

Article

Robustness and Complexity of Directed and Weighted Metabolic Hypergraphs

Pietro Traversa ^{1,2,3,*} , Guilherme Ferraz de Arruda ³ , Alexei Vazquez ⁴  and Yamir Moreno ^{1,2,3} 

¹ Institute for Bio computation and Physics of Complex Systems (BIFI), University of Zaragoza, 50018 Zaragoza, Spain

² Department of Theoretical Physics, University of Zaragoza, 50018 Zaragoza, Spain

³ CENTAI Institute, 10138 Turin, Italy

⁴ Nodes & Links Ltd., Salisbury House, Station Road, Cambridge CB1 2LA, UK

* Correspondence: pietro.traversa98@gmail.com

Abstract: Metabolic networks are probably among the most challenging and important biological networks. Their study provides insight into how biological pathways work and how robust a specific organism is against an environment or therapy. Here, we propose a directed hypergraph with edge-dependent vertex weight as a novel framework to represent metabolic networks. This hypergraph-based representation captures higher-order interactions among metabolites and reactions, as well as the directionalities of reactions and stoichiometric weights, preserving all essential information. Within this framework, we propose the communicability and the search information as metrics to quantify the robustness and complexity of directed hypergraphs. We explore the implications of network directionality on these measures and illustrate a practical example by applying them to a small-scale *E. coli* core model. Additionally, we compare the robustness and the complexity of 30 different models of metabolism, connecting structural and biological properties. Our findings show that antibiotic resistance is associated with high structural robustness, while the complexity can distinguish between eukaryotic and prokaryotic organisms.

Keywords: hypergraphs; complexity; robustness; metabolism; communicability; search information



Citation: Traversa, P.; Ferraz de Arruda, G.; Vazquez, A.; Moreno, Y. Robustness and Complexity of Directed and Weighted Metabolic Hypergraphs. *Entropy* **2023**, *25*, 1537. <https://doi.org/10.3390/e25111537>

Academic Editor: Chaoming Song

Received: 5 October 2023

Revised: 4 November 2023

Accepted: 9 November 2023

Published: 11 November 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

A metabolic network [1–5] is a highly organized system of chemical reactions that occur in living organisms to sustain life and regulate cellular processes. Metabolic networks are incredibly complex because of the large number of reactions and the intricate web of interactions between molecules. Chemical reactions take some metabolites, usually called reactants or substrates, and turn them into products which can be used by other reactions. This complexity allows organisms to perform various functions and respond to various challenges, but it makes understanding them much more challenging. The key functions of metabolism are the production of energy, the conversion of food into building blocks of proteins, lipids, nucleic acids, and carbohydrates, and the elimination of metabolic wastes.

Given the network structure of metabolism, many researchers have attempted to characterize and understand it through network theory. It has been shown that graphs whose nodes are metabolites and are connected by chemical reactions have a scale-free distribution [3] and have been described as “among the most challenging biological networks and, arguably, the ones with most potential for immediate applicability” [6]. Other attempts have tried to give more concrete answers by focusing on graphs with reactions as nodes or bipartite graphs but missing a fundamental aspect of chemical reactions. To take place, they require a collective interaction of reactants to create multiple products. Hence, these are high-order interactions that graphs cannot fully capture. As network theory has advanced, new structures have been devised that can capture high-order interactions. These structures, called hypergraphs, have been very successful in fields such

as social sciences [7–17], epidemiology [12,15,18–21], biology [22–29], etc. The potential of hypergraphs to describe cellular networks has been hypothesized in a perspective in 2009 [30]. Mapping into a hypergraph was also noted in [31–34], bringing attention to this new representation. Recently, Mulas et al. [35,36] applied hypergraphs to chemical networks, trying to capture the high-order nature of chemical reactions. In this paper, we take the concept of chemical hypergraphs and apply it to metabolic networks. In addition, we take it a step further by showing how including weights in the treatment allows no biological or structural information to be lost. Therefore, we argue that metabolic hypergraphs are the right framework to address and understand metabolism, allowing for a bridge between biology and network theory.

This article aims to lay the foundation for a theory of metabolic networks based on hypergraphs. We describe the method by which each metabolic network can be represented as a hypergraph and introduce two applicable measures, namely, communicability and search information.

