

Article

Local Phase Transitions in a Model of Multiplex Networks with Heterogeneous Degrees and Inter-Layer Coupling

Nedim Bayrakdar ¹, Valerio Gemmetto ¹ and Diego Garlaschelli ^{1,2,3,*} ¹ Lorentz Institute for Theoretical Physics, University of Leiden, 2333 CA Leiden, The Netherlands² IMT School of Advanced Studies Lucca, 55100 Lucca, Italy³ INdAM-GNAMPA Istituto Nazionale di Alta Matematica, 00185 Rome, Italy

* Correspondence: garlaschelli@lorentz.leidenuniv.nl

Abstract: Multilayer networks represent multiple types of connections between the same set of nodes. Clearly, a multilayer description of a system adds value only if the multiplex does not merely consist of independent layers. In real-world multiplexes, it is expected that the observed inter-layer overlap may result partly from spurious correlations arising from the heterogeneity of nodes, and partly from true inter-layer dependencies. It is therefore important to consider rigorous ways to disentangle these two effects. In this paper, we introduce an unbiased maximum entropy model of multiplexes with controllable intra-layer node degrees and controllable inter-layer overlap. The model can be mapped to a generalized Ising model, where the combination of node heterogeneity and inter-layer coupling leads to the possibility of local phase transitions. In particular, we find that node heterogeneity favors the splitting of critical points characterizing different pairs of nodes, leading to link-specific phase transitions that may, in turn, increase the overlap. By quantifying how the overlap can be increased by increasing either the intra-layer node heterogeneity (spurious correlation) or the strength of the inter-layer coupling (true correlation), the model allows us to disentangle the two effects. As an application, we show that the empirical overlap observed in the International Trade Multiplex genuinely requires a nonzero inter-layer coupling in its modeling, as it is not merely a spurious result of the correlation between node degrees across different layers.

Keywords: multiplex networks; maximum entropy models; World Trade Multiplex; mean-field Ising model



Citation: Bayrakdar, N.; Gemmetto, V.; Garlaschelli, D. Local Phase Transitions in a Model of Multiplex Networks with Heterogeneous Degrees and Inter-Layer Coupling. *Entropy* **2023**, *25*, 828. <https://doi.org/10.3390/e25050828>

Academic Editor: Panos Argyrakis

Received: 1 March 2023

Revised: 6 May 2023

Accepted: 9 May 2023

Published: 22 May 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The wide variety of different phenomena that occur around us are often the result of systems that emerge and (self-)organize dynamically. These systems consist of a multitude of basic constituents interacting with each other in complicated ways and forming complex patterns. Many of these systems can be represented as networks sustaining various processes. Examples of such systems include social networks, transportation networks, biological networks, financial networks, and technological networks. In particular, social, financial, and economic networks are an important class of systems that, in the wake of recent global crises (such as the 2007–2008 financial crisis, the COVID-19 pandemic, and the ongoing Ukraine crisis), have been attracting attention given the possibility of studying the propagation of shocks among their constituents. Generally, individuals, banks, firms, or countries can be represented as nodes, and the relationships among them can be represented as links [1–3]. Other types of economic and financial networks are obtained as some form of projection from time series data [3–7]. The study of these networks may increase our understanding of a variety of processes that take place through them, such as the spreading of diseases, the diffusion of (mis)information, the stability of financial markets, and the resilience of the economy.

The simplest approach is to map each constituent within a system onto a single node and to map each interaction between pairs of constituents onto a link of a single type,

regardless of the nature of the interaction. In this approach, all the links in a network are treated on an equal footing, making it a single-layer network representation, which might, however, lead to an oversimplification that fails to capture the details of a multirelational system. For instance, production and trade networks are the result of the functioning of global supply chains, involving the exchange of multiple products between firms and countries, which determines nontrivial dependencies between product-specific layers of the network. In order to realistically follow the propagation of shocks in the economy, knowledge of the nature of the links is essential. The inability to properly represent multirelational systems using single-layer networks has led to the introduction of so-called *multilayer networks* [8–12]. Multilayer networks allow us to describe multirelational systems by representing each type of relationship in a separate layer of the network, where each node is present in all layers, and the different types of connections are reported in the corresponding layers. Returning to the example of social networks, the different types of relationships between people, such as kinship, friendship, coworkership, etc., would each be represented by links in a different layer [13], and could be analyzed in their mutual dependencies.

However, in order to assess true dependencies across layers, one should use proper null models. In recent years, there has been an increase in attention towards null models of networks constructed as random graph ensembles [14–20]. A class of such models is the so-called Exponential Random Graph Models (ERGMs) [17–27]. ERGMs are used commonly within the social network analysis community, and have been more recently re-derived within a statistical physics maximum entropy framework [19,20,27]. This has allowed researchers to utilize techniques that are common in statistical physics. In the ERGM framework, one chooses the probability distribution on graphs such that it maximizes the entropy. This maximization is performed while the expected values of certain chosen graph properties are constrained to be equal to desired values.

Real-world multilayer networks have been compared against null ERGMs with independent layers [28,29]. This comparison has highlighted various properties of real multilayer networks that result from the interdependence of layers. Two such properties are the *overlap* and the *multiplexity* [9,28]. The overlap and the multiplexity essentially contain similar information and capture the correlation of a node's connectivity across two or more layers. For example, in a social network, people may communicate with their friends through multiple means of communication, such as talking on the phone, sending emails, or sending instant text messages. In this example, the layer that represents communication through email has a significant overlap with the layer of communication through text messages. A more specific example is a study of the so-called World Trade Multiplex (representing international trade in different commodities among countries [30]), which showed that, despite the fact that each layer of the multiplex is separately well described by a maximum entropy model with given node degrees [31–33], the observed trade overlap across different commodity-specific layers is significantly different from the overlap predicted by a null model with independent layers [28]. This result is not unexpected, since one can imagine that the trade of a certain product between two countries may increase/decrease the possibility of the trade of a *different* product between the same two countries. Other examples of networks displaying a significant overlap are airport networks, on-line social games, collaboration networks, and citation networks [34–36].

An important conclusion that has been reached after comparing real-world multiplexes against null models with independent layers is that a significant part of the observed overlap in many real networks could actually be spuriously created by the correlations among node degrees across different layers, even if the latter are conditionally independent of each other, instead of resulting from genuine inter-layer dependencies [28,29]. Indeed, if node degrees are correlated among layers, then there will be an increased probability of a link between two nodes being present in multiple layers, while the probability of a link occurring in one layer will not necessarily influence the presence of a link occurring in another layer. The measured overlap of the network therefore consists of a part resulting

from ‘spurious’ coupling between the layers and of a part resulting from genuine coupling between the layers. This spurious coupling increases as the density and/or heterogeneity of the degrees of the network increases. Real-world networks are often dense and have strongly heterogeneous degrees; therefore, the assessment of inter-layer coupling in these real-world networks will be severely affected.

The focus of this paper is the introduction of interdependencies between the layers of a multilayer network in the ERGM through the explicit inclusion of the overlap as an extra constraint. This inclusion of the overlap in the ERGM will aid us in understanding which (higher-order) properties of the network structure may be (highly) dependent on the overlap. Additionally, it will help us distinguish between the overlap in the network due to the correlation of single-node properties across layers and the overlap due to a genuine coupling between the layers. Finally, it will allow us to generate null models with the desired amount of spurious overlap and genuine overlap. It turns out that this problem is mathematically identical to solving the Ising model on a complete graph (which is also known as the mean-field Curie–Weiss model) and leads to a phase transition between a ‘multiplexed’ (magnetized) and a ‘non-multiplexed’ (non-magnetized) phase. However, the problem is more general because the locality of the constraints on the degrees of nodes will imply different parameter values, and hence different properties for the phase transitions relative to different pairs of nodes. For instance, it will, in general, not be possible to enforce a ‘zero-field’ spontaneous symmetry breaking condition for all pairs of nodes simultaneously. Therefore, for a given specification of the constraints, different pairs of nodes may realize different symmetry-broken values of their contribution to the overall inter-layer overlap. Crucially, this property arises only from the simultaneous presence of the two constraints (on the global overlap and on the heterogeneous local degrees), and would not be realized in the absence of one of them.

The rest of the paper is organized as follows: In Section 2, we mathematically define quantities and models that are relevant to this paper. This includes the derivation of a benchmark model, where the layers of the multiplex network are independent. In Section 3, we introduce, and solve analytically, our new model, where the layers of the multiplex are interdependent due to the inclusion of the overlap. Section 4 contains a discussion regarding the possible local phase transitions of the model. In Section 5, we explore our model by using various numerical methods. In Section 6, we briefly analyze the World Trade Multiplex, and show that the empirical overlap in this real-world network is not merely the result of the heterogeneity of the network, but requires a nonzero coupling between the layers in its modeling. Finally, we provide some concluding remarks in Section 7, and some technical details in Appendices A and B.

2. Background Theory

This section contains some background notions, definitions, and models.

2.1. Single-Layer Network Definitions

We will limit our discussion to the case of *binary and undirected* networks. A binary undirected network can be defined as a graph that is an ordered pair $G = (V, E)$, where $V = \{v_1, v_2, \dots, v_N\}$ is a set of N vertices or nodes, and E is a set of *unordered* pairs of different vertices called *edges* or *links*. Note that the definition of E depends on the relevant class of relations between the constituents of the system. The vertex $v_i \in V$ will be referred to simply as i throughout the rest of the paper. If $(i, j) \in E$, the vertices i and j are said to be connected, and may be referred to as *neighbors* of each other. The number of links L of the graph is given by the cardinality of E : $L = |E|$.

Matrix Representation

A graph G is represented by its *adjacency matrix* $G = \{g_{ij}\}$. This is an $N \times N$ matrix where

$$g_{ij} = \begin{cases} 1 & \text{if } (i, j) \in E, \\ 0 & \text{otherwise.} \end{cases} \quad (1)$$

We define E as containing pairs of *distinct* vertices, which means that a vertex cannot have a connection to itself (self-loop). It is then natural to define the diagonal elements as $g_{ii} \equiv 0$. Since we limit our discussion to undirected graphs, the adjacency matrix is always symmetric, $g_{ij} = g_{ji}$, and it therefore contains $N(N - 1)/2$ independent elements that fully specify the matrix and ultimately the graph.

Degrees and Degree Distribution

One of the main topics in the analysis of complex networks is the identification of the different roles that nodes play [37]. For instance, there are a variety of measures that characterize the structural importance of a node in a network. The degree $k_i(G)$ of the graph G is defined as the number of connections node i has to other nodes in the network.

$$k_i(G) = \sum_{j=1}^N g_{ij} \quad (2)$$

The list $\{k_i(G)\}_{i=1}^N$ of degrees is called the *degree sequence* of the graph G . The degree distribution $P(k)$ is defined as the fraction of nodes in the network with degree k . Real-world networks systematically show a degree distribution with heavy tails, where the degrees vary over a broad range, often spanning several orders of magnitude [38,39]. The majority of the vertices of these real-world networks have a small number of links to other vertices, while a few vertices have a relatively high number of links to other vertices, which are also referred to as ‘hubs’. An example is the World Wide Web, where some pages are incredibly popular and are pointed to by thousands of other pages, while generally, most pages are almost unknown. The heavy tails of real-world degree distributions can often be, but not necessarily, approximated by power laws of the form $P(k) \sim k^{-\gamma}$. In any case, vertices with a degree much larger than the average degree $\langle k \rangle$ occur with a non-negligible probability. This is a signature of a high level of statistical heterogeneity in real-world networks. Encoding this heterogeneity will be a crucial ingredient of our models.

2.2. Multiplex Network Definitions

A binary undirected multiplex network can be defined in terms of the previously defined single-layer networks. A multiplex network is a set $\vec{G} = \{G^\alpha\}_{\alpha=1}^M$ of M undirected binary graphs $G^\alpha = (V, E^\alpha)$ that share the same set of N nodes. In the context of multilayer networks, G^α is called a layer of \mathcal{M} , and will be referred to simply as α throughout the rest of the paper. Note that a multiplex network is a type of multilayer network that does not allow inter-layer connections between two layers α and β where $\alpha \neq \beta$.

Matrix Representation

The layer G^α and its intra-layer links can then be represented by the adjacency matrix $G^\alpha = \{g_{ij}^\alpha\}$. This is an $N \times N$ matrix where

$$g_{ij}^\alpha = \begin{cases} 1 & \text{if } (i, j) \in E^\alpha, \\ 0 & \text{otherwise.} \end{cases} \quad (3)$$

Multilinks in Multiplex Networks

In order to capture the information regarding the presence of the links between the pair of nodes (i, j) in any of the M layers, we define the object

$$m_{ij} \equiv (g_{ij}^1, g_{ij}^2, \dots, g_{ij}^M) \tag{4}$$

which is also known as the *multilink* of (i, j) . Additionally, we define the set \mathcal{M}_{ij} as the set that contains all the 2^M possible configurations of m_{ij} .

Multidegrees

The multidegree of a node $i \in V$ of a multiplex network \vec{G} is the object

$$\vec{k}_i(\vec{G}) \equiv (k_i^1(\vec{G}), k_i^2(\vec{G}), \dots, k_i^M(\vec{G})) \tag{5}$$

where

$$k_i^\alpha(\vec{G}) = \sum_{j \neq i}^N g_{ij}^\alpha \tag{6}$$

is the degree of the node i in the layer α [9,40]. From the vector definition of the multidegree, one can obtain a scalar quantity defined as the *layer-averaged degree*:

$$\bar{k}_i(\vec{G}) = \frac{1}{M} \sum_{\alpha=1}^M k_i^\alpha(\vec{G}), \tag{7}$$

which is the degree of node i averaged over all the M layers. Note that, in each layer α , the total layer-specific degree of all nodes equals twice the number of links in that layer, which we denote as L^α :

$$\sum_{i=1}^N k_i^\alpha(\vec{G}) = \sum_{i < j} g_{ij}^\alpha = 2L^\alpha(\vec{G}). \tag{8}$$

Summing the above relationship for the M layers, we get

$$M \sum_{i=1}^N \bar{k}_i(\vec{G}) = \sum_{\alpha=1}^M \sum_{i < j} g_{ij}^\alpha = 2 \sum_{\alpha=1}^M L^\alpha(\vec{G}) = 2L(\vec{G}), \tag{9}$$

where $L(\vec{G})$ denotes the total number of links over the entire multiplex:

$$L(\vec{G}) = \sum_{\alpha=1}^M \sum_{i < j} g_{ij}^\alpha. \tag{10}$$

Overlap

There are many properties that encode the interdependence between the layers of a multilayer network, but we will limit our discussion to one such property: the overlap. The overlap $O^{\alpha\beta}(\vec{G})$ between two layers α and β of the multiplex \vec{G} is defined as the number of links that appear in both layers α and β [34,41]:

$$O^{\alpha\beta}(\vec{G}) = \sum_{i < j} g_{ij}^\alpha g_{ij}^\beta \tag{11}$$

where, throughout the paper, using $\sum_{a < b}$ and $\prod_{a < b}$, we denote a *double sum* and a *double product* for all possible (unrepeated) pairs of values of the two indices, a and b (with $a \neq b$), respectively. The *global overlap* $O(\vec{G})$ is defined as the sum of $O^{\alpha\beta}(\vec{G})$ for all pairs of layers:

$$O(\vec{G}) = \sum_{\alpha < \beta} \sum_{i < j} g_{ij}^\alpha g_{ij}^\beta. \tag{12}$$

As the names of these properties suggest, they are a measure of how overlapping the layers of the multiplex network are.

2.3. Exponential Random Graph Models for Multiplexes

ERGMs are ensemble models, which means that they are defined as probability distributions over many possible (multiplex) networks. Given the observed (or desired) value $C_i^* \equiv C_i(\vec{G}^*)$ for K graph properties $\{C_i(\vec{G})\}_{i=1}^K$ defined on each possible multiplex \vec{G} (where \vec{G}^* represents a particular, e.g., real-world, multiplex of interest), an ERGM generates a probability distribution $P(\vec{G})$ over multiplex networks that maximizes the entropy, under the constraint that the expected value of $C_i(\vec{G})$ equals C_i^* , for all $i = 1, K$. This method provides us with a general framework for modeling maximally random (maximum entropy) multiplex networks, to be used as null models that can be compared against the empirical multiplex \vec{G}^* to detect higher-order patterns that are irreducible to the K enforced constraints. Maximizing the entropy subject to a set of constraints is also widely used in problems with incomplete information [42,43].

Let \mathcal{G}_N^M be the set of (binary undirected) multiplex networks consisting of N vertices and M layers (note that this set includes single-layer networks for $M = 1$), let $\vec{G} = \{G_1, G_2, \dots, G_M\} \in \mathcal{G}_N^M$ be a multiplex network in that set, and let $P(\vec{G})$ be the sought-for probability of \vec{G} within the ensemble. We want $P(\vec{G})$ to be such that the expectation value of each graph observable $C_i(\vec{G})$ (in the chosen set of K observables) is equal to the corresponding observed or desired value C_i^* . This type of probability distribution is also referred to as a *canonical ensemble*. The ideal probability distribution is the one that maximizes the Gibbs–Shannon entropy

$$S = - \sum_{\vec{G} \in \mathcal{G}_N^M} P(\vec{G}) \ln P(\vec{G}) \tag{13}$$

under the normalization condition

$$\sum_{\vec{G} \in \mathcal{G}_N^M} P(\vec{G}) = 1 \tag{14}$$

and the other K constraints

$$C_i^* = \langle C_i \rangle, \quad i = 1, \dots, K, \tag{15}$$

where

$$\langle C_i \rangle \equiv \sum_{\vec{G} \in \mathcal{G}_N^M} P(\vec{G}) C_i(\vec{G}). \tag{16}$$

The maximization of the entropy is achieved by introducing a global Lagrange multiplier η for the normalization condition and a specific multiplier θ_i for each constraint $\langle C_i \rangle = C_i^*$, $i = 1, \dots, K$. This leads to the parametric solution

$$P(\vec{G}, \vec{\theta}) = \frac{e^{-H(\vec{G}, \vec{\theta})}}{Z(\vec{\theta})} \tag{17}$$

where $H(\vec{G}, \vec{\theta})$ is the graph Hamiltonian

$$H(\vec{G}, \vec{\theta}) \equiv \sum_{i=1}^K \theta_i C_i(\vec{G}) = \vec{\theta} \cdot \vec{C}(\vec{G}) \tag{18}$$

and $Z(\vec{\theta})$ is the partition function determined by the normalization condition

$$Z(\vec{\theta}) \equiv e^{\eta+1} = \sum_{\vec{G} \in \mathcal{G}_N^M} e^{-H(\vec{G}, \vec{\theta})}. \tag{19}$$

The parametric form of $P(\vec{G}, \vec{\theta})$, if inserted back into Equation (13), leads to the explicit expression for the entropy:

$$S(\vec{\theta}) = - \sum_{\vec{G} \in \mathcal{G}_N^M} P(\vec{G}, \vec{\theta}) \ln P(\vec{G}, \vec{\theta}) = \vec{\theta} \cdot \langle \vec{C} \rangle + \ln Z(\vec{\theta}). \tag{20}$$

2.4. Maximum Likelihood Parameter Estimation

Equations (17)–(19) fully define the ERGM, apart from the specification of the parameters $\vec{\theta}$. In principle, by treating these Lagrange multipliers as free parameters, one can study the effects that the specification of certain graph observables $\{C_i\}$ has on other aspects of network structure [27,44–47]. This approach, however, does not allow one to consider ERGMs as null models of a particular real network [17,19]. In the latter case, maximum likelihood parameter estimation leads to the unique (given the choice of constraints) ERGM representing a null model for a particular real (multiplex) network \vec{G}^* , and hence, enforcing Equation (15) exactly, as we briefly recall below. This null model can then be used to detect statistically significant deviations of empirical structural properties of \vec{G}^* from the ensemble.

