 \)

(ii) There is only one positive right eigenvector \( \mathbf{v} = 1_n \) and it satisfies \( P 1_n = \lambda_1 1_n \).

(iii) The product sequence \( \lim_{t \to \infty} P^t \) converges to \( 1_n \pi^T \), where \( \pi \) is the left eigenvector of \( P \) associated with \( \lambda_1 \) i.e. \( P^T \pi = \lambda_1 \pi \). It is also known as the stationary probability distribution vector of the MC and its entries sum up to 1.

Proof of the theorem is given in [16].

The discrete multiplex consensus process is given in compact form by equation (8), that is \( \mathbf{y}(t + 1) = P\mathbf{y}(t) = P^t\mathbf{y}(0) \), where the supratransition matrix has the form (9). According to the Perron–Frobenius theorem, theorem 4, the product sequence \( P^t \) in (8) will converge to \( 1_n \pi^T \), when the matrix \( P \) is row stochastic, irreducible and primitive. If we plug in this expression for the product sequence in (8) we get the consensus value that the process will reach. That is a weighted average of the stationary probability distribution vector and the initial conditions vector \( \mathbf{y}_0 = \pi^T \mathbf{y}(0) \). So the key for reaching a consensus value is to make the matrix \( P \) satisfy all three conditions.

It is easy to prove that the supratransition matrix is a row stochastic matrix, since the class of stochastic matrices forms a convex compact set. The matrices \( D^\alpha \) for \( \alpha = 1, 2, \ldots, L \) are row stochastic and we have that \( c^1 + c^2 + \ldots c^L = 1 \), since \( C \) is also row stochastic. Therefore it follows that for any block of rows in (9):
For the suprtransition matrix to be irreducible the multiplex graph must be connected. Note that for a multiplex graph to be connected, the graph describing the interlayer connections $G_L$ must be connected, but the individual layers $G^a$ do not need to be connected. For example consider two disconnected subnetworks in one layer. In the multiplex graph they can still be connected if at least one pair of nodes from those two subnetworks are connected in another layer. Formally, this means that the graph formed as a union of the graphs in the individual layers $G = \bigcup_{i=1}^L G^i$ needs to be a connected graph. If this graph is disconnected then the multiplex graph will also be disconnected.

According to theorem 3, the suprtransition matrix will be primitive when, in addition to being irreducible, some of its diagonal entries are strictly positive. This will guarantee that the underlying graph is aperiodic. From observing the form of the suprtransition matrix $P$ given in (9) we note that, for some $p_{ii}^a$ to be positive, it is necessary that some $c_{aa}^a$ are positive. In other words the interlayer transition matrix $C$ must also be primitive. Also for those $c_{aa}^a$ that are positive the interlayer transition matrix $D^a$ must also be primitive, so that some of its entries are $d_{ii}^a$ are positive. We note that there are even more general conditions on when a matrix can be primitive than the one given in theorem 3, but they are more involved and will not be discussed here.

Appendix B. Proof of a relation between the eigenvalues of $C$ and $P$

We give a proof to a useful relation between the eigenvalues of the inter layer matrix and the supr transition matrix. Let $C = [c_{ij}]$ be $L \times L$ matrix; let $\lambda$ be an eigenvalue of $C$ and let $v = [v_1, \ldots, v_L]^T$ be the corresponding eigenvector. Assume that the $n \times n$ matrices $D^a$, $a = 1, \ldots, L$ have a common eigenvalue $\mu$ and a common eigenvector $u = [u_1, \ldots, u_n]^T$. This assumption is always valid for the eigenvector $1_n$ of stochastic transition matrices. Then the matrix:

$$P = \begin{bmatrix}
c_{11}D^1 & c_{12}D^2 & \ldots & c_{1L}D^L \\
c_{21}D^1 & c_{22}D^2 & \ldots & c_{2L}D^L \\
\vdots \\
c_{L1}D^1 & c_{L2}D^2 & \ldots & c_{LL}D^L
\end{bmatrix}$$
has the eigenvalue $\lambda \mu$ with eigenvector $w = [v_1u, v_2u, ..., v_Lu]^T$. Indeed

$$P_W = \begin{bmatrix}
    c^{11}v_1D^1u + c^{12}v_2D^2u + \ldots + c^{1L}v_LD^Lu \\
    c^{21}v_1D^1u + c^{22}v_2D^2u + \ldots + c^{2L}v_LD^Lu \\
    \vdots \\
    c^{L1}v_1D^1u + c^{L2}v_2D^2u + \ldots + c^{LL}v_LD^Lu
\end{bmatrix}
= \begin{bmatrix}
    (c^{11}v_1 + c^{12}v_2 + \ldots + c^{1L}v_L)\mu u \\
    (c^{21}v_1 + c^{22}v_2 + \ldots + c^{2L}v_L)\mu u \\
    \vdots \\
    (c^{L1}v_1 + c^{L2}v_2 + \ldots + c^{LL}v_L)\mu u
\end{bmatrix}
= \lambda \mu \begin{bmatrix}
    v_1u \\
    v_2u \\
    \vdots \\
    v_Lu
\end{bmatrix}^T = (\lambda \mu)w.

Note that this result holds always when the matrices $D^\alpha$, $\alpha = 1, \ldots, L$ are identical.

References

[1] DeGroot M H 1974 J. Am. Stat. Assoc. 69 118–21
[2] Olfati-Saber R, Fax J A and Murray R M 2007 Proc. IEEE 95 215–33
[3] Ren W and Cao Y 2011 Distributed Coordination of Multi-agent Networks: Emergent problems, Models, and Issues (Berlin: Springer)
[4] Nijmeijer H and Mareels I M 1997 IEEE Trans. Circuits Syst. 1 44 882–90
[5] Arenas A, Díaz-Guilera A, Kurths J, Moreno Y and Zhou C 2008 Phys. Rep. 469 93–153
[6] Newman M 2010 Networks: An Introduction (Oxford: Oxford University Press)
[7] Vicsek T, Czirók A, Ben-Jacob E, Cohen I and Shochet O 1995 Phys. Rev. Lett. 75 1226
[8] Jadabaie A, Lin J and Morse A S 2003 IEEE Trans. Autom. Control 48 988–1001
[9] Olfati-Saber R 2006 IEEE Trans. Autom. Control 51 401–20
[10] Szell M, Lambiotte R and Thurner S 2010 Proc. Natl Acad. Sci. 107 13636–41
[11] Buldyrev S V, Parshani R, Paul G, Stanley H E and Havlin S 2010 Nature 464 1025–8
[12] Morris R G and Barthelemy M 2012 Phys. Rev. Lett. 109 128703
[13] de Domenico M, Solé-Ribalta A, Cozzo E, Kivelä M, Moreno Y, Porter M A, Gómez S and Arenas A 2013 Phys. Rev. X 3 041022
[14] Gómez S, Díaz-Guilera A, Gómez-Gardeñes J, Perez-Vicente C J, Moreno Y and Arenas A 2013 Phys. Rev. Lett. 110 028701
[15] de Domenico M, Solé-Ribalta A, Gómez S and Arenas A 2014 Proc. Natl Acad. Sci. 111 8351–6
[16] Meyer C D 2000 Matrix Analysis and Applied Linear Algebra vol 2 (Philadelphia, PA: SIAM)
[17] Smilkov D and Kocarev L 2011 Phys. Rev. E 84 016104
[18] Lambiotte R, Sinatra R, Delvenne J C, Evans T, Barahona M and Latora V 2011 Phys. Rev. E 84 017102
[19] Radicchi F and Arenas A 2013 Nat. Phys. 9 717–20