The work is organized as follows. In Section 2, we give the mathematical definitions regarding metabolic hypergraphs. We also comment on previous studies in the field of metabolic networks and on how they can be viewed as a simplification of the metabolic hypergraph we propose here. In Section 3, we propose a generalization of communicability and search information for hypergraphs. We keep this section general enough so that these measures can be easily applied to any hypergraph, directed or undirected, weighted or not. We use metabolic hypergraphs as an example, and we report the results in Section 4. We conclude by commenting on the possibility that this framework offers of motivating further research in this area.

2. Metabolic Networks as Hypergraphs

In this section, we give a formal definition of metabolic hypergraphs and introduce the notation that is used to characterize them.

2.1. Hypergraphs Definition

A hypergraph $H = \{V, E\}$ is a set of vertices or nodes $v \in V$ and hyperedges $e \in E$. Each hyperedge is a subset of V such that different nodes interact with each other if and only if they belong to the same hyperedge. Thus, unlike traditional graphs, where edges connect pairs of nodes, hyperedges represent interactions involving multiple nodes. If the dimension $|e|$ of the hyperedges is 2, then the hypergraph is equivalent to a conventional graph. The total number of vertices is denoted as $N = |V|$ and the number of hyperedges as $M = |E|$.

To interpret metabolic networks as hypergraphs, we first need to define a special type of hypergraph introduced by Chitra et al. [37]. A hypergraph with edge-dependent vertex weights (EDVW) $H = \{V, E, W, \Gamma\}$ is a set of vertices or nodes $v \in V$, hyperedges $e \in E$, edge weights $w(e)$, and edge-dependent vertex weight $\gamma_e(v)$. If $\gamma_e(v) = \gamma(v) \forall e \in E$, then the hypergraph is said to have edge-independent vertex weight. All the weights are assumed to be positive. These types of weights are a unique property of some higher-order systems and are crucial for encoding in the hypergraph all the information contained in metabolic networks.

In this paper, we deal with directed hypergraphs, which are an extension of directed graphs. In a directed hypergraph, each hyperedge is associated with a direction similar to the direction of an arrow connecting two vertices in a directed graph. In this context, a hyperedge e_j is divided into a head set $H(e_j)$ and a tail set $T(e_j)$. Similarly to the arrow, the direction goes from the tail to the head set, with the difference that the directed hyperedge is connecting multiple vertices. A vertex can belong solely to either the head or the tail of a hyperedge, but not both. Unless explicitly stated otherwise, any hypergraph in this paper is considered to be directed.

Additionally, we define k_v^{out} , the out-degree of a vertex $v \in V$, as the number of hyperedge-tails that include v . Similarly, k_v^{in} denotes the in-degree of a vertex $v \in V$, the

number of hyperedge-heads in which v is contained. We also use $|H(e)|$ and $|T(e)|$ to represent the number of vertices belonging to $H(e)$ and $T(e)$, respectively.

Given a directed hypergraph $H = \{V, E\}$ of N vertices and M hyperedges, the incidence matrix is the matrix $\mathcal{I} \in \mathbb{R}^{N \times M}$ such that:

$$\mathcal{I}_{ij} = \begin{cases} 1 & \text{if } v_i \in H(e_j) \\ -1 & \text{if } v_i \in T(e_j) \\ 0 & \text{if } v_i \notin e_j \end{cases}, \quad (1)$$

where $H(e_j)$ and $T(e_j)$ are, respectively, the heads and the tails of the hyperedges e_j . We can rewrite the incidence matrix as

$$\mathcal{I} = \mathcal{I}_H - \mathcal{I}_T, \quad (2)$$

where we separate the contributions coming from the head (\mathcal{I}_H) and the tail (\mathcal{I}_T) of the hyperedges in order to work with positive signed matrices. It is useful to mathematically define sinks and sources. A source is a node (or hyperedge) that has zero in-degree (or empty tail) and non-zero out-degree (or empty head). A sink is a node (or hyperedge) that has non-zero in-degree (or empty tail) and zero out-degree (or empty head).

2.2. Metabolic Hypergraphs

In this article, we focus on metabolic networks. A metabolic network [2] is a set of biological processes that determines the properties of a cell. Several reactions are involved in metabolism, grouped into various metabolic pathways. A metabolic pathway is an ordered chain of reactions in which metabolites are converted into other metabolites or energy. For example, the glycolysis pathway is the set of reactions involved in the transformation of one molecule of glucose into two molecules of pyruvate, producing energy. Metabolic networks are among the most challenging and highest-potential biological networks [3,6]. The way to represent a metabolic network on a graph is not unique, and several approaches have been tried. One possible way is to consider metabolites (or reactions) as nodes and connect them if and only if they share a reaction (or metabolite). The resulting graph is undirected, and this may change the structural properties of the network in an undesirable way. In [38], the authors analyze the same dataset that we analyze for *E. coli* and propose a directed graph with reactions as nodes that take into account the directionality of the reactions, highlighting the difference with the undirected counterparts.