The log-likelihood of the particular multiplex \vec{G}^* is

$$\mathcal{L}(\vec{G}^*, \vec{\theta}) = \ln P(\vec{G}^*, \vec{\theta}) = - \sum_{i=1}^K \theta_i C_i^* - \ln Z(\vec{\theta}). \tag{21}$$

This function has the following properties [19]:

$$\frac{\partial \mathcal{L}(\vec{G}^*, \vec{\theta})}{\partial \theta_i} = \langle C_i \rangle - C_i^* \tag{22}$$

$$\frac{\partial^2 \mathcal{L}(\vec{G}^*, \vec{\theta})}{\partial \theta_i \partial \theta_j} = -\langle C_i C_j \rangle + \langle C_i \rangle \langle C_j \rangle. \tag{23}$$

Equation (22) means that the stationary points $\vec{\theta} = \vec{\theta}^*$ of \mathcal{L} are precisely those that satisfy the constraints (15), i.e.,

$$\langle C_i \rangle_{\vec{\theta}^*} = \sum_{\vec{G} \in \mathcal{G}_N^M} C_i(\vec{G}) P(\vec{G}, \vec{\theta}^*) = \sum_{\vec{G} \in \mathcal{G}_N^M} C_i(\vec{G}) \frac{e^{-\sum_{j=1}^K \theta_j^* C_j(\vec{G})}}{Z(\vec{\theta}^*)} = C_i(\vec{G}^*), \quad i = 1, \dots, K \tag{24}$$

where $\langle C_i \rangle_{\vec{\theta}^*}$ indicates that the ensemble average is evaluated at the values $\vec{\theta}^*$. Equation (23) indicates that \mathcal{L} is concave, since the matrix with entries $\partial^2 \mathcal{L} / \partial \theta_i \partial \theta_j$ has the form of a negative covariance matrix, and must therefore be non-positive definite [48]. The solutions $\vec{\theta}^*$ of the coupled equations $\langle C_i \rangle_{\vec{\theta}^*} = C_i^*$ in Equation (15) can therefore be found by maximizing the log-likelihood \mathcal{L} . If $\partial^2 \mathcal{L} / \partial \theta_i \partial \theta_j$ is negative definite, which will be true if the functions $C_i(\vec{G})$ are linearly independent [48] (i.e., the chosen constraints are non-redundant), then there will be, at most, one solution, and it will be the unique maximum of \mathcal{L} . Maximizing a concave function is generally easier than solving the system of coupled nonlinear equations in Equation (24). Once the solution $\vec{\theta} = \vec{\theta}^*$ is found, it can be used to generate a null model

of \vec{G}^* . Moreover, inserting the value $\vec{\theta}^*$ back into Equation (21) and using Equation (20), we obtain the important relation

$$\begin{aligned} \mathcal{L}(\vec{G}^*, \vec{\theta}^*) &= \ln P(\vec{G}^*, \vec{\theta}^*) \\ &= -\sum_{i=1}^K \theta_i^* C_i^* - \ln Z(\vec{\theta}^*) \\ &= -\sum_{i=1}^K \theta_i^* \langle C_i \rangle_{\vec{\theta}^*} - \ln Z(\vec{\theta}^*) \\ &= -S(\vec{\theta}^*), \end{aligned} \tag{25}$$

i.e., the maximized log-likelihood equals minus the entropy for the particular value $\vec{\theta}^*$, which in turn represents the ‘entropy of the data’ given the chosen constraints. This result allows one to easily calculate the entropy of the data $S(\vec{\theta}^*) = -\mathcal{L}(\vec{G}^*, \vec{\theta}^*)$ automatically as part of the likelihood maximization procedure, rather than as a much more complicated formal sum of all configurations, as in the general definition (13).

2.5. Benchmark: Independent Layers Model

As anticipated in the Introduction, our goal is that of considering how the empirical overlap between links in different layers of a multiplex is jointly determined by both a ‘genuine’ coupling between the M layers and a ‘spurious’ correlation resulting from the heterogeneous (and correlated across layers) degrees of the N nodes. As a null benchmark before inserting both components in an ERGM of a multiplex, we first consider only the layer-averaged degrees of all vertices as constraints, as defined in Equation (7). We can therefore create a null model of a real multiplex \vec{G}^* using the ERGM in combination with the maximum likelihood method. This model will be referred to as the *Average Configuration Model (ACM)*, and will allow us to study the sole effects of correlated heterogeneous degrees on the inter-layer overlap. The Hamiltonian of this model, denoted as H_0 , since it represents a benchmark for a more complicated model to be defined later, is

$$H_0(\vec{G}, \vec{\theta}) = M \sum_{i=1}^N \theta_i \bar{k}_i(\vec{G}) = \sum_{\alpha=1}^M \sum_{i<j} (\theta_i + \theta_j) g_{ij}^\alpha \tag{26}$$

where we have reparametrized by exposing M for convenience. The partition function is

$$\begin{aligned} Z_0(\vec{\theta}) &= \sum_{\vec{G} \in \mathcal{G}_N^M} e^{-\sum_{\alpha=1}^M \sum_{i<j} (\theta_i + \theta_j) g_{ij}^\alpha} \\ &= \sum_{\vec{G} \in \mathcal{G}_N^M} \prod_{\alpha=1}^M \prod_{i<j} e^{-(\theta_i + \theta_j) g_{ij}^\alpha} \\ &= \prod_{\alpha=1}^M \prod_{i<j} \sum_{g_{ij}^\alpha=0}^1 e^{-(\theta_i + \theta_j) g_{ij}^\alpha} \\ &= \prod_{\alpha=1}^M \prod_{i<j} [1 + e^{-(\theta_i + \theta_j)}] \\ &= \prod_{i<j} [1 + e^{-(\theta_i + \theta_j)}]^M. \end{aligned} \tag{27}$$

The probability distribution over the ensemble is then given by

$$P_0(\vec{G}, \vec{\theta}) = \prod_{\alpha=1}^M \prod_{i<j} \frac{e^{-(\theta_i + \theta_j) g_{ij}^\alpha}}{1 + e^{-(\theta_i + \theta_j)}}, \tag{28}$$

from which we see that pairs of nodes and pairs of layers are all independent of each other, each entry g_{ij}^α being an independent Bernoulli random variable with success probability $p_{ij}^\alpha(\vec{\theta})$ and expected value $\langle g_{ij}^\alpha \rangle_{\vec{\theta}}$ given by

$$p_{ij}^\alpha(\vec{\theta}) = \langle g_{ij}^\alpha \rangle_{\vec{\theta}} = \frac{e^{-(\theta_i + \theta_j)}}{1 + e^{-(\theta_i + \theta_j)}} \equiv p_{ij}(\vec{\theta}). \tag{29}$$

Clearly, $p_{ij}^\alpha(\vec{\theta}) = p_{ij}(\vec{\theta})$ is the probability that a link occurs between node i and j in layer α , which turns out to be independent of α given our choice of the layer-averaged (not layer-specific) degree as a constraint.

The log-likelihood of the multiplex \vec{G}^* is

$$\mathcal{L}_0(\vec{G}^*, \vec{\theta}) = -M \sum_{i=1}^N \theta_i \bar{k}_i^* - M \sum_{i < j} \ln [1 + e^{-(\theta_i + \theta_j)}], \tag{30}$$

where $\bar{k}_i^* = \bar{k}_i(\vec{G}^*)$. The parameter value θ_m^* maximizing the log-likelihood must satisfy

$$\left. \frac{\partial \mathcal{L}_0(\vec{G}^*, \vec{\theta})}{\partial \theta_m} \right|_{\vec{\theta} = \vec{\theta}^*} = -M \bar{k}_m^* + M \sum_{j \neq m} \frac{e^{-(\theta_m^* + \theta_j^*)}}{1 + e^{-(\theta_m^* + \theta_j^*)}} = 0 \quad \forall m \tag{31}$$

or equivalently,

$$\bar{k}_i^* = \sum_{j \neq i} \frac{e^{-(\theta_i^* + \theta_j^*)}}{1 + e^{-(\theta_i^* + \theta_j^*)}} \quad \forall i. \tag{32}$$

The above results show that, as expected from the general result reported in Equation (24), according to the maximum likelihood principle, the empirical layer-averaged degree $\bar{k}_i^* = \bar{k}_i(\vec{G}^*)$ of the real multiplex \vec{G}^* is equal to the ensemble average $\langle \bar{k}_i \rangle_{\vec{\theta}^*}$:

$$\begin{aligned} \bar{k}_i^* &= \sum_{j \neq i} p_{ij}(\vec{\theta}^*) \\ &= \frac{1}{M} \sum_{\alpha=1}^M \sum_{j \neq i} p_{ij}^\alpha(\vec{\theta}^*) \\ &= \frac{1}{M} \sum_{\alpha=1}^M \sum_{j \neq i} \langle g_{ij}^\alpha \rangle_{\vec{\theta}^*} \\ &= \langle \bar{k}_i \rangle_{\vec{\theta}^*}. \end{aligned} \tag{33}$$

The probability distribution $P_0(\vec{G}, \vec{\theta}^*)$ can then be written as a product of the layers:

$$P_0(\vec{G}, \vec{\theta}^*) = \prod_{\alpha=1}^M P_0^\alpha(G^\alpha, \vec{\theta}^*) \tag{34}$$

where P_0^α is the probability distribution over a single layer, i.e.,

$$P_0^\alpha(G^\alpha, \vec{\theta}^*) = \prod_{i < j} [p_{ij}(\vec{\theta}^*)]^{g_{ij}^\alpha} [1 - p_{ij}(\vec{\theta}^*)]^{1 - g_{ij}^\alpha}. \tag{35}$$

This means that each layer α can be generated by using the link probability $p_{ij}(\vec{\theta}^*)$ that is equal throughout the layers. This is again a consequence of exclusively constraining properties defined as the overall averages of the layers. This null model can be used as a benchmark to determine the expected value of the inter-layer overlap $O(\vec{G})$ defined in Equation (12), which is due solely to the correlation between the degree of the same node i across the M layers, and not to any genuine inter-layer dependency. This expected value is

$$\langle O \rangle_{\vec{\theta}^*} = \sum_{\alpha < \beta} \sum_{i < j} \langle g_{ij}^\alpha g_{ij}^\beta \rangle_{\vec{\theta}^*} = \sum_{\alpha < \beta} \sum_{i < j} \langle g_{ij}^\alpha \rangle_{\vec{\theta}^*} \langle g_{ij}^\beta \rangle_{\vec{\theta}^*} = \sum_{\alpha < \beta} \sum_{i < j} p_{ij}^2(\vec{\theta}^*), \tag{36}$$

where we have used the independence $\langle g_{ij}^\alpha g_{ij}^\beta \rangle_{\vec{\theta}^*} = \langle g_{ij}^\alpha \rangle_{\vec{\theta}^*} \langle g_{ij}^\beta \rangle_{\vec{\theta}^*}$ between layers $\alpha \neq \beta$. Deliberately, we have chosen the layer-averaged degree as the only constraint so that the expected degree of a node is the same across all layers, thereby creating a strong correlation between degrees in different layers, while keeping the layers themselves independent. Using Equations (25) and (30), we can calculate the entropy of the data, given the model, as

$$S_0(\vec{\theta}^*) = -\mathcal{L}_0(\vec{\theta}^*) = -\ln P_0(\vec{G}^*, \vec{\theta}^*) = M \sum_{i=1}^N \theta_i^* \bar{k}_i^* + M \sum_{i < j} \ln \left[1 + e^{-(\theta_i^* + \theta_j^*)} \right], \tag{37}$$

which only requires the knowledge of $\vec{\theta}^*$ and of the layer-averaged degrees $\bar{k}_i(\vec{G}^*), i = 1, N$.

3. The Overlapping Average Configuration Model

Having illustrated all the ingredients that are necessary to define and model basic properties of multiplex networks within a maximum entropy framework, in this section, we introduce a model of multiplex networks with genuinely interdependent layers. To this end, we incorporate the overlap as an extra constraint in the ERGM, and study the model in combination with the maximum likelihood method. This model is a generalization of the previous ACM benchmark, and will therefore be referred to as the *Overlapping Average Configuration Model (OACM)*, as it includes not only the intra-layer degrees, but also the inter-layer coupling, as building blocks.

3.1. Constructing the Hamiltonian

We want to define a model of a multiplex with M layers, N vertices, and given expected layer-averaged degrees (as defined in Equation (7)) and global inter-layer overlap (as defined in Equation (12)). The Hamiltonian of our ERGM is, in this case,

$$H(\vec{G}, \vec{\theta}, J) = M \sum_{i=1}^N \theta_i \bar{k}_i(\vec{G}) - \frac{4J}{M} O(\vec{G}) = \sum_{i < j} \sum_{\alpha=1}^M (\theta_i + \theta_j) g_{ij}^\alpha - \frac{4J}{M} \sum_{i < j} \sum_{\alpha < \beta} g_{ij}^\alpha g_{ij}^\beta \tag{38}$$

where $(\vec{\theta}, J)$ are the Lagrange multipliers coupled to the $N + 1$ constraints. We have defined the Lagrange multiplier for the overlap as $-4J/M$ for later convenience. Clearly, $H(\vec{G}, \vec{\theta}, J) = H_0(\vec{G}, \vec{\theta})$ where H_0 is the benchmark Hamiltonian of the ACM without overlap defined in Equation (26). Using the multilink m_{ij} defined in Equation (4) and defining

$$\theta_{ij} \equiv \theta_i + \theta_j, \tag{39}$$

the Hamiltonian in Equation (38), this can be written as a sum of the pairs of vertices:

$$H(\vec{G}, \vec{\theta}, J) = \sum_{i < j} h_{ij}(m_{ij}, \theta_{ij}, J) \tag{40}$$

where

$$h_{ij}(m_{ij}, \theta_{ij}, J) \equiv (\theta_i + \theta_j) \sum_{\alpha=1}^M g_{ij}^\alpha - \frac{4J}{M} \sum_{\alpha < \beta} g_{ij}^\alpha g_{ij}^\beta \tag{41}$$

will be referred to as the *pair Hamiltonian*. As we shall see in a moment, the pair Hamiltonian can be mapped exactly to a mean-field Ising model coupling the M layers homogeneously. To arrive at this mapping, we transform the Boolean variables $g_{ij}^\alpha \in \{0, 1\}$ to new ‘spin’ variables $\sigma_{ij}^\alpha \in \{-1, 1\}$, as follows:

$$g_{ij}^\alpha = \frac{1}{2}(\sigma_{ij}^\alpha + 1). \tag{42}$$

From now on, we assume that M is large (multiplex with several layers) and expand expressions accordingly. By defining

$$s_{ij} \equiv \{\sigma_{ij}^1, \sigma_{ij}^2, \dots, \sigma_{ij}^M\} \tag{43}$$

as the multilink for the node pair (i, j) in terms of the $\sigma_{ij}^\alpha = \pm 1$ variables, we see that Equation (42) can be used to transform Equation (41) into

$$h_{ij}(s_{ij}, \theta_{ij}, J) = \left(\frac{\theta_{ij}}{2} - J\right) \sum_{\alpha=1}^M \sigma_{ij}^\alpha - \frac{J}{M} \sum_{\alpha < \beta} \sigma_{ij}^\alpha \sigma_{ij}^\beta - \frac{JM}{2} + \frac{M\theta_{ij}}{2}. \tag{44}$$

If we define

$$B_{ij} \equiv J - \frac{\theta_{ij}}{2}, \tag{45}$$

$$v_{ij} \equiv -MB_{ij} + \frac{JM}{2}, \tag{46}$$

then the pair Hamiltonian finally reduces to

$$h_{ij}(s_{ij}, B_{ij}, J) = -B_{ij} \sum_{\alpha=1}^M \sigma_{ij}^\alpha - \frac{J}{M} \sum_{\alpha < \beta} \sigma_{ij}^\alpha \sigma_{ij}^\beta + v_{ij}. \tag{47}$$

From the above expression, we see that, for every specific pair of nodes (i, j) , the variables σ_{ij}^α can be thought of as Ising spins residing in the M nodes of a fully connected graph, where every Ising spin interacts with every other $M - 1$ spins and is coupled to a ‘field’ B_{ij} . In terms of the multiplex networks being modeled, this means that for every specific pair of nodes (i, j) , the edges connecting i and j throughout the M layers are all coupled to a common ‘external’ field B_{ij} , and are also coupled to each other with a homogeneous interaction strength J/M . A positive coupling $J > 0$ favors more overlap (i.e., more alignment between links in different layers), while $J < 0$ disfavors the overlap. The term v_{ij} is an inessential overall shift in energy independent of the spin configuration. This model is identical to the mean-field Ising or Curie–Weiss model. This exact mapping is what we use in Appendix in order to solve the model analytically, and in particular, to show the existence, for each pair of nodes, of a phase transition separating a ‘magnetized’ phase and a ‘non-magnetized’ phase, which here represent a ‘multiplexed’ phase (where links in different layers tend to ‘align’ to each other) and a ‘non-multiplexed’ phase, respectively.

The full Hamiltonian (40) is a summation of the Hamiltonians of *non-interacting* Ising systems, each for a distinct pairs of nodes. Note, however, that despite the independence of different pairs of nodes, the pair Hamiltonians $h_{ij}(s_{ij}, B_{ij}, J)$ share some parameters: J is common to all such Hamiltonians, and $h_{ij}(s_{ij}, B_{ij}, J)$ and (say) $h_{ik}(s_{ik}, B_{ik}, J)$ also share the parameter θ_i , because the latter appears in both B_{ij} and B_{ik} . This is the result of the original constraint on the degree of each node, which results in the same Lagrange multiplier θ_i appearing in all pair Hamiltonians involving the same node i . These common parameters imply that, even if all pairs of nodes are independent, the control parameters of all pair Hamiltonians cannot be chosen independently, resulting in a correlated phenomenology for the various pairs of nodes. In particular, as we shall see, each pair of nodes can undergo *locally* the typical phase transition of the mean-field Ising model, but the features of these pair-specific phase transitions are all nontrivially related to each other.

We also note, from Equations (44) and (47), that if $J = \theta_{ij}/2$ (or equivalently, $B_{ij} = 0$), then the pair Hamiltonian (hence the graph probability) becomes invariant upon a global ‘spin flip’ ($\sigma_{ij}^\alpha \rightarrow -\sigma_{ij}^\alpha \forall \alpha$), which here corresponds to the replacement of each existing link with a missing link ($g_{ij}^\alpha = 1 \rightarrow g_{ij}^\alpha = 0 \forall \alpha$) and, vice versa, of each missing link with an existing link ($g_{ij}^\alpha = 0 \rightarrow g_{ij}^\alpha = 1 \forall \alpha$). This is due to the vanishing of the ‘external field’ B_{ij} that, when present, selects a preferred ‘spin direction’ (up versus down), which here means a preferred density (high versus low). We expect that with the parameter

choice $J = \theta_{ij}/2$, the pair of nodes (i, j) gains an expected $1/2$ density of links across the M layers, i.e., an expected number of links equal to $M/2$, corresponding to half the maximum number of links for that node pair. Additionally, if J is smaller than the critical value, this expected number of links is also the typical value, and basically, the model is not fundamentally different from a model without constraints, where the intermediate density is produced as a result of a completely uniform probability distribution for the multilink. However, if J exceeds the critical value, the intermediate average density is no longer the typical one realized by individual graphs sampled from the model: rather, it is the ensemble average of two typical (high and low) values of the realized density, just like in the equivalent spin system, below the Curie temperature, and without an external field one would typically observe, with the same probability, overall positive and negative magnetization with a zero ensemble average. The numerical simulations access the typical realized values, while the equations still govern the expected value. This situation corresponds to a ‘symmetry-broken’ phase, where the typical realizations are less symmetric than the Hamiltonian that generates them. However, here, the heterogeneity of the degrees implies different values of the external field $B_{ij} = J - \theta_{ij}/2$, which means that the zero-field spontaneous symmetry breaking condition cannot, in general, be realized for all pairs of nodes simultaneously, leading to a phenomenology governed by the interplay between the values of J and $\{\theta_i\}_{i=1}^N$, and ultimately between the values of the inter-layer overlap and the node degrees.

3.2. Calculating the Partition Function

The partition function defined in (19) can be written as the product

$$Z(\vec{\theta}, J) = \sum_{\vec{G} \in \mathcal{G}_N^M} e^{-H(\vec{G}, \vec{\theta}, J)} = \sum_{\vec{G} \in \mathcal{G}_N^M} \prod_{i < j} e^{-h_{ij}(s_{ij}, \theta_{ij}, J)} = \prod_{i < j} z_{ij}(\theta_{ij}, J), \tag{48}$$

where $z_{ij}(\theta_{ij}, J)$ is the *pair partition function*, which is a sum of the set \mathcal{S}_{ij} of all 2^M possible multilinks for (i, j) :

$$z_{ij}(\theta_{ij}, J) \equiv \sum_{s_{ij} \in \mathcal{S}_{ij}} e^{-h_{ij}(s_{ij}, \theta_{ij}, J)}. \tag{49}$$

The multiplex probability can be written in terms of the multilink probabilities $P_{ij}(s_{ij}, \theta_{ij}, J)$:

$$P(\vec{G}, \vec{\theta}, J) = \prod_{i < j} P_{ij}(s_{ij}, \theta_{ij}, J) \tag{50}$$

where

$$P_{ij}(s_{ij}, \theta_{ij}, J) \equiv \frac{e^{-h_{ij}(s_{ij}, \theta_{ij}, J)}}{z_{ij}(\theta_{ij}, J)}. \tag{51}$$

The complete partition function and multiplex probability can therefore be obtained as products of pair-specific quantities, where each multilink can be regarded as a configuration of a Curie–Weiss system. To obtain an explicit expression for $z_{ij}(\theta_{ij}, J)$, we use a Hubbard–Stratonovich transformation and the Laplace theorem [49] in the limit $M \rightarrow \infty$. The details are provided in Appendix A and are a generalization of the approach used in [50]. The final result is

$$z_{ij}(\theta_{ij}, J) = 2^M e^{-\frac{M}{2}\theta_{ij} - 2JMu_{ij}(u_{ij}-1)} \cosh^M \left(2Ju_{ij} - \frac{\theta_{ij}}{2} \right), \tag{52}$$

where u_{ij} is the solution to the equation

$$u_{ij} = \frac{1}{2} + \frac{1}{2} \tanh \left(2Ju_{ij} - \frac{\theta_{ij}}{2} \right). \tag{53}$$

The solutions to the above equation will be discussed in the next section.

Now, given a particular real multiplex network \vec{G}^* , the log-likelihood, as defined, in general, in Equation (21), is

$$\mathcal{L}(\vec{\theta}, J) = \ln P(\vec{G}^*, \vec{\theta}, J) = \sum_{i < j} \left[-h_{ij}(s_{ij}^*, \theta_{ij}, J) - \ln z_{ij}(\theta_{ij}, J) \right]. \tag{54}$$

At a stationary point of \mathcal{L} , the derivatives of \mathcal{L} with respect to every Lagrange multiplier must equal zero. As we show in Appendix B, this leads to the maximum likelihood equations

$$\sum_{j \neq i} \sum_{\alpha=1}^M g_{ij}^{*\alpha} = M \sum_{j \neq i} u_{ij}^* \quad \forall i \tag{55}$$

$$\frac{4}{M} \sum_{i < j} \sum_{\alpha < \beta} g_{ij}^{*\alpha} g_{ij}^{*\beta} = 2M \sum_{i < j} (u_{ij}^*)^2 \tag{56}$$

where u_{ij}^* , being the solution to Equation (53) with $(\vec{\theta}, J)$ replaced by $(\vec{\theta}^*, J^*)$, is implicitly related to the maximum likelihood parameters $(\vec{\theta}^*, J^*)$. Note that the quantities on the LHS of Equations (55) and (56) are precisely the quantities that we constrained from the start, namely, $M\bar{k}_i^*$ and $4O^*/M$, respectively. According to the maximum likelihood principle, these empirical quantities must equal their respective ensemble averages, $M\langle \bar{k}_i \rangle_{\theta^*, J^*}$ and $4\langle O \rangle_{\theta^*, J^*} / M$, which appear on the RHS. The quantity u_{ij}^* can therefore be considered as an *average* probability of a link occurring between the nodes i and j , which is equal throughout the M layers and is, therefore, a measure of the density of links in the multilink m_{ij} . This is similar to how we identified p_{ij} to be the connection probability in the ACM, which was based solely on the constraints \bar{k}_i . In support of this idea, we see that, in the case $J^* = 0$, the Lagrange multipliers $\vec{\theta}^*$ reduce it to the value $\vec{\theta}^0 \equiv \vec{\theta}^*|_{J^*=0}$, such that

$$u_{ij}^*|_{J^*=0} = \frac{1}{2} \left[1 + \tanh \left(-\frac{\theta_i^0 + \theta_j^0}{2} \right) \right] = \frac{e^{-(\theta_i^0 + \theta_j^0)}}{1 + e^{-(\theta_i^0 + \theta_j^0)}} = p_{ij}(\vec{\theta}^0) \tag{57}$$

which is identical to the expression in Equation (29), providing the link probability p_{ij} obtained in Section 2.5 in the absence of the constraint for the overlap. The quantity u_{ij}^* can therefore possibly be interpreted as a *mean-field* quantity that *globally* incorporates the layer interdependence that was introduced through the overlap O^* , but *locally* treats the layers as if they were independent. A characteristic of mean-field theories is that the effects of all elements of a system on a given element are approximated by a single, average effect.

Formally, we can calculate the entropy of the data, given the model, as the maximized likelihood using Equations (25) and (54):

$$\begin{aligned} S(\vec{\theta}^*, J^*) &= -\mathcal{L}(\vec{\theta}^*, J^*) \\ &= -\ln P(\vec{G}^*, \vec{\theta}^*, J^*) \\ &= H(\vec{G}^*, \vec{\theta}^*, J^*) + \sum_{i < j} \ln z_{ij}(\theta_{ij}^*, J^*) \\ &= M \sum_{i=1}^N \theta_i^* \bar{k}_i^* - \frac{4J^*}{M} O^* + \sum_{i < j} \ln z_{ij}(\theta_{ij}^*, J^*), \end{aligned} \tag{58}$$

which requires the knowledge of the parameters $\vec{\theta}^*$ and J^* (which are, however, defined only implicitly through u_{ij}^*). Comparing the above expression with Equation (37), we see that $S(\vec{\theta}^0, 0) = S_0(\vec{\theta}^0)$, as expected, i.e., the model with $J^* = 0$ has the same entropy as the equivalent ACM with no overlap, for the same value of $\vec{\theta}^0$. Similarly, $\mathcal{L}(\vec{\theta}^0, 0) = \mathcal{L}_0(\vec{\theta}^0)$ for the maximized likelihood in the two models. In order to understand the relationship between the entropies of the two models when $J^* \neq 0$, let us first note that a positive (resp. negative) coupling strength J^* means that the empirical overlap O^* is larger (resp. smaller) than the expected overlap under the null model with $J^* = 0$, i.e.,

$$O^* \leq \langle O \rangle_{\vec{\theta}^0} \Leftrightarrow J^* \leq 0 \tag{59}$$

where we have used the notation in Equation (36). However, one should not naively conclude from the combination of Equations (58) and (59) that the entropy of the model with $J^* < 0$ is larger than the entropy of the model with $J^* = 0$, because the two partition functions are different, and also because the two entropies are calculated for different Lagrange multipliers, i.e., $\vec{\theta}^* \neq \vec{\theta}^0$ when $J^* \neq 0$. In fact, we can actually show that the entropy of the model with $J^* \neq 0$ is always smaller than the one for the model with $J^* = 0$. To see this, we introduce the relative entropy (or Kullback–Leibler divergence) between the two models, as follows:

$$R(\vec{\theta}^0, \vec{\theta}^*, J^*) \equiv \sum_{\vec{G} \in \mathcal{G}_N^M} P(\vec{G}, \vec{\theta}^*, J^*) \ln \frac{P(\vec{G}, \vec{\theta}^*, J^*)}{P_0(\vec{G}, \vec{\theta}^0)} \geq 0, \tag{60}$$

where the last inequality is a well-known property of the relative entropy, and the equality is realized if, and only if, $P_0(\vec{G}, \vec{\theta}^0)$ and $P(\vec{G}, \vec{\theta}^*, J^*)$ are identical, which, in turn, requires $J^* = 0$, yielding $\vec{\theta}^0 = \vec{\theta}^*$ and $R(\vec{\theta}^0, \vec{\theta}^0, 0) = 0$. For $J^* \neq 0$, we can write

$$\begin{aligned} R(\vec{\theta}^0, \vec{\theta}^*, J^*) &= \sum_{\vec{G} \in \mathcal{G}_N^M} P(\vec{G}, \vec{\theta}^*, J^*) \ln P(\vec{G}, \vec{\theta}^*, J^*) - \sum_{\vec{G} \in \mathcal{G}_N^M} P(\vec{G}, \vec{\theta}^*, J^*) \ln P_0(\vec{G}, \vec{\theta}^0) \\ &= -S(\vec{\theta}^*, J^*) + \sum_{\vec{G} \in \mathcal{G}_N^M} P(\vec{G}, \vec{\theta}^*, J^*) \left[H_0(\vec{G}, \vec{\theta}^0) + \ln Z_0(\vec{\theta}^0) \right] \\ &= -S(\vec{\theta}^*, J^*) + \sum_{\vec{G} \in \mathcal{G}_N^M} P_0(\vec{G}, \vec{\theta}^0) \left[H_0(\vec{G}, \vec{\theta}^0) + \ln Z_0(\vec{\theta}^0) \right] \\ &= -S(\vec{\theta}^*, J^*) + \sum_{\vec{G} \in \mathcal{G}_N^M} P_0(\vec{G}, \vec{\theta}^0) \ln P_0(\vec{G}, \vec{\theta}^0) \\ &= -S(\vec{\theta}^*, J^*) + S_0(\vec{\theta}^0), \end{aligned} \tag{61}$$

where we have used the fact that $H_0(\vec{G}, \vec{\theta}^0) = M \sum_{i=1}^N \theta_i^0 \bar{k}_i(\vec{G})$ has the same expectation value, equal to $M \sum_{i=1}^N \theta_i^0 \bar{k}_i(\vec{G}^*)$, under both $P(\vec{G}, \vec{\theta}^*, J^*)$ and $P_0(\vec{G}, \vec{\theta}^0)$:

$$\begin{aligned} \sum_{\vec{G} \in \mathcal{G}_N^M} P(\vec{G}, \vec{\theta}^*, J^*) H_0(\vec{G}, \vec{\theta}^0) &= M \sum_{i=1}^N \theta_i^0 \left[\sum_{\vec{G} \in \mathcal{G}_N^M} P(\vec{G}, \vec{\theta}^*, J^*) \bar{k}_i(\vec{G}) \right] \\ &= M \sum_{i=1}^N \theta_i^0 \bar{k}_i(\vec{G}^*) \\ &= M \sum_{i=1}^N \theta_i^0 \left[\sum_{\vec{G} \in \mathcal{G}_N^M} P_0(\vec{G}, \vec{\theta}^0) \bar{k}_i(\vec{G}) \right] \\ &= \sum_{\vec{G} \in \mathcal{G}_N^M} P_0(\vec{G}, \vec{\theta}^0) H_0(\vec{G}, \vec{\theta}^0). \end{aligned} \tag{62}$$

Now, applying the inequality $R(\vec{\theta}^0, \vec{\theta}^*, J^*) \geq 0$ in Equation (60) to Equation (62), we get

$$0 \leq S(\vec{\theta}^*, J^*) \leq S_0(\vec{\theta}^0), \tag{63}$$

confirming that the entropy of the model with $J^* \neq 0$ is always smaller than the one for the model with $J^* = 0$, consistent with the fact that the former is more constrained than the latter.

4. Local Phase Transitions in the Model

The number of solutions of Equation (53) depends on the values of the parameters $\theta_{ij} = \theta_i + \theta_j$ and J . We illustrate this fact in Figure 1, where both the LHS and the RHS of

Equation (53) are plotted as a function of u_{ij} for various values of θ_{ij} and J . The appearance of multiple solutions signals the existence of *phase transitions* in the limit when the number M of layers diverges, which determine abrupt changes in the value of u_{ij} and, therefore, also in the properties of the multilink m_{ij} and the structure of the multiplex as a whole. The configurations for m_{ij} that are separated by a phase transition are the *phases* of the multilink. The point where multiple solutions appear or vanish is the *bifurcation point*.

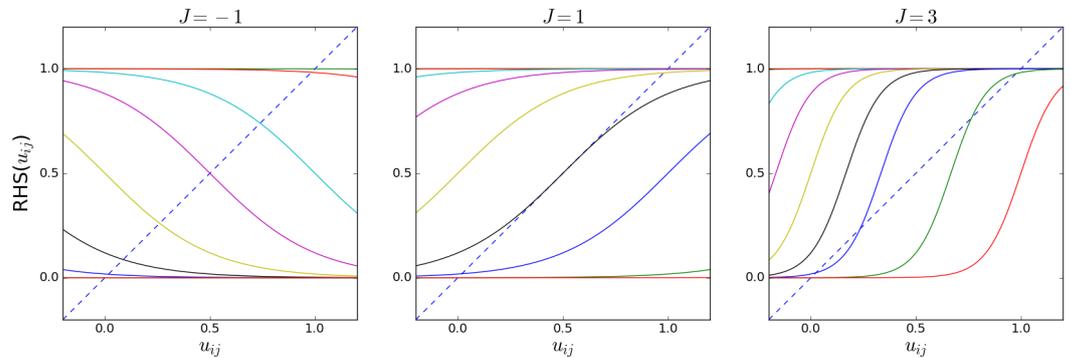


Figure 1. A graphical illustration of the solution(s) of Equation (53). The solid lines show the RHS of Equation (53) as a function of u_{ij} for the different parameters $\theta_{ij} \in \{-12, -8, -4, -2, 0, 2, 4, 8, 12\}$, while the dashed line shows the LHS, which equals u_{ij} itself. For a given parameter value, the solutions of Equation (53) are the intersection between the dashed and the corresponding solid line. Each panel corresponds to a different value of J (in the rest of the paper, we will consider only $J \geq 0$).