However, reactions are intrinsically higher-order interactions since they can occur only when all reactants are present. In Figure 1, we illustrate the way to map a chemical reaction network into a hypergraph. The resulting hypergraph is a directed hypergraph with edge-dependent vertex weight, which we will refer to as a metabolic hypergraph for brevity. More formally, we define a metabolic hypergraph as a 3-tuple $H = \{V, E, \mathcal{S}\}$, where $V = \{v_1, v_2, \dots, v_N\}$ is a set of N metabolites (vertices) and E is a set of oriented reactions (hyperedges). Each $e \in E$ is a pair $(T(e), H(e))$, the tail and the head of the hyperedge which correspond, respectively, to the inputs and outputs of the reaction. Note that $T(e)$ or $H(e)$ can also be empty sets. This is the case for external reactions that introduce inside the cell the ingested metabolites (the tail is an empty set) and external reactions that secrete metabolites (the head is an empty set). We also call the former source reactions and the latter sink reactions, and their effect on the measurements is discussed in more detail in Section 3. \mathcal{S} is the stoichiometry matrix associated with the chemical network, and it represents the EDVW of the hypergraph. Indeed, one can notice that \mathcal{S} can be rewritten using the EDVW matrix Γ as $\mathcal{S} = \Gamma \circ \mathcal{I}$, where \mathcal{I} is the directed incidence matrix and " \circ " is the element-wise matrix product.

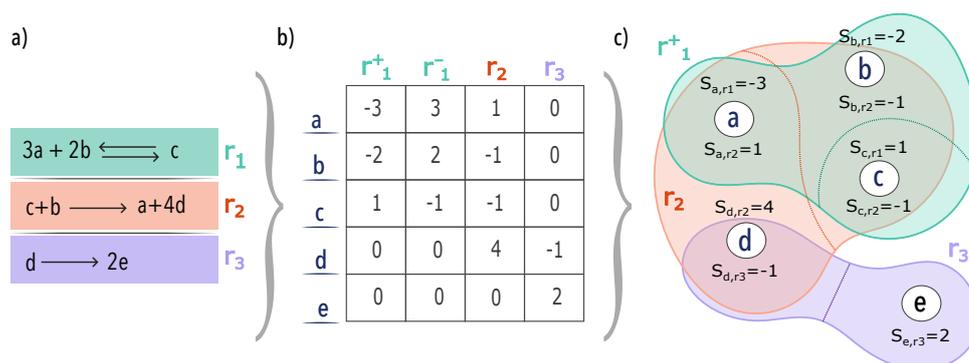


Figure 1. An example of a metabolic network mapped into a hypergraph with edge-dependent vertex weight. In (a), we present a small network composed of three reactions and five metabolites. The first reaction r_1 is reversible and is represented with the double arrow. In (b), we show the corresponding stoichiometry matrix. Reactants are negative and products are positive. Note that we need to split the reversible reaction into two irreversible reactions r_1^+ and r_1^- to write it in matrix form. This stoichiometry matrix is the weighted incidence matrix of the hypergraph with edge-dependent vertex weights shown in (c). For the sake of visualization, only the hyperedge r_1^+ is shown. The hyperedge r_1^- is just the same but with the opposite sign. Note that weights are both positive and negative, meaning that the hypergraph is directed. Indeed, we separate the head and tail of each hyperedge with a dashed line.