Figure 1 shows that, at the interval $0 \leq u_{ij} \leq 1$, there can be either one, two, or three solutions, and that for $\theta_{ij} \rightarrow +\infty$ or $\theta_{ij} \rightarrow -\infty$ there is always one solution, namely, $u_{ij} = 0$ or $u_{ij} = 1$, respectively. The number of solutions depends on whether the slope (derivative) of the RHS (which depends on the parameters) exceeds the slope of the LHS (which is always equal to 1) of Equation (53) at their intersection. From now on, we will consider only the case $J \geq 0$, which corresponds to a tendency to create an increased inter-layer overlap compared with the model with $J = 0$. The case $J < 0$ corresponds to the opposite case where the overlap is suppressed, which we do not discuss here. New solutions appear or vanish at the point where Equation (53) is satisfied *and* the derivatives of the LHS and RHS of Equation (53) are equal:

$$1 = J \left[1 - \tanh^2 \left(2Ju_{ij} - \frac{\theta_{ij}}{2} \right) \right]. \tag{64}$$

Equation (64) cannot be satisfied if $0 \leq J \leq 1$, since $0 \leq \tanh^2(x) < 1$ for $x \in \mathbb{R}$, and, therefore, if $J \leq 1$, a phase transition is impossible, and there is a unique solution for u_{ij} . When $J > 1$, Equation (64) gives us two potential solution branches, $u_{ij}^\pm = \frac{1}{2} \pm \frac{1}{2} \sqrt{1 - 1/J}$, where we have used $2u_{ij} - 1 = \tanh(2Ju_{ij} - \theta_{ij}/2)$. Equation (53) can be written as $\theta_{ij} = 4Ju_{ij} - \ln[u_{ij}/(1 - u_{ij})]$ using the identity $\tanh^{-1} x = \frac{1}{2} \ln[(1 + x)/(1 - x)]$. By then substituting u_{ij}^\pm into this expression for θ_{ij} , we obtain the equations for the two curves in the (J, θ_{ij}) plane that mark the points where additional solutions appear or vanish:

$$\theta_{ij}^+(J) = \frac{2\sqrt{J}}{\sqrt{J} - \sqrt{J-1}} - \ln \left(\frac{\sqrt{J} + \sqrt{J-1}}{\sqrt{J} - \sqrt{J-1}} \right), \tag{65}$$

$$\theta_{ij}^-(J) = \frac{2\sqrt{J}}{\sqrt{J} + \sqrt{J-1}} - \ln \left(\frac{\sqrt{J} - \sqrt{J-1}}{\sqrt{J} + \sqrt{J-1}} \right), \tag{66}$$

as shown in Figure 2. In the region between the two curves, there are three solutions to Equation (53). Note that the ‘zero-field’ condition $\theta_{ij} = 2J$ is always in that region when $J > 1$. This means that the condition $J > 1$ is sufficient to ensure that the system is in the magnetized (symmetry-broken) phase when in the absence of the external field. However,

when $\theta_{ij} \neq 2J$, the condition $J > 1$ is necessary but not sufficient. In particular, generally, it may happen that, for a given value of $J > 1$, different pairs of nodes will be in different (magnetized or non-magnetized) phases depending on the value of θ_{ij} . This shows that the system can undergo a multitude of separate phase transitions if the parameters $\{\theta_{ij}\}$ remain fixed and J is varied.

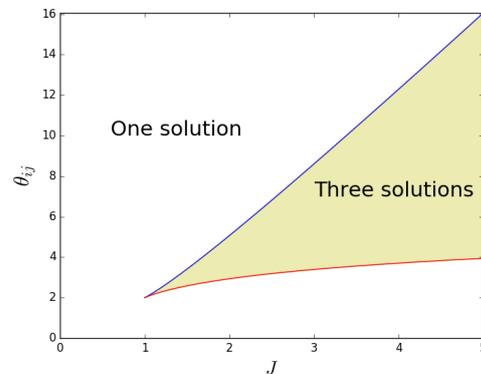


Figure 2. The upper (blue) and lower (red) curves correspond to Equations (65) and (66), respectively, which delimit the region of phase space (yellow area), for which Equation (53) has three solutions. Note that the ‘zero-field’ condition $\theta_{ij} = 2J$ is always in the yellow area when $J > 1$, so the condition $J > 1$ is sufficient to ensure that the system in zero field is in the magnetized (symmetry-broken) phase.

In the magnetized phase, the phenomenon of symmetry breaking will occur: the typical realized values of the ‘magnetization’ will not coincide with the corresponding ensemble average. In the zero-field case ($\theta_{ij} = 2J$), the symmetry breaking is ‘spontaneous’, i.e., not induced by any field pointing in a preferred direction, while in the nonzero-field case, the symmetry is broken by the field itself. This well-known property of the Ising model has specific implications for our problem here. Indeed, while certain values of θ_{ij} , J may solve the maximum likelihood Equations (55) and (56), the corresponding solutions to Equation (53) may not necessarily maximize the likelihood, and are therefore not ‘valid’ (or *stable*). Once the values θ_{ij}^* and J^* that solve the maximum likelihood equations are found, the graph probability corresponding to this set of values can be written as a function of the configuration of the graph (or the collection of configurations of the multilinks m_{ij}), and one can check which typical configurations (those minimizing the Hamiltonian) arise. As Figure 1 suggests, in the regime where there are three solutions, u_{ij} , one value will be relatively high (which corresponds to a relatively high density of links in m_{ij}), another value will be relatively low (which corresponds to a relatively low density of links in m_{ij}), and the third value will be between the other two, corresponding to an intermediate density of links in m_{ij} . By inspecting the (pair) Hamiltonian in Equation (47) in terms of the $\sigma_{ij}^\alpha = 2g_{ij}^\alpha - 1$ variable, it becomes clear which of the three solutions u_{ij}^* are viable (stable). In the case where $B_{ij} = 0$, or equivalently, when $\theta_{ij} = 2J$, the (pair) Hamiltonian is symmetric with respect to a change in sign, $\sigma_{ij}^\alpha \rightarrow -\sigma_{ij}^\alpha$, which means that the high- and low-density solutions are equal. This is the symmetry-broken situation we have discussed in Section 3.1. In this case, the intermediate-density solution will result in a lower value for the Hamiltonian than the high- and low-density solutions. The viable (stable) solutions are therefore the high- and low-density ones. In the case where $B_{ij} \neq 0$, it is clear that the high-density solution minimizes the Hamiltonian when $B_{ij} > 0$ and maximizes it when $B_{ij} < 0$. The low-density solution minimizes the Hamiltonian when $B_{ij} < 0$ and maximizes it when $B_{ij} > 0$. The intermediate solution will, however, never minimize the Hamiltonian when $B \neq 0$, and is therefore never viable (stable). From these considerations, it becomes clear that a phase transition, corresponding to a sudden change in u_{ij} , may only happen when we cross from a negative (positive) B_{ij} to a positive (negative) B_{ij} (when $J > 1$). Figure 3 shows the symmetric stable solutions u_{ij} in the case where $B_{ij} = 0$, with the bifurcation occurring at $J = 1$. In case of the positive field $B_{ij} = +1$, it shows a single stable solution

curve, which is the high-density solution (in the case where $B_{ij} = -1$, this image would be flipped with respect to the $u_{ij}^* = 1/2$ axis). The right panel in Figure 3 shows that the value of the stable solution u_{ij} jumps when B_{ij} crosses from positive to negative, as expected.

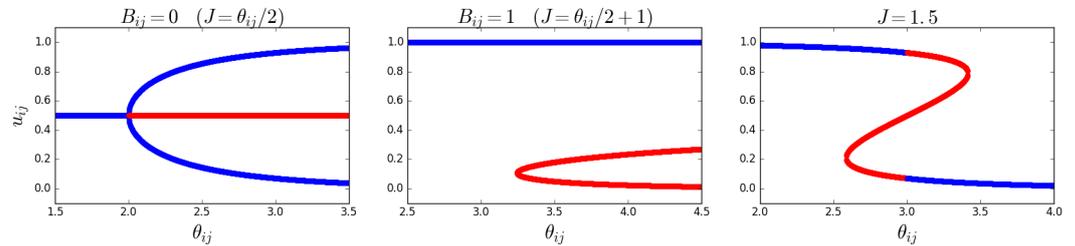


Figure 3. Solutions for u_{ij} as a function of θ_{ij} for different parameter values. The blue and red segments of the curve(s) correspond to the stable and unstable solutions of Equation (53), respectively. **Left panel:** $B_{ij} = 0$ (with J varying accordingly). **Middle panel:** $B_{ij} = 1$ (with J varying accordingly). **Right panel:** constant value of $J = 1.5$, which translates to a non-constant B_{ij} .

Combining the above considerations for all multilinks simultaneously, and adding the other constraint on the layer-averaged degrees, the multiplex will undergo a sequence of phase transitions, determining a hierarchy of increasingly ordered (magnetized, or rather ‘multiplexed’ in this case) phases where, for an increasing number of pairs of nodes, the links in different layers will tend to ‘align’ to each other (for $J > 1$). The separations between these phase transitions will depend on the values of the enforced layer-averaged degrees, which determine $\bar{\theta}^*$. The fully ordered phase, where all the M layers of the multiplex are perfectly aligned, and are, therefore, basically an identical copy of each other. We might say that, in this case, the effective number of independent layers is $M_{\text{eff}} \approx 1$, and the expected overlap is maximal and proportional to the expected number $\langle L \rangle_{\bar{\theta}^*, J^*} = \sum_{\alpha=1}^M \sum_{i < j} u_{ij}^*$ of links in the entire multiplex:

$$\langle O \rangle_{\bar{\theta}^*, J^*} \approx \sum_{\alpha < \beta} \sum_{i < j} u_{ij}^* = (M/2) \langle L \rangle_{\bar{\theta}^*, J^*}, \tag{67}$$

since $\langle g_{ij}^\alpha g_{ij}^\beta \rangle_{\bar{\theta}^*, J^*} \approx \langle g_{ij}^\alpha \rangle_{\bar{\theta}^*, J^*} \langle g_{ij}^\beta \rangle_{\bar{\theta}^*, J^*} = u_{ij}^*$ for most pairs, i.e., α, β , of layers. In the opposite extreme, we have a fully disordered phase where no pair of nodes is multiplexed (for instance, if $J < 1$), so the effective number of independent layers is maximal ($M_{\text{eff}} \approx M$), and the expected overlap is basically of the order of that given by Equation (36) for the model with $J^* = 0$, i.e.,

$$\langle O \rangle_{\bar{\theta}^*, J^*} \approx \sum_{\alpha < \beta} \sum_{i < j} (u_{ij}^*)^2, \tag{68}$$

since $\langle g_{ij}^\alpha g_{ij}^\beta \rangle_{\bar{\theta}^*, J^*} \approx \langle g_{ij}^\alpha \rangle_{\bar{\theta}^*, J^*} \langle g_{ij}^\beta \rangle_{\bar{\theta}^*, J^*} = (u_{ij}^*)^2$ for most pairs of layers. The relationship between $\langle O \rangle_{\bar{\theta}^*, J^*}$ and $\langle L \rangle_{\bar{\theta}^*, J^*}$ will depend on the specific values of $\{u_{ij}^*\}_{i < j}$, so ultimately, on the enforced degree sequence. Between these two extremes, if the phases are well separated (which here means that the enforced degrees of different nodes have very different values), there will be intermediate regimes where $\langle O \rangle_{\bar{\theta}^*, J^*}$ and $\langle L \rangle_{\bar{\theta}^*, J^*}$ scale in a way that is between the two limiting scalings. All these general considerations will be confirmed in the next sections with numerical, analytical, and empirical analyses.

5. Numerical Analysis

Equations (53), (55), and (56) are the key equations of our OACM model. These equations are generally, however, very difficult to solve. Therefore, before creating a null model for a real-world network by solving the maximum likelihood equations to find the Lagrange multipliers, we shall first treat the Lagrange multipliers as free parameters in order to explore and analyze the properties of the model as a function of these parameters. This analysis shall be performed by utilizing the Metropolis–Hastings algorithm [51]. This algorithm can be used to sample the exponential probability distribution defined by

the Hamiltonian of the model. By sampling the distribution, we numerically obtain various properties of the graph ensemble, which may then be compared to our analytical results in order to test the validity of the latter. Note that the sampling of the exponential distribution defined by a specific Hamiltonian may also be regarded as the simulation of a multiplex that corresponds to that Hamiltonian.

5.1. Exploring the Parameter Space

In order to explore the space of parameters, we are primarily interested in the difference between statistically *homogeneous* networks and statistically *heterogeneous* ones. To this end, we will explore the parameter space $(\theta_1, \dots, \theta_N, J)$ of the model by specifying a value for J and sampling certain transformed parameters x_1, \dots, x_N from a distribution for each class, where $x_i \equiv e^{-\theta_i}$. The quantity x_i will be referred to as the ‘fitness’, or ‘hidden variable’, of node i . The broader the distribution of the fitness, the more heterogeneous the resulting network structure.

5.1.1. Homogeneous Fitness: Erdős–Rényi Graphs with Overlap

The simplest distribution from which we can sample x_1, \dots, x_N is the delta distribution centered at x , such that $x_1 = x_2 = \dots = x_N \equiv x$ and, therefore, $\theta_1 = \theta_2 = \dots = \theta_N \equiv \theta = -\ln x$, resulting in statistically homogeneous networks. With this choice of parameters, our model is an extension of the Erdős–Rényi model, which is a random graph model that can be derived within the ERGM by solely constraining the total number of links in the network, and where all links occur with the same probability. As we shall see, the extension derives from the fact that the extra constraint on the overlap can lead to a symmetry-breaking phase transition, although the broken symmetry might not manifest at first sight. Indeed, since the parameters are the same for all pairs of nodes, the condition for the existence of multiple solutions is also the same, and, therefore, there is a unique phase transition where, depending on the values of θ and J , pairs of nodes are either all ‘magnetized’ or all ‘non-magnetized’. Similarly, since here $\theta_{ij} = \theta_i + \theta_j = 2\theta \forall i, j$, the spontaneous symmetry-breaking condition discussed in Section 3.1 for the vanishing of the external field is the same for all pairs of nodes, and given by $J = \theta$. In the symmetry-broken (magnetized) phase, for all pairs of nodes, the expected value of $\sum_{\alpha=1}^M g_{ij}^{\alpha}$ (or equivalently, of the ‘magnetization’ $\sum_{\alpha=1}^M \sigma_{ij}^{\alpha}$) is the same, and is always between the two typical (high-density and low-density) realized values. However, since all pairs are independent, the actual realized values of $\sum_{\alpha=1}^M g_{ij}^{\alpha}$ are also independent across pairs, so on average, over the entire network, the magnetization will realize both the low-density and high-density values, with equal probability. In other words, different pairs of nodes are i.i.d. realizations of the same system. This is a peculiar situation where the realized values of L and O (which represent sums of all pairs of nodes) will still coincide with their expected values as if no symmetry breaking was present, even if different pairs of nodes actually realize different symmetry-broken values that are individually different from the expected value. The net result is an expected number of links $\langle L \rangle = MN(N-1)/4$ equal to half the maximum one, or equivalently, an average zero magnetization in the associated spin system. Similar considerations apply to the case $J \neq \theta$, with the difference that, in that case, the symmetry is not broken spontaneously, but by the direction of the external field (value of θ), which implies that the two typical realized values of the magnetization for a given pair of nodes are no longer symmetric around the expected value. Still, both typical values will be realized, independently and with their probabilities, across the entire network, because different pairs of nodes are still independent. So, irrespective of the value of J and θ , we expect to observe realized values of L and O that correspond again to what one would observe without symmetry breaking, using the ensemble averages for each pair, irrespective of the phase of the system. All these considerations are confirmed below.

By looking at Equation (38), we can see that a uniform θ essentially means that instead of constraining the average layer degrees \bar{k}_i , we constrain the total number of links L in the

multiplex network. In this case, the combined maximum entropy and maximum likelihood equations become

$$u = \frac{1}{2} + \frac{1}{2} \tanh(2J^*u - \theta^*) \tag{69}$$

$$\sum_{i < j} \sum_{\alpha=1}^M g_{ij}^{*\alpha} = \frac{MN(N-1)}{2} u^* = \langle L \rangle_{\theta^*, J^*} \tag{70}$$

$$\frac{4}{M} \sum_{i < j} \sum_{\alpha < \beta} g_{ij}^{*\alpha} g_{ij}^{*\beta} = MN(N-1)(u^*)^2 = \frac{4}{M} \langle O \rangle_{\theta^*, J^*} \tag{71}$$

where $u^* = u(\theta^*, J^*)$ is the solution to Equation (69). Note that we now have a single equation for u , confirming the existence of a single *global* phase transition across the multiplex network, rather than separate local phase transitions for every multilink m_{ij} . Additionally, we note that if u^* can be considered as the density (and the link probability) of the network, then the value of u^* is exactly the same as the value of the density p in the Erdős–Rényi model [14,27], which solely constrains the number of links in the network. The difference between our model and the Erdős–Rényi model is that our model contains the possibility of a phase transition. However, since the number of links $\langle L \rangle$ also determines the overlap $\langle O \rangle$, the two quantities cannot be tuned independently of each other.

By using the Metropolis–Hastings algorithm, we have sampled our ERGM for multiplexes with $M = 100$ layers and $N = 100$ nodes for various values of θ and/or J . If we repeat the simulations for $J = 1.5$ and $\theta = 1.4, \theta = 1.5$, and $\theta = 1.6$, the system must undergo a phase transition as per Figure 3. We expect an abrupt change in the value of u^* , and according to Equations (70) and (71), we therefore expect an abrupt change in the equilibrium value of both L and O . Figure 4 shows simulations for $\theta \in \{1.4, 1.5, 1.6\}$ confirming the transition from a relatively high to a low density as the value of the field $B = J - \theta$ changes sign. These simulations have been repeated for different combinations of values for J and θ around the point where B changes sign, confirming the results shown here. Note that the middle plot in Figure 4 shows that the algorithm converges to multiplexes with a density of $1/2$, confirming that, when $B = 0$, L is approximately half of the total amount of possible links in the multiplex, as we expected above.

In Figure 5 we test the prediction, given by Equations (70) and (71), of the quadratic relationship $\langle O \rangle = \langle L \rangle^2 / N^2$. Note that this quadratic trend is predicted irrespective of the value of $J > 0$, and even coincides with what Equation (68) predicts in the case $J = 0$ for a homogeneous multiplex with constant θ , as considered here. So, in this case, the expected relationship between $\langle O \rangle$ and $\langle L \rangle$ is not informative regarding the phase transition, although the specific values picked up by the system along the curve are. Indeed, we again simulate multiplexes with $M = 100$ layers, $N = 100$ nodes, and a variety of values for θ and J . Each simulation results in a value for $\langle L \rangle$ and a value for $\langle O \rangle$, which we plot against each other. These points are then compared to the theoretical points predicted by Equations (69)–(71) for the chosen parameter values, and added to Figure 5. We see that the relationship between simulated quantities is in agreement with the one predicted by the model. As we had anticipated, this is the result of the fact that different pairs of nodes are i.i.d. realizations of the same system, so that the ensemble average is realized as a sample average of the pairs of nodes across the network, even if in the symmetry-broken phase, the ensemble average of $\sum_{\alpha=1}^M g_{ij}^{\alpha}$ is not representative of any of the values realized locally for individual pairs of nodes. Therefore, the only scaling we observe coincides with the one given in Equation (68) for the ‘non-magnetized’ regime in the case where θ is the same for all nodes. The only, although very important, signature of the phase transition we see in Figure 5 is the fact that, for $J > 1$ and $\theta \neq J$, both the simulated data and the corresponding theoretical predictions ‘drift away’ from the intermediate values of $\langle L \rangle$ (which are still obtained for $\theta = J$) towards either low ($\theta > J$) or high ($\theta < J$) values of $\langle L \rangle$. This is because the realized multiplex networks are either low-density or high-density,

which is an indication of a phase transition occurring when increasing the value of J , exactly as predicted by Figure 3.

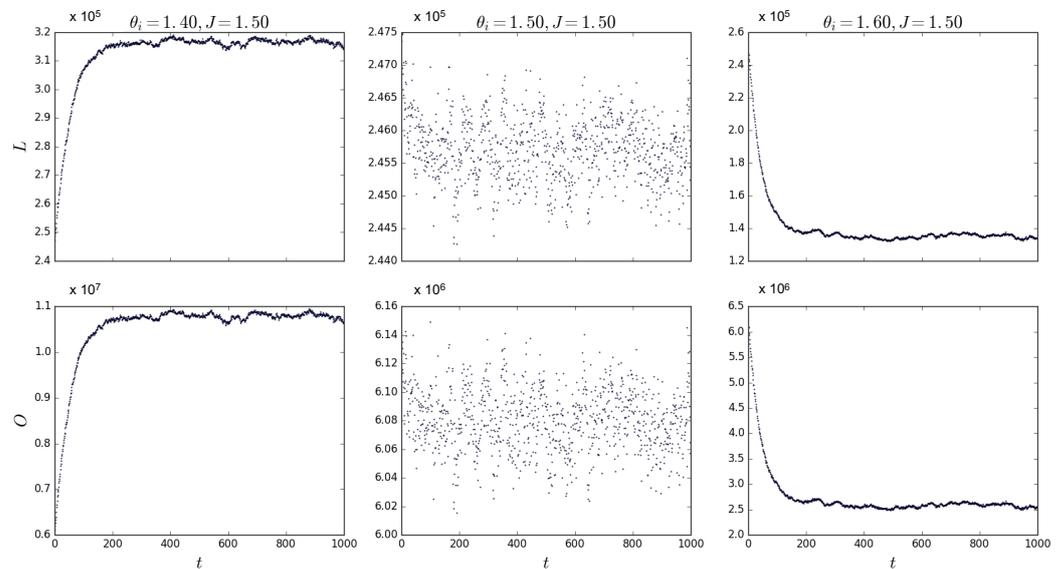


Figure 4. Total number of links L (top panels) and inter-layer overlap O (bottom panels) as a function of simulation time using the Metropolis–Hastings algorithm for $J = 1.5$, $N = 100$, $M = 100$. **Left** panels: $\theta = 1.4$. **Middle** panels: $\theta = 1.5 = J$ (symmetry-broken case). **Right** panels: $\theta = 1.6$. For fixed J , varying θ determines a phase transition from a high-density phase to a low-density phase.

We conclude our discussion of the homogeneous case by noting that, given an empirical multiplex \vec{G}^* of interest, the entropy of the data given, in general, by Equation (58) reduces, in this case, to

$$\begin{aligned}
 S(\theta^*, J^*) &= M\theta^* \sum_{i=1}^N \bar{k}_i^* - \frac{4J^*}{M} O^* + \sum_{i<j} \ln z_{ij}(2\theta^*, J^*) \\
 &= 2\theta^* L^* - \frac{4J^*}{M} O^* + \frac{N(N-1)}{2} \ln z(2\theta^*, J^*), \tag{72}
 \end{aligned}$$

where we have used Equation (9) (denoting, via $L^* = L(\vec{G}^*)$, the total number of links in the multiplex, which also equals the expected value $\langle L \rangle_{\theta^*, J^*}$) and the fact that the pair partition function z_{ij} , given by Equation (52), has the same value $z(2\theta^*, J^*) \equiv z_{ij}(2\theta^*, J^*)$ for all the $N(N-1)/2$ pairs of nodes. From Equation (72), we see that the entropy is determined, as expected, by both L^* and O^* . At the same time, we know that O^* depends uniquely and quadratically on L^* in this homogeneous model. The values achieved by the entropy are, therefore, bound by the relationship between L^* and O^* , which here is the same irrespective of the value of J^* , including when $J^* = 0$. In any case, the entropy also depends on the specific values of (θ^*, J^*) , and Equation (63) guarantees that an upper bound for $S(\theta^*, J^*)$ is given by the entropy $S_0(\theta^0)$ of the ACM model with $J^* = 0$ and $\theta^* = \theta^0$ (clearly, the homogeneity implies that $\theta_i^0 = \theta^0$ for all $i = 1, N$ in the ACM model as well).

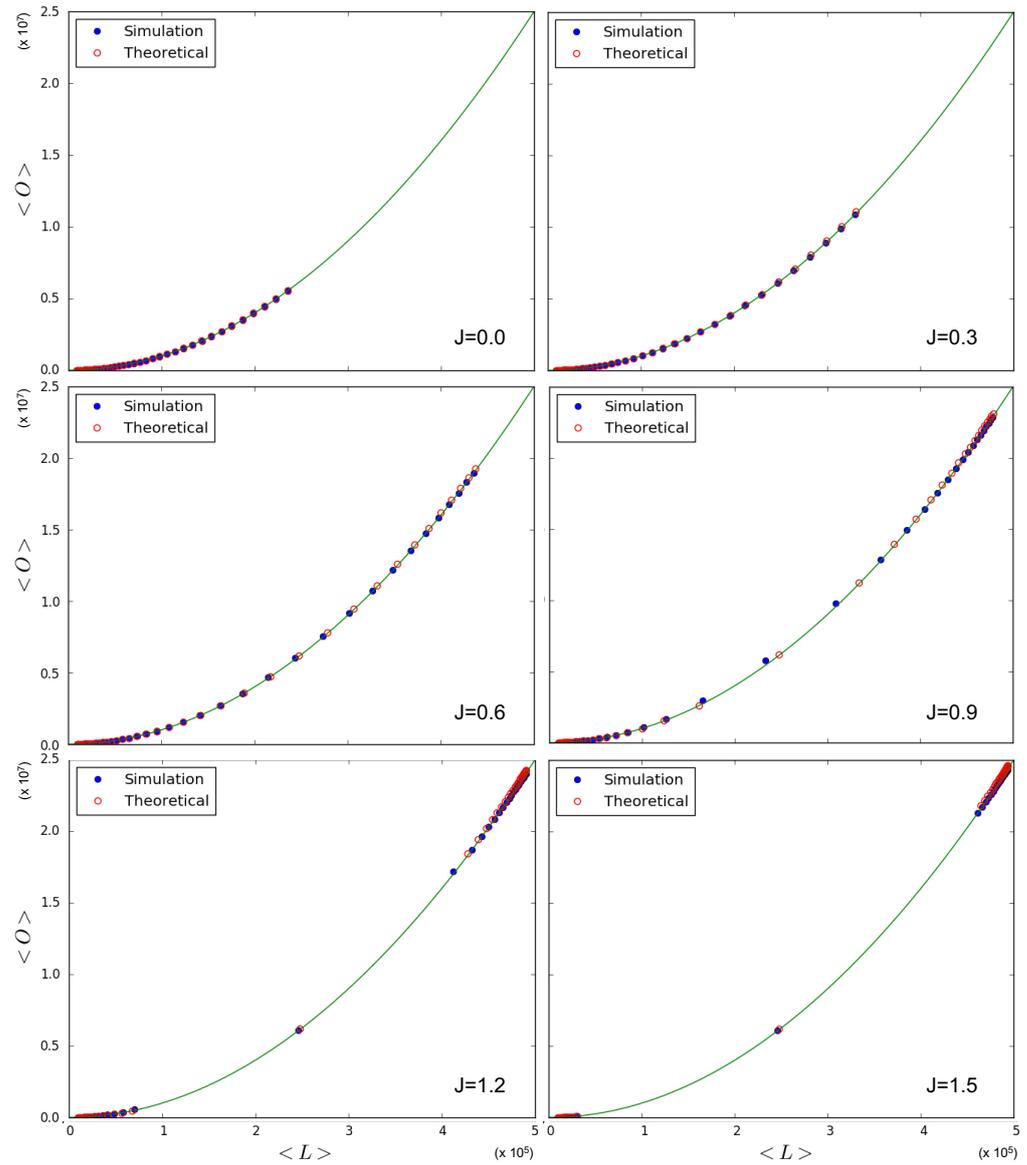


Figure 5. Relationship between the expected inter-layer overlap $\langle O \rangle$ and the total number of links $\langle L \rangle$ in homogeneous multiplexes with $N = 100$, $M = 100$, and $\theta_i = \theta$ for all $i = 1, N$. The blue points correspond to simulations obtained via the Metropolis–Hastings algorithm for $J \in \{0.0, 0.3, 0.6, 0.9, 1.2, 1.5\}$ and $\theta \in [0.05, 2.00]$ in steps of $\Delta\theta = 0.05$. The open red circles are the corresponding theoretically predicted points. The solid curve corresponds to the quadratic trend $\langle O \rangle = \langle L \rangle^2 / N^2$ predicted for all $J \geq 0$. Multiple solutions for u_{ij}^* first appear when $J > 1$, but the system keeps following the quadratic trend, albeit drifting away from the central point obtained for the zero-field case $\theta = J$ (corresponding to a spontaneously broken symmetry).

5.1.2. Power-Law-Distributed Fitness: Scale-Free Networks with Overlap

We now move away from the homogeneous case and consider a situation where the fitness values $\{x_i\}_{i=1}^N$ are drawn from a heavy-tailed distribution, in particular, a power law. This choice will produce a high degree of heterogeneity. In the ACM (see Section 2.5), the expected degree distribution is determined by the Lagrange multipliers θ_i , or equivalently, the transformed hidden variables $x_i = e^{-\theta_i}$. If x is distributed according to a power law, the expected degree distribution shall be distributed according to a power law as well, with the modulo as an upper cut-off. Since our OACM is an extension of the ACM, we will still sample x_i from a power law distribution $P(x) \sim x^{-\gamma}$ for various values

of γ , even though the expected degree distribution is not solely determined by the hidden variables $\{x_i\}$, but depends on J as well. In any case, a higher level of heterogeneity in the hidden variables x_i will lead to a higher level of heterogeneity in the degrees. Since the parameter space is rather large ($N + 1$ -dimensional), we define

$$x_i = zx_{0,i} \quad (73)$$

where z is a scaling factor. We sample $x_{0,i}$ only *once* from every chosen distribution. The value of x_i is varied by varying the scaling factor z . The parameter space to be explored will then be (z, J) , which is 2-dimensional. We deduce that

$$\theta_i = -\ln(zx_{0,i}) \quad (74)$$

which shows that an increasing z leads to a decreasing θ_i . In the ACM, we have shown that the link probability is equal to $p_{ij} = x_i x_j / (1 + x_i x_j)$, which means that larger values of x_i lead to a larger expected degree, so that increasing all the fitness values will increase the density in the network. This qualitative relationship still holds with the addition of the constraint on the expected overlap (for fixed J).

The complexity of Equations (53), (55), and (56) does not allow us to easily derive the expected relationship between the overlap and the number of links in the network, as was the case when θ_i was constant. It is, however, possible to visualize the relationship between the overlap and the number of links by using the Metropolis–Hastings algorithm. Figure 6 shows this relationship, where x_i is sampled from power law distributions with various values of γ , alongside the expected quadratic term previously observed to occur for homogeneous values of the fitness x_i (delta distribution). We see that the overlap for a given number of links is higher in the cases where x is drawn from a power law distribution than when x is drawn from a delta distribution, even though the coupling parameter J is kept constant. The cause of this difference lies in the level of heterogeneity of the fitness distribution: unlike the homogeneous case, now different pairs of nodes have very different values of $\theta_{ij} = \theta_i + \theta_j$, and, therefore, the condition $J = \theta_{ij}/2$ for the vanishing of the ‘external field’ B_{ij} (spontaneous symmetry-breaking condition) cannot be realized simultaneously by all pairs. The figure also shows the effect of different exponents of the power law distributions of the fitness. A smaller value of γ leads to a higher overlap for a given number of links. By increasing the value of γ , the power law distribution becomes more sharply peaked, and will therefore lead to more homogeneous networks. Note, however, that increasing the value of the coupling parameter J itself also leads to an increase in the overlap for a given number of links for the same distribution.

Importantly, the phase transition now occurs for different pairs of nodes as J is varied. Some pairs of nodes will be in the non-magnetized phase, while others will be in the magnetized phase. The effective number M_{eff} of independent layers will, in general, depend on the choice of parameters. Among the magnetized pairs, the realized values of the overlap are no longer those corresponding to the ensemble average (as in the homogeneous case), but typically to the symmetry-broken solution with lower energy (hence dictated by the value of θ_{ij}), because no other pair of nodes will, in general, exist with the same parameters and such that the two symmetry-broken values are averaged by the resulting value of the realized overlap. In particular, while for $0 < J < 1$ all node pairs are in the non-magnetized phase, as J increases from 1 towards larger values, the pairs of nodes that first undergo the phase transition are the ones with values $\theta_i + \theta_j$ that fall between the limits set by Equations (65) and (66). As those equations and Figure 2 show, there are more and more combinations $\theta_i + \theta_j$ entering the magnetized phase as J increases. When J is sufficiently large, all pairs will be magnetized. Clearly, for any two pairs of nodes, (i, j) and (i, k) , that share the same node, i , the values of $\theta_i + \theta_j$ and $\theta_i + \theta_k$ will be correlated, as they share the same term θ_i . This means that the pairs of nodes entering the magnetized phase typically have nodes in common, even if it would be incorrect to say that individual nodes enter the magnetized phase ‘one by one’, while this is certainly correct for individual node pairs, if the sum $\theta_i + \theta_j$ is different across all of them.

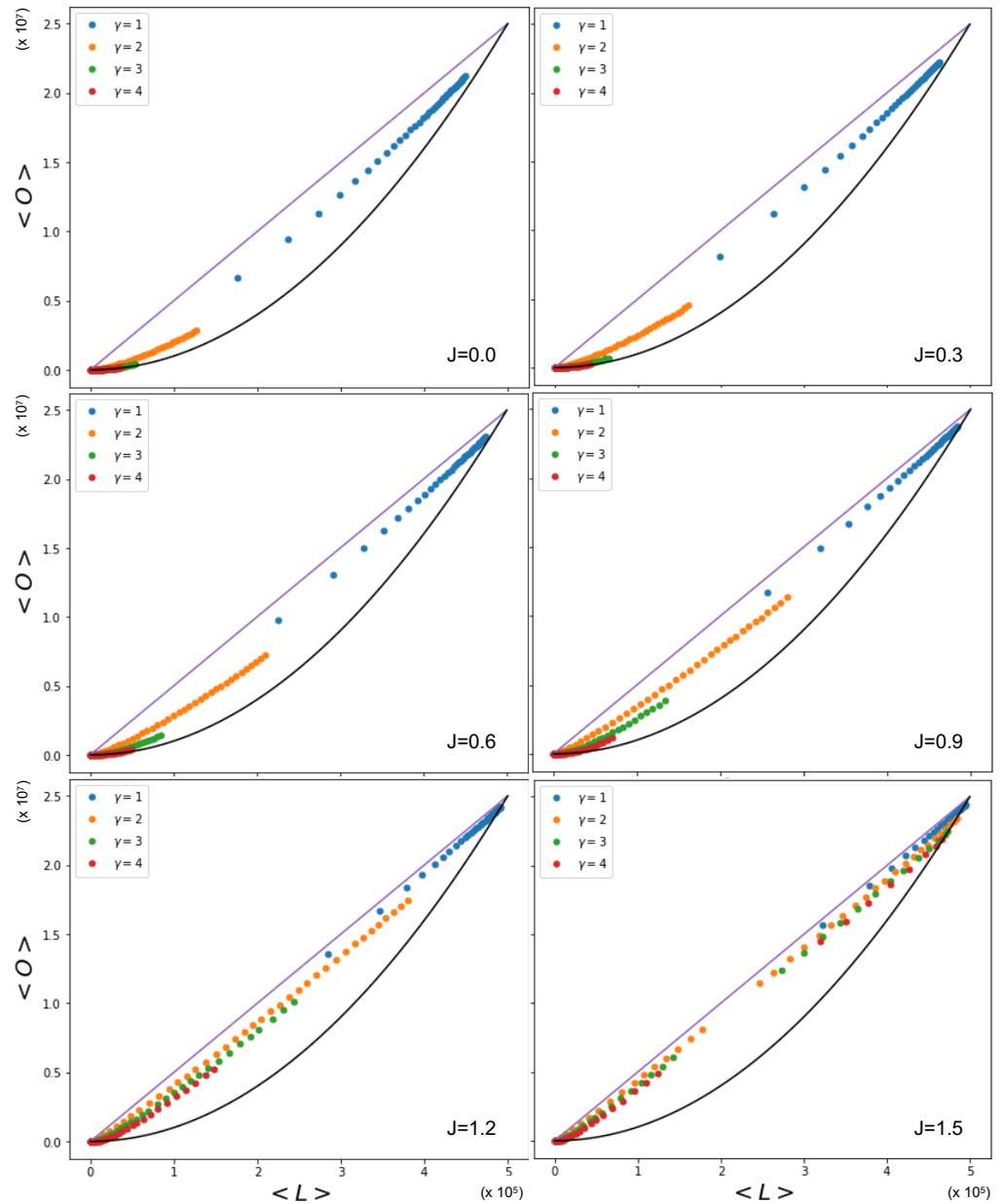


Figure 6. Relationship between the expected inter-layer overlap $\langle O \rangle$ and the total number of links $\langle L \rangle$ in heterogeneous multiplexes with $N = 100$, $M = 100$, and $x_{0,i}$ sampled from a power law distribution with different values for γ . The colored points correspond to simulations obtained via the Metropolis–Hastings algorithm for $J \in \{0.0, 0.3, 0.6, 0.9, 1.2, 1.5\}$ and $z \in [0.05, 2.00]$ in steps of $\Delta z = 0.05$. The straight line corresponds to the upper limit $\langle O \rangle = M\langle L \rangle / 2$ calculated in Equation (67). The solid curve corresponds to the quadratic trend $\langle O \rangle = \langle L \rangle^2 / N^2$ (achieved by homogeneous multiplexes with constant x_i), which here turns out to mark a lower bound. For increasing values of J , and especially as $J > 1$, the system moves closer to the upper bound. For $J = 1.5$, we see that the points are concentrating towards high-density and low-density (symmetry-broken) regimes, drifting away from the intermediate values, like in the homogeneous case. However, this is now the combined result of the behavior of statistically different pairs of nodes, each having a different zero-field condition $\theta_i + \theta_j = 2J$, so the spontaneous symmetry breaking cannot be realized for all node pairs simultaneously.