2.3. Literature Background

There are different techniques for studying metabolic networks. Popular methods employ kinetic metabolic models [39,40] and stochastic chemical kinetics [41] to study the dynamics of metabolites concentrations in metabolism. While these models are crucial to understanding the complex dynamics of metabolic networks, they require the notion of the kinetic rates constant, the rates at which metabolites are consumed per reaction, which are usually not available [42]. What instead is generally known are the reactions, the stoichiometry coefficient, and the structure of the metabolic network. Thus, several graph representations of metabolic networks have been tried. The most common one is the reaction adjacency matrix (RAG) defined as $A^{\text{RAG}} = \hat{S}^T \hat{S}$ [32,38], where \hat{S} is the boolean version of the stoichiometry matrix. The biggest limitation of this model is that it is undirected, while we know that the direction of reactions is chemically very important. A big improvement was proposed in [38], where the authors proposed a flux-dependent graph model that accounts suitably for the directness of the reactions. However, graph representations of these systems are still missing a crucial point, which is the fact that reactions are higher-order objects which involve the interactions of all input metabolites to produce output metabolites. Therefore, hyperedges are the natural mathematical object for encoding reactions. Mulas et al. [36] already took a step in this direction by defining a Laplace operator for chemical hypergraphs. The last step we make is to incorporate into the hypergraph model the weights associated with metabolites and reactions, using a similar framework to the EDVW defined in [37]. In this last modification to the model, it is crucial to include biological and chemical constraints in the model.

The main advantage of the metabolic hypergraph framework we propose is that it captures all the physical properties that a metabolic network displays: the directness of reactions, the higher-order interactions, and the chemical properties like mass conservation, due to the inclusion of weights. This framework represents a link between network theory and biology. Another commonly employed method for analyzing large-scale metabolic network models is constraint-based metabolic modeling, such as flux balance analysis. Flux balance analysis (FBA) is used to obtain steady-state reaction rates that are consistent with a metabolic network and linear constraints on the reaction rates, without necessitating any knowledge about the kinetic parameters [43]. FBA is a method of finding steady-state

solutions [44,45], yet, one needs to perform additional analyses to determine the relevance of each reaction or metabolite to the solutions obtained. For this scope, hypergraph theory provides a lot of tools that could be used alongside FBA.

We remark that the previous graph representation of metabolic networks can be seen as a pairwise projection of a metabolic hypergraph. For example, the RAG is an undirected projection of the hypergraph, as in [46], and the flux-dependent graph [38] is similar to the normalized adjacency matrix defined in [47] but extends to directed and weighted hypergraphs. Projections are a pairwise simplification and can perform well depending on the task, but they do not contain all the information. For example, in [32], the authors start with a hypergraph formalism and project it to a reaction adjacency to evaluate the number of extreme pathways in metabolic networks.

2.4. Dataset

In our experiments, metabolic hypergraphs are generated from the stoichiometry matrix of the models stored in the BiGG database [48]. We analyze 30 different models, with an increasing number of nodes describing different organisms (see Table A1 in Appendix A for the exact number of nodes and reactions of each BiGG model). We chose the metabolic networks in order to have a reasonable variety of organisms, and we avoided very large networks because of the computational costs. The majority of the data are composed of bacteria that can be divided into classes like antibiotic-resistant, aerobic or anaerobic, Gram-positive or Gram-negative. The other organisms are eukaryotes, and one is in the Archaea domain. All data are publicly available on the BiGG models web page [49] in different formats. In this analysis, the *.json* format is used. The data contain information on metabolites, reactions, and genes. Metabolites are identified by a BiGG ID, consisting of an abbreviation defining their type, for example, “h” for hydrogen and “ATP” for adenosine triphosphate, and a subscript indicating the compartment to which they belong. Regarding the reactions, in addition to their IDs, the metabolites belonging to them are given, with their respective stoichiometric coefficients. We work in the convention in which a metabolite with a positive stoichiometric coefficient is a product; otherwise, it is a reactant. In the BiGG dataset, the direction of the reactions is also determined using the parameters “lower_bound” and “upper_bound”. These parameters are associated with each reaction and correspond to the maximal flux of metabolites that can flow through. Values of lower_bound = 0 and upper_bound > 0 mean that the reaction is annotated correctly, following the convention. On the contrary, if lower_bound < 0 and upper_bound = 0, the reactions are written with inverted orientations. These two parameters combined also determine if a reaction is reversible or not. If a reaction is reversible, both the direct and inverse reactions are present and will be characterized by a lower_bound < 0 and upper_bound > 0. We recall that we treat reversible reactions as two distinct hyperedges, see Figure 1 for a visual example. It is important to notice that few reactions have lower_bound = 0 and upper_bound = 0. In practice, this implies that no flux of metabolites can flow through, so those reactions are discarded. The origin of the reaction bounds depends on the BiGG models considered. For example, both models for *Mycobacterium tuberculosis H37Rv* have some reactions with lower_bound = 0 and upper_bound = 0 identified via flux variability analysis (FVA). In the case of the *Synechococcus elongatus PCC 7942*, the bounds are obtained experimentally. We decide to proceed with this convention, but using relaxed reaction boundaries to include these reactions is also a valid option.