Figure 6 indeed shows the effect of the changing number of magnetized node pairs as J increases above 1. We note that, for larger and larger J , the relationship between $\langle O \rangle$ and $\langle L \rangle$ tends towards the ‘maximally multiplexed’ linear extreme (shown as a straight line) given in Equation (67). At the same time, we see that the ‘non-multiplexed’ case ($J < 1$) described by Equation (68) now realizes values of the overlap that are very different from the quadratic trend achieved by the homogeneous model (also shown as a solid curve in Figure 6), which now turns out to represent a lower bound. We can ‘zoom in’ to better see this difference by looking at Figure 7, where, by using Equations (53), (55), and (56), we additionally calculate the theoretically predicted values of $\langle O \rangle$ and $\langle L \rangle$ and compare them to the simulation data, where $x_{0,i}$ is sampled from a power law distribution with $\gamma = 1$ (the results for $\gamma \in \{2, 3, 4\}$ are qualitatively similar and are therefore not shown here). The figure confirms a strong deviation from the curve for the homogeneous model, even when $J = 0$ (signaling a much higher but spurious overlap, arising only from the rising correlation among node degrees across different layers), and a close agreement with the maximally overlapping value in Equation (67) already for $J = 1.5$ (corresponding to a further increase in overlap, arising from an additional, genuine coupling between layers).

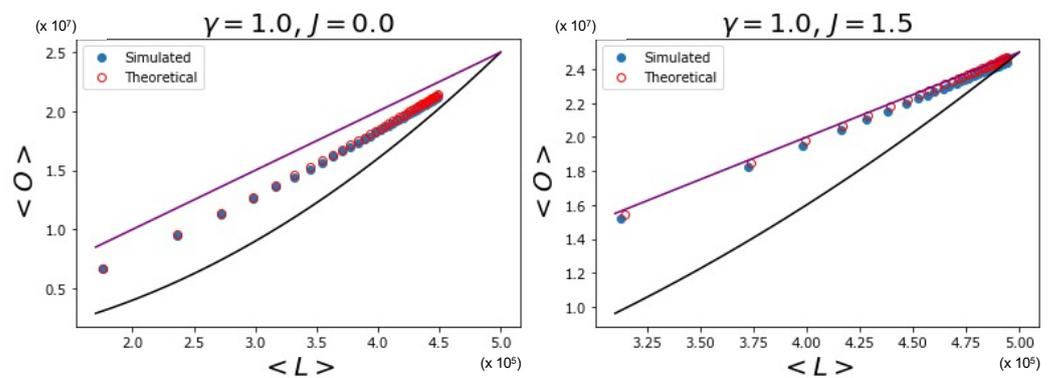


Figure 7. Relationship between the expected inter-layer overlap $\langle O \rangle$ and the total number of links $\langle L \rangle$ in heterogeneous multiplexes with $N = 100$, $M = 100$, and $x_{0,i}$ sampled from a power law distribution with $\gamma = 1$. The blue points correspond to simulations obtained via the Metropolis–Hastings algorithm for $z \in [0.05, 2.00]$ in steps of $\Delta z = 0.05$ with $J = 0$ (left panel) and $J = 1.5$ (right panel). The red open circles are the theoretically predicted values corresponding to the same parameters used in the simulations. The straight line corresponds to the upper limit $\langle O \rangle = M\langle L \rangle / 2$ calculated in Equation (67). The solid curve corresponds to the quadratic trend $\langle O \rangle = \langle L \rangle^2 / N^2$ (achieved by homogeneous multiplexes with constant x_i), which here turns out to mark a lower bound. We see that, compared with the homogeneous lower bound, the heterogeneity of nodes increases the overlap dramatically, even in the absence of true coupling ($J = 0$). When coupling is present, the overlap is additionally increased and already approaches the upper bound for $J = 1.5$.

5.1.3. Log-Normally Distributed Fitness

The delta and power law distributions we have considered so far represent examples of completely homogeneous and extremely heterogeneous (especially for $\gamma = 1$) distributions, respectively. We now consider the log-normal distribution as a third example between these two extremes. This analysis will indeed lead to results that are in some sense intermediate between what we have observed so far, and useful for interpreting the real-world case that we will present later on. A log-normal distribution is the distribution of a random variable whose logarithm is normally distributed (i.e., if the random variable x is log-normally distributed, then $y = \ln x$ follows a normal distribution). The probability density for a log-normal distribution is

$$P(x) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-(\ln x - \mu)^2 / (2\sigma^2)}, \tag{75}$$

where μ and σ correspond to the mean and the standard deviation of the normal distribution of $\ln x$. We will vary the value of x_i by again introducing a scaling factor that can be changed such that $x_i = zx_{0,i}$ and $\theta_i = -\ln(zx_{0,i})$, where we sample $x_{0,i}$ *once* from the log-normal distribution for a variety of values for μ and σ .

The log-normal distribution allows us to inspect the transition in the relationship between the overlap and the number of links from the quadratic lower limit to the linear upper limit by varying the value of σ . Indeed, when $0 < \sigma \ll 1$, the normal distribution of $\ln x_{0,i}$ is sharply peaked. By decreasing the value of σ towards 0, $\ln x_{0,i}$ (and, therefore, $x_{0,i}$ as well) shall approach a delta distribution. This is the distribution that led us to the quadratic lower limit for the relationship between the overlap and the number of links in the network. Conversely, when $\sigma \gg 1$, the log-normal distribution approaches a distribution with a power law tail with $\gamma = 1$. This distribution led us to the linear upper limit between the overlap and the number of links in the network (when J was sufficiently large). By increasing the value of σ from 0 to a sufficiently large value (e.g., $\sigma = 10$), we can therefore increase the heterogeneity of the network from a completely homogeneous network achieving the quadratic lower limit to an extremely heterogeneous network close to the linear upper limit relationship in the simulation data.

Figure 8 shows the relationship between the average overlap and the number of links in the network with simulation data that were obtained by using the Metropolis–Hastings algorithm for a variety of values for J and σ . Again, the linear upper limit is illustrated as a straight line and the quadratic lower limit as a solid curve. The figure confirms that in the case where $J = 0$, the data points that correspond to $x_{0,i}$ being sampled from a log-normal distribution with a relatively low value for σ are either on or close to the quadratic lower limit curve. On the other hand, the case where $\sigma = 10$ results in data points where the overlap in the network for a given number of links is almost maximal, and therefore approaches the linear upper limit. This first set of results confirms the strong role of node heterogeneity in determining increased correlations between the degrees of the same node across different layers, which, in turn, increase the inter-layer overlap even without any explicit coupling ($J = 0$), and hence, in a ‘spurious’ manner. On the other hand, when we increase the value of J , the data points corresponding to relatively low values of σ (e.g., $\sigma = 10^{-5}$ and $\sigma = 10^{-3}$) stay on or close to the quadratic lower limit, a finding similar to the results in Section 5.1.1, showing that the symmetry-broken values realized by different pairs of nodes, when averaged across the network, restore the ensemble average because the node pairs are all independent and (almost) identically distributed. Remarkably this means that, in a certain sense, node homogeneity ‘suppresses’ the effects of the true inter-layer coupling ($J > 0$) on the realized overlap. For the intermediate value $\sigma = 1.0$, the data are distributed close to the quadratic lower limit curve only for low values of J , while increasing the value of J leads to a more linear trend, eventually approaching the linear upper limit. In this case, the coupling is effective in producing a higher realized overlap. In the case where $\sigma = 10$, the linear trend is instead achieved already for $J = 0.0$ (although the points are aligned below it); hence, increasing the value of J barely influences the value of the overlap for a given number of links.

Therefore the effect of increasing J in networks with a moderate heterogeneity is a transition from multiplex configurations with densities of all levels towards multiplex configurations with either low or high density, which is a result of the phase transition. It also shows that a very high level of heterogeneity leads to an overlap in the network that is already close to maximal for a given number of links, irrespective of the phase transition and the value of J . However, in the case where we have an intermediate level of heterogeneity ($\sigma = 1.0$), we observe that the effect of the coupling can be relatively strong, and we can therefore construct networks with a combination of the overlap and number of links falling between the extreme linear upper limit and the quadratic lower limit in a controlled, systematic manner. Note that Figure 8 also shows that, as J increases above 1, the (symmetry-broken) realized data start to ‘drift away’ from the intermediate densities, in a way similar to what we observed in Figure 5, but in a more pronounced manner. This is due to the fact that, as J increases, a larger number of multilinks shall be either in the low-density or high-density phase.

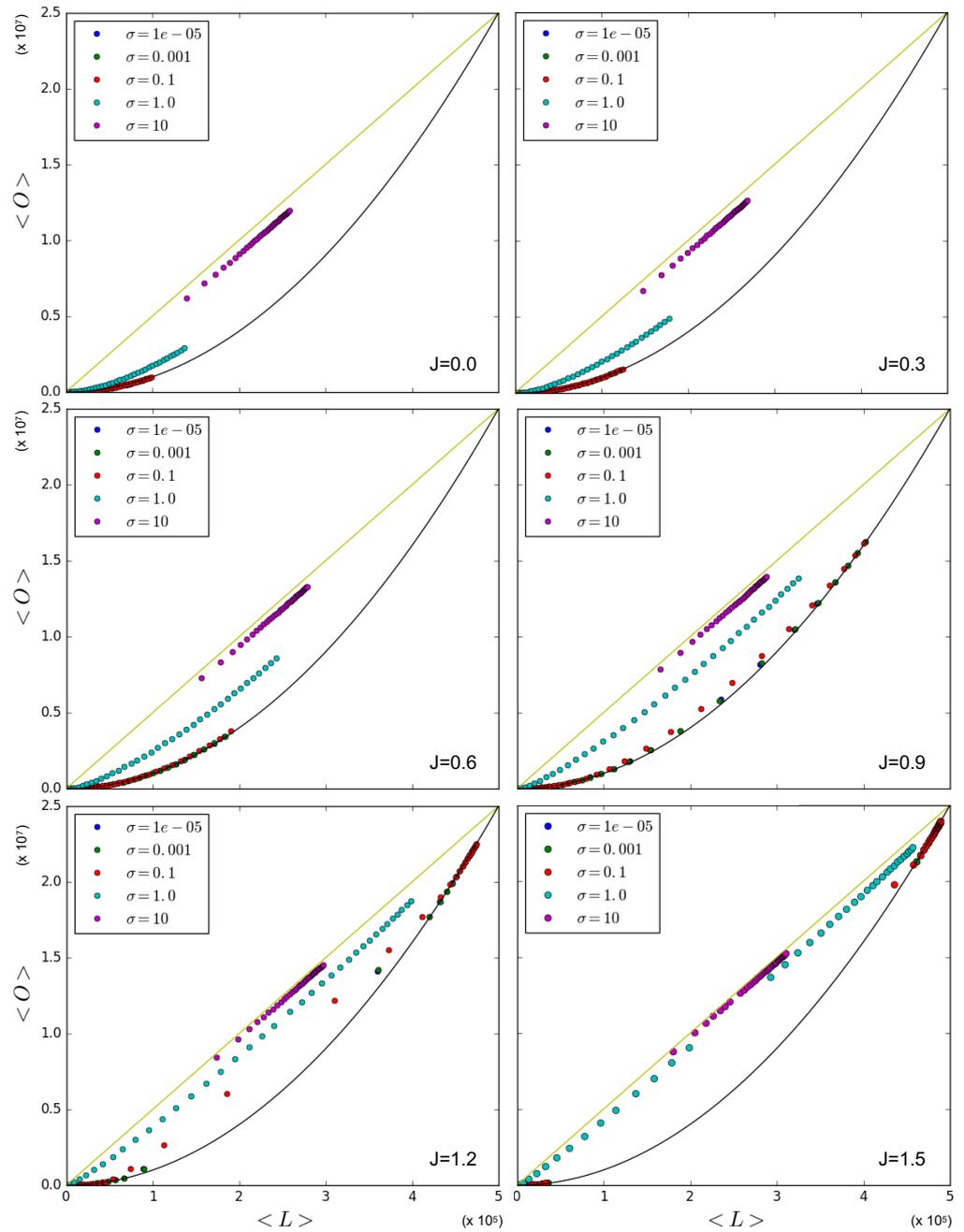


Figure 8. Relationship between the expected inter-layer overlap $\langle O \rangle$ and the total number of links $\langle L \rangle$ in heterogeneous multiplexes with $N = 100$, $M = 100$, and $x_{0,i}$ sampled from a log-normal distribution with different values for σ . The colored points correspond to simulations obtained via the Metropolis–Hastings algorithm for $J \in \{0.0, 0.3, 0.6, 0.9, 1.2, 1.5\}$ and $z \in [0.05, 2.00]$ in steps of $\Delta z = 0.05$. The straight line corresponds to the upper limit $\langle O \rangle = M\langle L \rangle/2$ calculated in Equation (67). The solid curve corresponds to the quadratic trend $\langle O \rangle = \langle L \rangle^2/N^2$ (achieved by homogeneous multiplexes with constant x_i), which here marks a lower bound achieved when $\sigma \rightarrow 0^+$. For increasing values of J (genuine coupling) and σ (spurious coupling), the system moves closer to the upper bound. For $J > 1$, we see that, starting from the multiplexes with smaller values of σ , the points are concentrating towards high-density and low-density (symmetry-broken) regimes, drifting away from the intermediate values, like in the homogeneous and power law cases. To realize this separation for larger values of σ , a larger value of J is required.

Again, in Figure 9 (which is the counterpart of Figure 7), we ‘zoom in’, and, using Equations (53), (55), and (56), we show the theoretically predicted values of $\langle O \rangle$ and $\langle L \rangle$ and compare them to the simulation data, where $x_{0,i}$ is sampled from a log-normal distribution with $\sigma = 1$, for $J = 0$ and $J = 1.5$. The results for $\sigma \in \{10^{-5}, 10^{-3}, 10^{-1}, 10^1\}$ are not shown here since relatively low and high values for σ lead to results similar to those we have shown in Sections 5.1.1 and 5.1.2, respectively. Figure 9 confirms that the theoretical predictions are in good agreement with the simulation data, apart from the expected ‘drifting away’ of symmetry-broken values from the corresponding ensemble average.

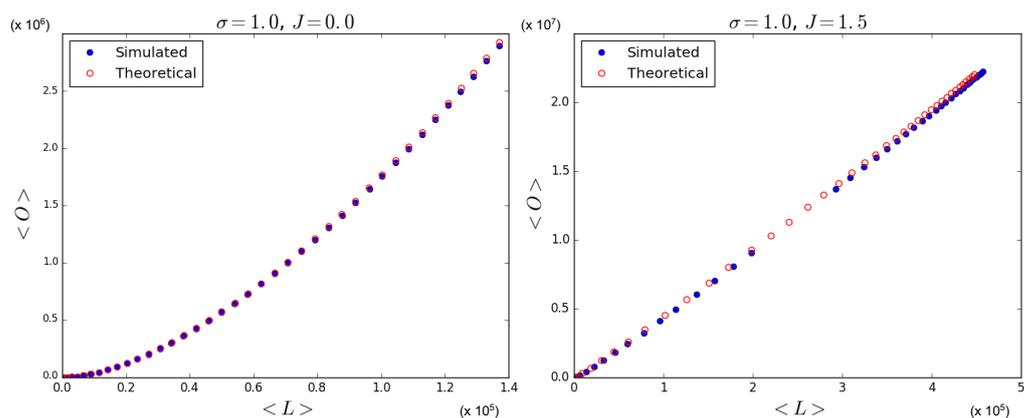


Figure 9. Relationship between the expected inter-layer overlap $\langle O \rangle$ and the total number of links $\langle L \rangle$ in heterogeneous multiplexes with $N = 100$, $M = 100$, and $x_{0,i}$ sampled from a log-normal distribution with $\sigma = 1$. The blue points correspond to simulations obtained via the Metropolis–Hastings algorithm for $z \in [0.05, 2.00]$ in steps of $\Delta z = 0.05$ with $J = 0$ (left panel) and $J = 1.5$ (right panel). The red open circles are the theoretically predicted values corresponding to the same parameters used in the simulations.

6. Analysis of the World Trade Multiplex

In this section, we finally consider an application of the model to a real-world economic network. Since our models lead to multiplex networks with independent pairs of nodes (i.e., independent multilinks) even when links are correlated across layers, it is important that the real-world network is consistent with this assumption. For instance, networks constructed from time series data [3–5] are not viable, because the known (and strong) correlations between the time series corresponding to different vertices generate dependencies between pairs of nodes (and higher-order patterns) through the triangular inequality [6,7]. For this reason, we select the World Trade Multiplex as an ideal case study for the present analysis, because each separate layer of that network has been successfully modeled in the past via maximum entropy models of networks with given degrees [31–33]. At the same time, it has been shown that certain structural properties of commodity-specific layers are very similar across the different layers of the multiplex [30], and that this similarity (in particular, the correlation among the degrees of the same node in different layers) generates a large spurious component of the inter-layer overlap [28,29], which is not necessarily due to a genuine coupling. In this sense, our analysis here will add a natural novel aspect to the modeling of the network, namely, the explicit comparison with a model with nontrivial coupling among layers, which has not been considered so far. We use the UN-COMTRADE dataset that represents the multiplex network of international trade (<https://comtradeplus.un.org>, accessed on 2 September 2019). The different layers of this multiplex network represent different commodities. The vertices in this network represent different countries, and a link exists between two countries in a given layer if there is trade between them in that commodity. The data include $N = 206$ countries and $M = 96$ commodities. Some examples of traded commodities are meat, fish, dairy products, coffee, and tobacco [30,33].

Using the international trade data, we wish to identify a possibly nontrivial overlap by creating (L, O) plots similar to the ones depicted in Figures 6 and 7 or 8. We therefore repeatedly filter the network such that each layer α has the same number of links $L^\alpha \equiv L^0$ (where $\alpha = 1, \dots, M$), and calculate the corresponding overlap O for the specified value of L^0 (note that this means that the total number of links in the entire multiplex is $L = ML^0$). The criterion we follow is choosing the L^0 strongest (highest weight) links in every layer to obtain data with comparable degrees across layers, as in our models. Note that, by using this filtering method, the highest possible density we can achieve is limited by the density of the sparsest layer in the unfiltered network. The results are shown in Figure 10, which indicates that the overlap for a given number of links appears to be around halfway between the quadratic lower limit curve and the linear upper limit curve. This suggests that the degree of heterogeneity of the network is intermediate, similar to that realized by log-normally distributed fitness values, as in our example considered above.

As anticipated, we are currently unable to solve the maximum likelihood equations in order to obtain the joint values of all the Lagrange multipliers in the full OACM model with $J \neq 0$. However, after filtering the original empirical network such that every layer has L^0 links, we can use the values of the hidden variables x_i^* for the null model corresponding to the absence of inter-layer coupling, i.e., to $J^* = 0$. As we have shown in Equations (57), this assumption reduces our model to the ACM discussed in Section 2.5. The maximum likelihood equations in this case are much easier to solve, and can be found using one of the numerical algorithms available at <https://meh.imtlucca.it> (accessed on 1 May 2023). This procedure is repeated for a range of values for L^0 . The cumulative distribution of the hidden variables x_i^* are plotted in Figure 10 for various values of L^0 . The figure qualitatively shows that the shape of the cumulative distribution of x is fat-tailed and indeed similar to the one for a log-normal distribution. Moreover, it does not vary with L^0 , apart from an overall change of scale.

The null model with $J^* = 0$, when compared to the data for the same choice of L_0 , allows us to detect the presence of nontrivial coupling among the layers, when present. Indeed, from Figure 10, we see that the filtered networks have a relatively high overlap, the data points being distributed along a similar trend as the one corresponding to a nonzero J in our previous heterogeneous examples. By using the values of the hidden variables for the model with $J^* = 0$, we can calculate the corresponding expected number of links and the expected overlap under the null hypothesis of no coupling between the layers, but the same average degree sequence in the real network. The results are shown in Figure 10, alongside the curve corresponding to the empirical data. We see that the assumption $J = 0$ leads to an insufficiently overlapping multiplex, demonstrating the necessity of a model that introduces dependencies between the layers of a network. The difference between the two curves can be quantified by fitting both to the curve

$$O = AL^\alpha \quad (76)$$

where A is a proportionality factor and α is an exponent (not to be confused with the label of a layer of the multiplex). For the empirical data, we find a steeper increase characterized by an exponent $\alpha_{\text{empirical}} = 1.19$, while for the predictions from the ACM, we find $\alpha_{\text{CM}} = 1.06$ (see Figure 10). The difference between the two values implies that the difference between the realized and expected overlap increases as L increases, confirming that the observed overlap in the WTM is not only the spurious result of the correlated heterogeneity of the degrees of countries, but reflects genuine ($J^* > 0$) inter-layer dependencies.

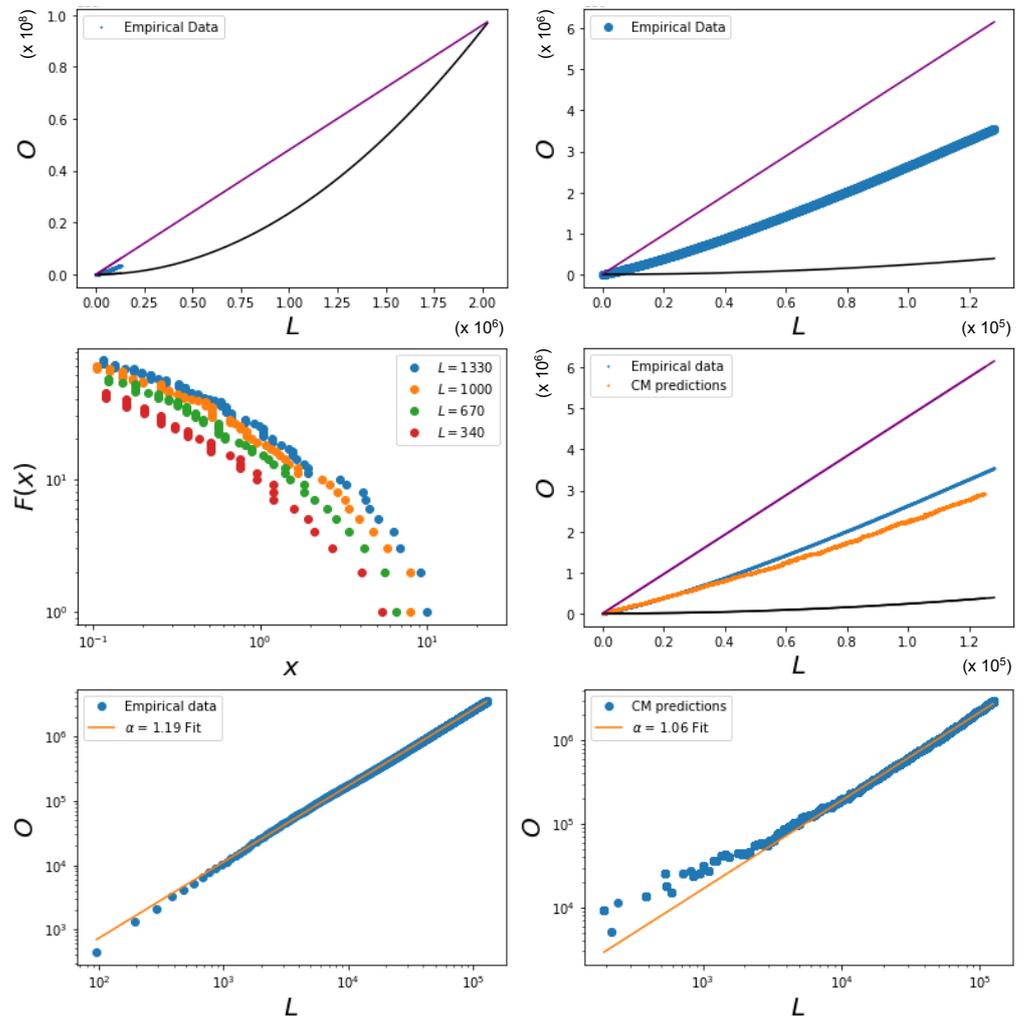


Figure 10. Comparison of the empirical World Trade Multiplex (WTM) with the zero-coupling ($J^* = 0$) benchmark provided by the Average Configuration Model (ACM). The WTM consists of $N = 206$ nodes, each representing a country, and $M = 96$ layers, each representing a commodity group. The filtered data were obtained by retaining the same number L^0 of strongest links in each layer (hence $L = ML^0$ links in the entire multiplex), and varying L^0 . **Top left:** relationship between the expected inter-layer overlap $\langle O \rangle$ and the total number of links (L) in the WTM (blue), compared with the upper limit $\langle O \rangle = M\langle L \rangle / 2$ calculated in Equation (67) (purple straight line) and the quadratic trend $\langle O \rangle = \langle L \rangle^2 / N^2$ achieved by homogeneous multiplexes (black solid curve). **Top right:** zoomed-in version of the top left panel, showing that the empirical data follow an intermediate scaling between the two extremes. **Center left:** cumulative distributions reporting the number $F(x)$ of nodes with hidden variable larger than x in the ACM, obtained for different values of L^0 (see legend). **Center right:** same as the top right panel with the addition of the relationship produced by the ACM benchmark, showing that the empirical WTM (blue) has a higher overlap than the corresponding null model having zero inter-layer coupling but the same degree heterogeneity (orange). **Bottom left:** log–log plot of the relationship between the overlap and the number of links in the empirical WTM, along with a power law fit of the form $O = AL^\alpha$, where the fitted exponent is $\alpha = 1.19$. **Bottom right:** log–log plot of the same relationship in the ACM benchmark with no coupling, along with a power law fit of the form $O = AL^\alpha$, where the fitted exponent is $\alpha = 1.06$.

We conclude with a discussion about the entropy in the heterogeneous case, analogous to the one we made in Section 5.1.1 in the homogeneous case. Here we note that, given a

multiplex \vec{G}^* of interest, the entropy $S(\vec{\theta}^*, J^*)$ of the data, given the OACM model, is the one given by Equation (58), which in the heterogeneous case cannot be, in general, reduced to a simpler formula. However, if we define the minimum and maximum values of the hidden variables as

$$\theta_{\min}^* \equiv \min_{i=1,N} \{\theta_i^*\}, \quad \theta_{\max}^* \equiv \max_{i=1,N} \{\theta_i^*\}, \quad (77)$$

respectively, we can bound the entropy as follows:

$$S_{\min}(\vec{\theta}^*, J^*) \leq S(\vec{\theta}^*, J^*) \leq S_{\max}(\vec{\theta}^*, J^*) \quad (78)$$

where we have defined

$$S_{\min}(\vec{\theta}^*, J^*) \equiv 2\theta_{\min}^* L^* - \frac{4J^*}{M} O^* + \sum_{i < j} \ln z_{ij}(\theta_i^* + \theta_j^*, J^*), \quad (79)$$

$$S_{\max}(\vec{\theta}^*, J^*) \equiv 2\theta_{\max}^* L^* - \frac{4J^*}{M} O^* + \sum_{i < j} \ln z_{ij}(\theta_i^* + \theta_j^*, J^*). \quad (80)$$

The bounds in Equation (78) are alternative to the general ones in Equation (63), and arguably more useful to characterize how the entropy is effectively constrained by, once again, the relationship between L^* and O^* . The latter, unlike the homogeneous case, is not necessarily quadratic, and can follow the diverse trends we have shown in Figures 6, 8 and 10. In particular, the power law relationship captured by Equation (76) for the empirical WTM provides a convenient way of bounding $S(\vec{\theta}^*, J^*)$ via Equations (78)–(80).

7. Conclusions

In this paper we have introduced a maximum entropy model, or ERGM, of multiplex networks with given degrees and inter-layer overlap. The model allowed us to separately control the effects of the correlations between node degrees across different layers (which lead to a spurious overlap) and that of a genuine inter-layer coupling. The nature of the enforced constraints is such that different pairs of nodes are statistically independent, even if the parameters governing them are correlated via those of the nodes they share.

For each pair of nodes, the model can be mapped exactly to a mean-field Ising model featuring a magnetization-like phase transition, which includes the possibility of (spontaneous) symmetry breaking. Given the difficulty of solving the maximum likelihood equations to obtain the values of the Lagrange multipliers corresponding to a particular real network, we first treated the Lagrange multipliers as free parameters in order to explore and analyze the properties of multiplex systems as a function of these parameters using numerical methods. Additionally, the numerical results were compared to our analytical results in order to test the validity of the latter. We have shown that the analytical equations are highly accurate. The combined result, at the level of the entire multiplex, of the properties of all node pairs is nontrivial and crucially depends on the values of the node-specific parameters, which ultimately depend on the enforced degrees.

In the fully homogeneous case, the phase transition occurs at the same critical point for all node pairs simultaneously, because the parameters are identical for all nodes. However, the independence of different node pairs implies that, even in the magnetized phase, the realized values of the inter-layer overlap and total number of links coincide with the ensemble average. This happens because different node pairs realize all the possible symmetry-broken values independently, so that an average of the realized values for a large number of independent node pairs asymptotically equals the ensemble average. The value of J has little effect on the relationship connecting the overlap to the number of links, which remains similar to what we observed for the case $J = 0$, showing that node homogeneity suppresses the effects of a genuine inter-layer coupling.

In the heterogeneous case, the phenomenology is very different since, despite the fact that node pairs are still independent, they are now governed by different parameters, and the ensemble average for a given pair can no longer be realized as an average of the realized values of pairs with the same parameters. This implies that the observed overlap

and number of links will depend on the realized symmetry-broken values, whose typical value does not coincide in general with the ensemble average, and is determined by the node-specific parameters (hence, ultimately by the degrees). Moreover different pairs of nodes are, in general, found in different phases, so the multiplex displays, as a function of the parameters, a hierarchy of phase transitions. We have found that increasing the value of the coupling parameter J generally increases the (genuine) overlap for a given number of links, if there is enough node heterogeneity. However, we have also shown that increasing the heterogeneity of the network increases the (spurious) overlap for a given number of links as well. This is a consequence of the presence of large hubs that appear in a correlated manner across layers, due to the increased heterogeneity of the network. Additionally, every multilink that is connected to these hubs has a relatively low critical threshold for the coupling parameter J . Therefore, these multilinks have a higher probability to be in the high density phase, which leads to a higher overlap as well, which corresponds to increasing the amount of genuine correlation. In general, the overlap for a given number of links can be increased by increasing either the heterogeneity of the network or the value of the coupling parameter, with a subtle interplay between the two. In principle, this can be used in order to create multiplexes with a specific degree of overlap for a given number of links, provided their combination is within the theoretical limits discussed in Section 5.

Finally, by using a dataset that represents the empirical multiplex network of international trade in several commodity-specific layers, we have used the model to disentangle the spurious overlap arising from the documented strong correlation of node degrees across layers [28,29] from the genuine overlap arising from actual inter-layer coupling. We have found that the assumption that there is no coupling between the layers ($J = 0$), which reduces our model to the ACM, results in a multiplex with insufficient inter-layer overlap. This means that the empirical overlap is not merely the spurious result of the correlated heterogeneity of the network, but requires a true nonzero coupling between layers.

Our results demonstrate the subtleties of the interplay between node heterogeneity and inter-layer dependencies in multiplex networks, highlighting the need for null models that can control these factors separately. In this paper, we have introduced perhaps the simplest, although already very rich, model of this type. Our model can be seen as a minimal one, to be further generalized in the future.

Author Contributions: Conceptualization, D.G. and V.G.; methodology, D.G., V.G. and N.B.; software, N.B.; data curation, V.G. and N.B.; writing—original draft preparation, N.B.; writing—review and editing, V.G. and D.G.; visualization, N.B.; supervision, D.G.; project administration, V.G.; funding acquisition, D.G. All authors have read and agreed to the published version of the manuscript.

Funding: The APC was funded by Stichting Econophysics, Leiden, The Netherlands.

Data Availability Statement: For the empirical analysis of the World Trade Multiplex, the publicly available UN-COMTRADE dataset was analyzed in this study. The data are available at <http://comtrade.un.org/> (accessed on 2 September 2019). The codes used for the numerical calculation of the parameters maximizing the likelihood are available at <https://meh.imtlucca.it> (accessed on 1 May 2023).

Acknowledgments: This work is supported by the European Union—Horizon 2020 Program under the scheme ‘INFRAIA-01-2018-2019—Integrating Activities for Advanced Communities’, Grant Agreement n.871042, ‘SoBigData++: European Integrated Infrastructure for Social Mining and Big Data Analytics’ (<http://www.sobigdata.eu>, accessed on 1 May 2023). This work is also supported by the European Union-NextGenerationEU-National Recovery and Resilience Plan (Piano Nazionale di Ripresa e Resilienza, PNRR), project ‘SoBigData.it-Strengthening the Italian RI for Social Mining and Big Data Analytics’-Grant IR0000013 (3264, 28/12/2021). We also acknowledge support from the project NetRes—‘Network analysis of economic and financial resilience’, Italian DM n. 289, 25-03-2021 (PRO3 Scuole), CUP D67G22000130001 (<https://netres.imtlucca.it>, accessed on 1 May 2023).