All the metabolites present in the BiGG models were kept; we did not discard dead-end metabolites.

Lastly, we highlight that some hyperedges may have an empty tail or head. These hyperedges correspond to reactions involved in the transportation of metabolites from the outside of the cell to the inside or vice versa. For example, EX_h2o_e (H₂O exchange) is the reaction that takes the water from the environment and brings it into the metabolism. The metabolites outside the metabolism are not present in the BiGG models, and for this reason, the reaction appears as “ → h2o_e”, with an empty tail. Therefore, sometimes they may

represent sinks and sources in the hypergraph. By source, we mean a node or hyperedge from which you can start and leave but never go back, while a sink is a trapping node or hyperedge that, if it is reached, is impossible to leave.

3. Measurements

In this section, we define two measures of the chemical hypergraph based on the notion of paths or walks on hypergraphs. A *walk* of length l from node v_0 to node v_l is defined as a sequence of alternating nodes and hyperedges $(v_0, e_1, v_1, e_2, v_2, \dots, e_l, v_l)$. We also define the *dual walk* from hyperedge e_0 to hyperedge e_l of length l as the alternating sequence of alternating nodes and hyperedges $(e_0, v_1, e_1, v_2, e_2, \dots, v_l, e_l)$. We are interested in both metabolites and reactions, which is why it is useful also to consider the dual walk.

3.1. Hypergraph Communicability

We are usually interested in understanding how paths are distributed because that is how information and interactions spread. In social systems, for example, the more paths connecting two nodes, the easier is for information to spread from one to another. Also, if one path of connection fails, the information can still be spread through other paths, even if they are longer than the path that failed. For this reason, the notion of paths and communication between nodes can also be related to the robustness of the network. However, having a robust network is not always positive. The same reasoning about the spreading of information applies to the spreading of viruses. If a network is robust, it is way more difficult to design containment strategies for the virus, since shutting down a connection might not be enough because of the presence of alternative paths. A way to measure how nodes communicate within a network is called communicability, and we extend this definition to hypergraphs.

The communicability [50,51] between two pairs of nodes p and q is defined as the weighted sum of all walks starting from node p and ending at node q , as in

$$G_{pq} = \sum_{k=0}^{\infty} c_k n_{pq}^k \quad (3)$$

where n_{pq}^k is the number of walks from p to q and c_k is the penalization for long paths. The most common choice is $c_k = \frac{1}{k!}$ so that you recover an exponential expansion. For a graph, n_{pq}^k can be easily found by taking the k -power of the adjacency matrix, $(A^k)_{pq}$. Hypergraphs do not have a unique definition of adjacency matrix; we thus have to use the definition of walk given above. The vertex-to-vertex communicability for a hypergraph with incidence matrix \mathcal{I} is defined as

$$G_{pq}^V = \sum_{k=0}^{\infty} \frac{\left((\mathcal{I}_T \mathcal{I}_H^t)^k \right)_{pq}}{k!}, \quad (4)$$

or, in matrix form,

$$G^V = e^{\mathcal{I}_T \mathcal{I}_H^t}, \quad (5)$$

where $(\cdot)^t$ indicates the transpose of the matrix. In metabolic hypergraphs, we are also interested in how reactions communicate with each other. For this reason, we define the hyperedge-to-hyperedge communicability based on the notion of dual path

$$G_{pq}^E = \sum_{k=0}^{\infty} \frac{\left((\mathcal{I}_H^t \mathcal{I}_T)^k \right)_{pq}}{k!}, \quad (6)$$

or, in matrix form,

$$G^E = e^{\mathcal{I}_H^t \mathcal{I}_T}. \quad (7)$$

The Estrada index [50,52] of a hypergraph H is generalized as

$$\begin{aligned} EE^V(H) &= \text{Trace}(G^V), \\ EE^E(H) &= \text{Trace}(G^E). \end{aligned} \quad (8)$$

One can notice that the matrices $\mathcal{I}_T \mathcal{I}_H^t$ and $\mathcal{I}_H^t \mathcal{I}_T$ have the same spectrum except for the number of zero eigenvalues because of the difference in size. This means that for $M > N$, for example (which is usually the case in metabolic hypergraphs), the Estrada index defined on nodes and the one defined on the hyperedges are related by $EE^E(H) = EE^V(H) + (M - N)$. We use the Estrada index defined on the nodes to measure the hypergraph robustness, also known as natural connectivity, as

$$\bar{\lambda}^V = \log \left(\frac{EE(H)^V}{N} \right). \quad (9)$$

The same definition holds for $\bar{\lambda}^E$ with the proper normalization.