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Hubbard–Stratonovich Transform

The pair Hamiltonian of our OACM in Equation (47) can be rewritten as

$$h_{ij}(s_{ij}, B_{ij}, J) = -B_{ij} \sum_{\alpha=1}^M \sigma_{ij}^\alpha - \frac{J}{2M} \left(\sum_{\alpha=1}^M \sigma_{ij}^\alpha \right)^2 + \frac{J}{2} + v_{ij}. \tag{A1}$$

We want to obtain an expression for the pair partition function:

$$\begin{aligned} z_{ij}(B_{ij}, J) &= \sum_{s_{ij} \in \mathcal{S}_{ij}} e^{-h_{ij}(s_{ij}, B_{ij}, J)} \\ &= \sum_{s_{ij} \in \mathcal{S}_{ij}} \exp \left[\frac{J}{2M} \left(\sum_{\alpha=1}^M \sigma_{ij}^\alpha \right)^2 + B_{ij} \sum_{\alpha=1}^M \sigma_{ij}^\alpha - \frac{J}{2} - v_{ij} \right] \\ &= e^{-J/2} e^{-v_{ij}} \sum_{s_{ij} \in \mathcal{S}_{ij}} \exp \left[\left(\sqrt{\frac{J}{2M}} \sum_{\alpha=1}^M \sigma_{ij}^\alpha \right)^2 + B_{ij} \sum_{\alpha=1}^M \sigma_{ij}^\alpha \right]. \end{aligned} \tag{A2}$$

The argument of the exponent in the above expression can be linearized by using the Gaussian integral

$$e^{a^2} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} d\zeta_{ij} e^{-\zeta_{ij}^2/2 + \sqrt{2}a\zeta_{ij}}. \tag{A3}$$

In our case, by choosing $a = \sqrt{J/(2M)} \sum_{\alpha=1}^M \sigma_{ij}^\alpha$ the partition function factorizes with respect to the individual summations of σ_{ij}^α :

$$\begin{aligned} z_{ij}(B_{ij}, J) &= \frac{e^{-J/2-v_{ij}}}{\sqrt{2\pi}} \sum_{s_{ij} \in \mathcal{S}_{ij}} \int_{-\infty}^{\infty} d\zeta_{ij} e^{-\zeta_{ij}^2/2} \exp \left[\sum_{\alpha=1}^M \sigma_{ij}^\alpha \left(\sqrt{\frac{J}{M}} \zeta_{ij} + B_{ij} \right) \right] \\ &= \frac{e^{-J/2-v_{ij}}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} d\zeta_{ij} e^{-\zeta_{ij}^2/2} \sum_{\sigma_{ij}^1 \in \{-1,1\}} \dots \sum_{\sigma_{ij}^M \in \{-1,1\}} \prod_{\alpha=1}^M \exp \left[\sigma_{ij}^\alpha \left(\sqrt{\frac{J}{M}} \zeta_{ij} + B_{ij} \right) \right] \\ &= \frac{e^{-J/2-v_{ij}} 2^M}{\sqrt{2\pi}} \int_{-\infty}^{\infty} d\zeta_{ij} e^{-\zeta_{ij}^2/2} \left[\cosh \left(\sqrt{\frac{J}{M}} \zeta_{ij} + B_{ij} \right) \right]^M. \end{aligned} \tag{A4}$$

Performing the change of variable $\sqrt{J/M} \zeta_{ij} = Jy_{ij}$ we obtain

$$z_{ij}(B_{ij}, J) = 2^M \sqrt{\frac{JM}{2\pi}} e^{-J/2} e^{-v_{ij}} \int_{-\infty}^{\infty} d\zeta_{ij} \left[\Phi_{J,B_{ij}}(y_{ij}) \right]^M \tag{A5}$$

where

$$\Phi_{J,B_{ij}} \equiv e^{-Jy_{ij}^2/2} \cosh(Jy_{ij} + B_{ij}). \tag{A6}$$

We are interested in the large M limit. To proceed in the calculation of $z_{ij}(B_{ij}, J)$, it is useful to define the quantity

$$f_{ij}(B_{ij}, J) \equiv - \lim_{M \rightarrow \infty} \frac{1}{M} \ln z_{ij}(B_{ij}, J) = - \lim_{M \rightarrow \infty} \ln z_{ij}^{1/M}(B_{ij}, J) \tag{A7}$$

which is the *free energy per layer*. By inserting the result (A5) into (A7), we obtain

$$\begin{aligned} f_{ij}(B_{ij}, J) &= - \ln 2 - \lim_{M \rightarrow \infty} \frac{1}{M} \left[\ln \left(e^{-J/2} \sqrt{\frac{JM}{2\pi}} \right) - v_{ij} + \ln \left(\int_{-\infty}^{\infty} dy_{ij} \left[\Phi_{J,B_{ij}}(y) \right]^M \right) \right] \\ &= - \ln 2 + \frac{J}{2} - B_{ij} - \ln \left[\lim_{M \rightarrow \infty} \left(\int_{-\infty}^{\infty} dy_{ij} \left[\Phi_{J,B_{ij}}(y) \right]^M \right)^{1/M} \right]. \end{aligned} \tag{A8}$$

In order to obtain a more explicit form of the function $f_{ij(B_{ij}, J)}$, we use the Laplace theorem [49]. Let $\phi(y)$ and $\psi(y)$ be continuous and positive functions within a range $c \leq y \leq d$, then

$$\lim_{M \rightarrow \infty} \left[\int_c^d \psi(y) (\phi(y))^M \right]^{1/M} = \max_{c \leq y \leq d} \phi(y). \tag{A9}$$

For $\psi(y) = 1$ and $\phi(y) = \Phi_{J, B_{ij}}(y)$, this results in

$$f_{ij}(B_{ij}, J) = -\ln 2 + \frac{J}{2} - B_{ij} - \ln \left[\max_{-\infty \leq y_{ij} \leq \infty} \Phi_{J, B_{ij}}(y_{ij}) \right]. \tag{A10}$$

The derivative of $\Phi_{J, B_{ij}}(y_{ij})$ with respect to y_{ij} is zero at its maximum:

$$\frac{d\Phi_{J, B_{ij}}(y_{ij})}{dy_{ij}} = J e^{-Jy_{ij}^2/2} \sinh(Jy_{ij} + B_{ij}) - Jy_{ij} e^{-Jy_{ij}^2/2} \cosh(Jy_{ij} + B_{ij}) = 0. \tag{A11}$$

The variable y_{ij} therefore obeys the equation

$$y_{ij} = \tanh(Jy_{ij} + B_{ij}). \tag{A12}$$

Note that this equation is identical to the one obtained for the magnetization in the Ising Model, and, depending on the values of J and B_{ij} , there are either one or three solutions that satisfy Equation (A12). The free energy f_{ij} can now be written as a function of J and B_{ij} :

$$f_{ij}(B_{ij}, J) = -\ln 2 + \frac{J}{2} - B_{ij} + \frac{J}{2} (y_{ij})^2 - \ln [\cosh(Jy_{ij} + B_{ij})]. \tag{A13}$$

We then finally arrive at the pair partition function

$$\begin{aligned} z_{ij}(B_{ij}, J) &= e^{-Mf_{ij}} \\ &= 2^M e^{-v_{ij}} e^{-JM(y_{ij})^2/2} \cosh^M(Jy_{ij} + B_{ij}) \end{aligned} \tag{A14}$$

which, returning to the variables θ_{ij} , coincides with Equation (52) in the main text, where u_{ij} is the solution to Equation (53).

Appendix B. Maximum Likelihood

To determine the parameters $(\vec{\theta}^*, J^*)$ that maximize the log-likelihood of the OACM given in Equation (54), we first calculate the derivatives

$$-\frac{\partial \mathcal{L}(\vec{\theta}, J)}{\partial \theta_k} = \sum_{i < j} \frac{\partial h_{ij}(m_{ij}^*, \theta_{ij}, J)}{\partial \theta_k} + \sum_{i < j} \frac{\partial \ln z_{ij}(\theta_{ij}, J)}{\partial \theta_k}, \quad k = 1, \dots, N \tag{A15}$$

$$-\frac{\partial \mathcal{L}(\vec{\theta}, J)}{\partial J} = \sum_{i < j} \frac{\partial h_{ij}(m_{ij}^*, \theta_{ij}, J)}{\partial J} + \sum_{i < j} \frac{\partial \ln z_{ij}(\theta_{ij}, J)}{\partial J}. \tag{A16}$$

We then set the derivatives with respect to θ_k to zero:

$$-\frac{\partial \mathcal{L}(\vec{\theta}, J)}{\partial \theta_k} \Big|_{\vec{\theta}^*, J^*} = \sum_{\alpha=1}^M \sum_{j \neq k} g_{jk}^{\alpha} - M \sum_{j \neq k} u_{jk}^* = 0 \tag{A17}$$

where we utilize the fact that g_{ij}^{α} and u_{ij} are symmetric with respect to the indices i, j , i.e.,

$$\sum_{i < j} g_{ij}^{\alpha} \delta_i^k = \sum_{j=k+1}^N g_{jk}^{\alpha}, \quad \sum_{i < j} g_{ij}^{\alpha} \delta_j^k = \sum_{j=1}^{k-1} g_{jk}^{\alpha}. \tag{A18}$$

Similarly, we set the derivative with respect to J to zero:

$$-\left. \frac{\partial \mathcal{L}(\vec{\theta}, J)}{\partial J} \right|_{\vec{\theta}^*, J^*} = \sum_{i < j} \left(-\frac{4}{M} \sum_{\alpha < \beta} g_{ij}^{*\alpha} g_{ij}^{*\beta} + 2M(u_{ij}^*)^2 \right) = 0. \quad (\text{A19})$$

Taken together, the above calculations lead to the maximum likelihood Equations (55) and (56) in the main text.

References

- Krackhardt, D. Cognitive social structures. *Soc. Netw.* **1987**, *9*, 109–134. [[CrossRef](#)]
- Padgett, J.F.; Ansell, C.K. Robust Action and the Rise of the Medici, 1400–1434. *Am. J. Sociol.* **1993**, *98*, 1259–1319. [[CrossRef](#)]
- Bardoscia, M.; Barucca, P.; Battiston, S.; Caccioli, F.; Cimini, G.; Garlaschelli, D.; Saracco, F.; Squartini, T.; Caldarelli, G. The physics of financial networks. *Nat. Rev. Phys.* **2021**, *3*, 490–507. [[CrossRef](#)]
- Lacasa, L.; Luque, B.; Ballesteros, F.; Luque, J.; Nuno, J.C. From time series to complex networks: The visibility graph. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 4972–4975. [[CrossRef](#)] [[PubMed](#)]
- Tsiotas, D.; Magafas, L.; Argyrakakis, P. An electrostatics method for converting a time-series into a weighted complex network. *Sci. Rep.* **2021**, *11*, 11785. [[CrossRef](#)]
- MacMahon, M.; Garlaschelli, D. Community Detection for Correlation Matrices. *Phys. Rev. X* **2015**, *5*, 021006. [[CrossRef](#)]
- Anagnostou, I.; Squartini, T.; Kandhai, D.; Garlaschelli, D. Uncovering the mesoscale structure of the credit default swap market to improve portfolio risk modelling. *Quant. Financ.* **2021**, *21*, 1501–1518. [[CrossRef](#)]
- De Domenico, M.; Solé-Ribalta, A.; Cozzo, E.; Kivelä, M.; Moreno, Y.; Porter, M.A.; Gómez, S.; Arenas, A. Mathematical formulation of multilayer networks. *Phys. Rev. X* **2013**, *3*, 041022. [[CrossRef](#)]
- Battiston, F.; Nicosia, V.; Latora, V. Structural measures for multiplex networks. *Phys. Rev. E* **2014**, *89*, 032804. [[CrossRef](#)]
- Kivelä, M.; Arenas, A.; Barthelemy, M.; Gleeson, J.P.; Moreno, Y.; Porter, M.A. Multilayer networks. *J. Complex Netw.* **2014**, *2*, 203–271. [[CrossRef](#)]
- Battiston, F.; Nicosia, V.; Latora, V. The new challenges of multiplex networks: Measures and models. *Eur. Phys. J. Spec. Top.* **2017**, *226*, 401–416. [[CrossRef](#)]
- Boccaletti, S.; Bianconi, G.; Criado, R.; Del Genio, C.I.; Gómez-Gardenes, J.; Romance, M.; Sendina-Nadal, I.; Wang, Z.; Zanin, M. The structure and dynamics of multilayer networks. *Phys. Rep.* **2014**, *544*, 1–122. [[CrossRef](#)] [[PubMed](#)]
- Verbrugge, L.M. Multiplexity in adult friendships. *Soc. Forces* **1979**, *57*, 1286–1309. [[CrossRef](#)]
- Erdos, P.; Rényi, A. On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci.* **1960**, *5*, 17–60.
- Albert, R.; Barabási, A.L. Statistical mechanics of complex networks. *Rev. Mod. Phys.* **2002**, *74*, 47. [[CrossRef](#)]
- Watts, D.J.; Strogatz, S.H. Collective dynamics of ‘small-world’ networks. *Nature* **1998**, *393*, 440. [[CrossRef](#)]
- Squartini, T.; Garlaschelli, D. Analytical maximum-likelihood method to detect patterns in real networks. *New J. Phys.* **2011**, *13*, 083001. [[CrossRef](#)]
- Squartini, T.; Mastrandrea, R.; Garlaschelli, D. Unbiased sampling of network ensembles. *New J. Phys.* **2015**, *17*, 023052. [[CrossRef](#)]
- Squartini, T.; Garlaschelli, D. *Maximum-Entropy Networks: Pattern Detection, Network Reconstruction and Graph Combinatorics*; Springer: Berlin/Heidelberg, Germany, 2017.
- Cimini, G.; Squartini, T.; Saracco, F.; Garlaschelli, D.; Gabrielli, A.; Caldarelli, G. The statistical physics of real-world networks. *Nat. Rev. Phys.* **2019**, *1*, 58–71. [[CrossRef](#)]
- Holland, P.W.; Leinhardt, S. An exponential family of probability distributions for directed graphs. *J. Am. Stat. Assoc.* **1981**, *76*, 33–50. [[CrossRef](#)]
- Besag, J. Spatial interaction and the statistical analysis of lattice systems. *J. R. Stat. Soc. Ser. B (Methodol.)* **1974**, *36*, 192–236. [[CrossRef](#)]
- Frank, O.; Strauss, D. Markov graphs. *J. Am. Stat. Assoc.* **1986**, *81*, 832–842. [[CrossRef](#)]
- Contractor, N.S.; Wasserman, S.; Faust, K. Testing multitheoretical, multilevel hypotheses about organizational networks: An analytic framework and empirical example. *Acad. Manag. Rev.* **2006**, *31*, 681–703. [[CrossRef](#)]
- Wasserman, S.; Faust, K. *Social Network Analysis: Methods and Applications*; Cambridge University Press: Cambridge, UK, 1994; Volume 8.
- Carrington, P.J.; Scott, J.; Wasserman, S. *Models and Methods in Social Network Analysis*; Cambridge University Press: Cambridge, UK, 2005; Volume 28.
- Park, J.; Newman, M.E. Statistical mechanics of networks. *Phys. Rev. E* **2004**, *70*, 066117. [[CrossRef](#)]
- Gemmetto, V.; Garlaschelli, D. Multiplexity versus correlation: The role of local constraints in real multiplexes. *Sci. Rep.* **2015**, *5*, 9120. [[CrossRef](#)]
- Gemmetto, V.; Squartini, T.; Picciolo, F.; Ruzzenenti, F.; Garlaschelli, D. Multiplexity and multireciprocity in directed multiplexes. *Phys. Rev. E* **2016**, *94*, 042316. [[CrossRef](#)]

30. Barigozzi, M.; Fagiolo, G.; Garlaschelli, D. Multinetwork of international trade: A commodity-specific analysis. *Phys. Rev. E* **2010**, *81*, 046104. [[CrossRef](#)]
31. Squartini, T.; Fagiolo, G.; Garlaschelli, D. Randomizing world trade. I. A binary network analysis. *Phys. Rev. E* **2011**, *84*, 046117. [[CrossRef](#)]
32. Fagiolo, G.; Squartini, T.; Garlaschelli, D. Null models of economic networks: The case of the world trade web. *J. Econ. Interact. Coord.* **2013**, *8*, 75–107. [[CrossRef](#)]
33. Mastrandrea, R.; Squartini, T.; Fagiolo, G.; Garlaschelli, D. Reconstructing the world trade multiplex: The role of intensive and extensive biases. *Phys. Rev. E* **2014**, *90*, 062804. [[CrossRef](#)]
34. Szell, M.; Lambiotte, R.; Thurner, S. Multirelational organization of large-scale social networks in an online world. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 13636–13641. [[CrossRef](#)] [[PubMed](#)]
35. Cardillo, A.; Gómez-Gardenes, J.; Zanin, M.; Romance, M.; Papo, D.; Del Pozo, F.; Boccaletti, S. Emergence of network features from multiplexity. *Sci. Rep.* **2013**, *3*, 1344. [[CrossRef](#)] [[PubMed](#)]
36. Menichetti, G.; Remondini, D.; Panzarasa, P.; Mondragón, R.J.; Bianconi, G. Weighted multiplex networks. *PLoS ONE* **2014**, *9*, e97857. [[CrossRef](#)] [[PubMed](#)]
37. Freeman, L.C. A set of measures of centrality based on betweenness. *Sociometry* **1977**, *40*, 35–41. [[CrossRef](#)]
38. Clauset, A.; Shalizi, C.R.; Newman, M.E. Power-law distributions in empirical data. *SIAM Rev.* **2009**, *51*, 661–703. [[CrossRef](#)]
39. Barabási, A.L.; Albert, R. Emergence of scaling in random networks. *Science* **1999**, *286*, 509–512. [[CrossRef](#)]
40. Berlingerio, M.; Coscia, M.; Giannotti, F.; Monreale, A.; Pedreschi, D. Foundations of multidimensional network analysis. In Proceedings of the Advances in Social Networks Analysis and Mining (ASONAM), 2011 International Conference, Kaohsiung, Taiwan, 25–27 July 2011; pp. 485–489.
41. Bianconi, G. Statistical mechanics of multiplex networks: Entropy and overlap. *Phys. Rev. E* **2013**, *87*, 062806. [[CrossRef](#)]
42. Jaynes, E.T. Information theory and statistical mechanics. *Phys. Rev.* **1957**, *106*, 620. [[CrossRef](#)]
43. Jaynes, E.T. On the rationale of maximum-entropy methods. *Proc. IEEE* **1982**, *70*, 939–952. [[CrossRef](#)]
44. Garlaschelli, D.; Loffredo, M.I. Multispecies grand-canonical models for networks with reciprocity. *Phys. Rev. E* **2006**, *73*, 015101. [[CrossRef](#)]
45. Newman, M.E.; Girvan, M. Finding and evaluating community structure in networks. *Phys. Rev. E* **2004**, *69*, 026113. [[CrossRef](#)] [[PubMed](#)]
46. Anand, K.; Bianconi, G. Entropy measures for networks: Toward an information theory of complex topologies. *Phys. Rev. E* **2009**, *80*, 045102. [[CrossRef](#)] [[PubMed](#)]
47. Garlaschelli, D.; Loffredo, M.I. Generalized bose-fermi statistics and structural correlations in weighted networks. *Phys. Rev. Lett.* **2009**, *102*, 038701. [[CrossRef](#)]
48. Coolen, T.; Annibale, A.; Roberts, E. *Generating Random Networks and Graphs*; Oxford University Press: Oxford, UK, 2017.
49. Pólya, G.; Szegő, G. *Problems and Theorems in Analysis: Series, Integral Calculus, Theory of Functions*; Aeppli, D., Translator; Springer: Berlin/Heidelberg, Germany, 1972.
50. Park, J.; Newman, M.E. Solution of the two-star model of a network. *Phys. Rev. E* **2004**, *70*, 066146. [[CrossRef](#)]
51. Hastings, W.K. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika* **1970**, *57*, 97–109. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.