Since computing the exponential of very large matrices might be a difficult numerical task, we use an approximation for the calculation of the robustness based on eigenvalue decomposition. For simplicity, let us call $A^V = \mathcal{I}_T \mathcal{I}_H^t$ (the same reasoning holds for $A^E = \mathcal{I}_H^t \mathcal{I}_T$) and order the spectrum of A^V in such a way that $\lambda_1 > \lambda_2 > \lambda_3 > \dots \lambda_N$. Then, the natural connectivity or robustness of the hypergraph becomes

$$\begin{aligned} \bar{\lambda}^V &= \log \left(\sum_{i=1}^N e^{\lambda_i} \right) - \log(N) \\ &= \log \left[e^{\lambda_1} \left(1 + \sum_{i=2}^N e^{\lambda_i - \lambda_1} \right) \right] - \log(N) \\ &= \lambda_1 + \log \left(1 + \sum_{i=2}^N e^{\lambda_i - \lambda_1} \right) - \log(N) \\ &= \lambda_1 - \log(N) + \mathcal{O} \left(e^{-(\lambda_1 - \lambda_2)} \right). \end{aligned}$$

Thus, if the spectral gap is large enough, the natural connectivity is dominated by the largest eigenvalue. Since the correction is exponential, this approximation is usually quite good. As a consequence of the common spectrum of $\mathcal{I}_H^t \mathcal{I}_T$ and $\mathcal{I}_T \mathcal{I}_H^t$, the difference in robustness is approximately $\bar{\lambda}^V - \bar{\lambda}^E \approx \log \left(\frac{M}{N} \right)$, which is usually quite small. It is worth noting that the largest eigenvalue scales with the system size, i.e., with the number of nodes and hyperedges. The normalization factor $-\log(N)$ mitigates this scaling effect, but a partial correlation is still expected. This correlation was present in the original graph definition of natural connectivity [53]. This could be a problem when comparing systems with very different scales. For this reason, in this paper, we are comparing hypergraphs with a similar system size.

This generalization of communicability applies also to undirected hypergraphs by substituting \mathcal{I}_H and \mathcal{I}_T with the undirected incidence matrix I .

3.2. Hypergraph Search Information

Rosvall et al. [54,55] introduced the concept of search information, as a measure of complexity in urban graphs. The idea is to measure the number of binary questions one has to make in order to locate the shortest path connecting a node s to a node t . As a consequence, this measure is based on walks like the communicability, but with the crucial difference that it considers only the shortest paths. This allows us to link the search information with the notion of complexity. While alternative pathways tend to make the network more robust, they also make the probability of finding the shortest path decrease

and the complexity increase. This trade-off is the reason that motivated us to consider communicability and search information together.

In [54], the search information is defined as a matrix S with entries

$$S(i, j)^V = -\log_2 \left(\sum_{\{p(i, j)\}} P(p(i, j)) \right), \quad (10)$$

where $\{p(v_i, v_j)\}$ is the set of all shortest paths from node v_i to node v_j .

The original definition was made for undirected and unweighted ordinary graphs, so a very different structure from directed hypergraphs with edge-dependent vertex weight, but the meaning remains the same. What changes is the probability of following the shortest path. The probability of making a step is proportional to the stoichiometric coefficients of the starting and arriving nodes, similar to what has been performed in the normalized flow graph in [38]. The probability of taking a step in a directed hypergraph with EDVW is

$$\begin{aligned} P(v \rightarrow e) &= \frac{\gamma_e(v)}{\sum_h \gamma_h(v)}, \\ P(e \rightarrow v) &= \frac{\gamma_e(v)}{\sum_n \gamma_e(n)}. \end{aligned} \quad (11)$$

The probability of following a path is derived via multiplication of the single-step probability,

$$P(v_0, v_l) = P(v_0 \rightarrow e_1)P(e_1 \rightarrow v_1) \dots P(e_l \rightarrow v_l). \quad (12)$$

It is important to note that the search information might be ill defined if the hypergraph has sources or sinks. For example, by definition, there are no paths from a sink node v_{sink} to any other nodes v , making the definition of $S(v_{\text{sink}}, v)$ unclear in this case. What we do to solve the problem is to set $S(v_{\text{sink}}, v) = 0$ and then not count sink and source nodes when computing the average. With this convention, the access, hide, and average search information are defined as

$$\begin{aligned} A^V(s) &= \frac{1}{N - N_{\text{sources}}} \sum_t S^V(s, t) \\ H^V(t) &= \frac{1}{N - N_{\text{sinks}}} \sum_s S^V(s, t) \\ \bar{S}^V &= \frac{1}{(N - N_{\text{sinks}})(N - N_{\text{sources}})} \sum_{s, t} S^V(s, t). \end{aligned} \quad (13)$$

As a consequence, the access information of a sink and the hide information of a source will be set to zero. Following [54], we introduce an additional normalization factor $\log_2 N$ to take into account size effects. With this additional term, we did not observe any correlation between the average search information and the number of metabolites or reactions. We denote the normalized average search information as $\sigma^V = \frac{\bar{S}^V}{\log_2 N}$. The interpretation of these measures is very intuitive. The access information measures how easy it is to reach the other nodes in the network, while the hide information estimates how hidden a node is. Consequently, very central and connected nodes in the hypergraph have low hide information because there are a lot of paths leading to them, but they have relatively high access information because there are also many paths departing from such nodes.

4. Results and Discussion

In this section, we apply the previously defined metrics to a range of metabolic hypergraphs. As illustrated in Figure 1, these hypergraphs were constructed by starting with metabolic networks obtained from the BiGG dataset [48]. The metabolic networks

were selected to have a reasonable variety of organisms. The primary goal of this section is to demonstrate the practical application of our framework and the defined measurements.

4.1. Exploring the *E. coli* Core Model: A Practical Example

To provide a tangible illustration of our methodology, we focus on the BiGG model known as *e_coli_core* [56]. This model represents a small-scale version of *Escherichia coli* str. K-12 substr. MG1655, making it an ideal candidate for demonstrating the performance of our metrics and understanding their limitations. Additionally, an Escher map for this model is available online [57].

In Figure 2, we show the access vs. hide information for reactions and metabolites. Regarding the reactions (Figure 2a), the measure correctly identifies the Biomass reaction as a central hub. Reactions are plotted with different colors based on the biological pathway they belong to. We can clearly see the behavior of sinks and sources in the reactions belonging to the extracellular exchange pathway. The pathways do not tend to separate into clusters, indicating that they all have a similar complexity. This could be an effect of the simplicity of this model or could be a property shared by all organisms. We did not investigate further since the scope of this section was just to provide a practical example, but it could be worth it to explore it in future work.

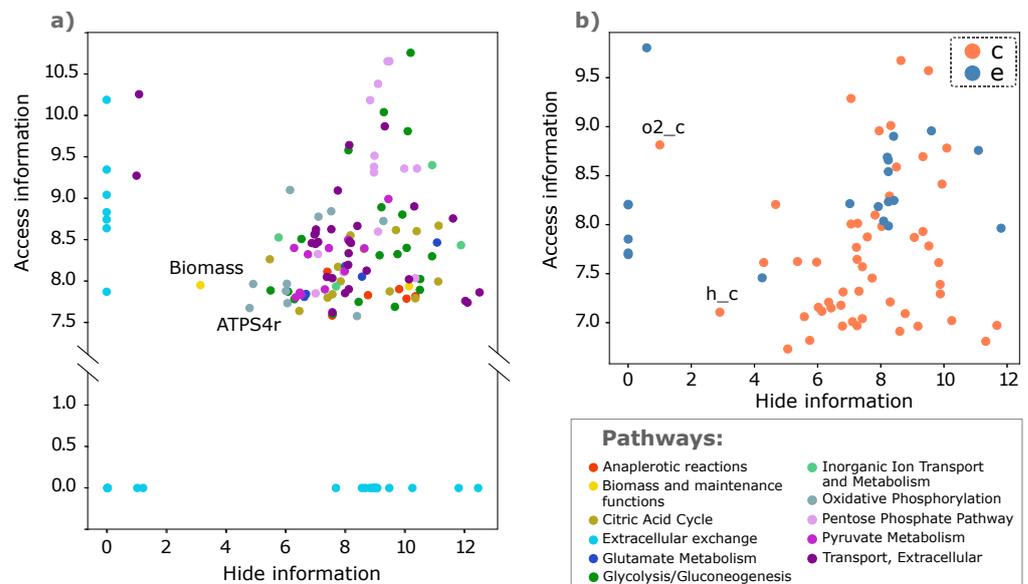


Figure 2. Access vs. hide information for reactions (a) and metabolites (b). Reactions are colored differently according to the pathway they belong to. Note that the y axis is cut for visualization purposes. Metabolites are divided into compartments; c stands for cytosol compartment and e for extracellular space.

We also comment on the reactions that are ranked the highest according to average communicability. The average communicability is defined as $\bar{G}_e = \frac{1}{M} \sum_{h \in E} G_{he}^E$ and is shown in Figure 3. Notably, the Biomass reaction (first highest average communicability) and ATP synthase (second highest average communicability) are correctly identified as central reactions within the metabolism. The Biomass reaction is responsible for cell growth, while ATP synthase plays a crucial role in ATP synthesis, the primary energy source for the organism. The production of ATP is mainly due to the consumption of oxygen that occurs through the reaction CYTBD (cytochrome oxidase bd—sixth highest average communicability). When oxygen is unavailable, *Escherichia coli* can still survive due to the activation of the anaerobic pathway, which derives energy from the reaction THD2 (NAD(P) transhydrogenase—third highest average communicability).

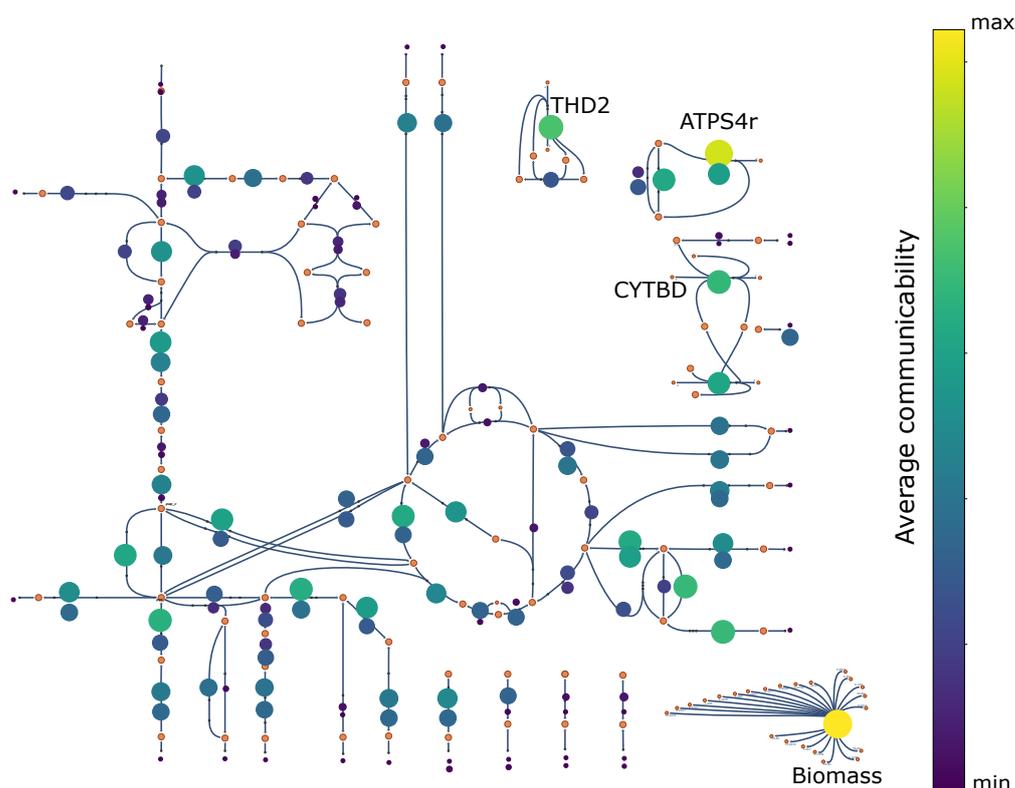


Figure 3. Reactions' average communicability for the *e_coli_core* model. A simplified Escher map is used as a background to help with the visualization. For a more accurate version of the map, visit [57].

Regarding the metabolites (Figure 2b), we observe a clear distinction between those belonging to the cytosol compartment and those located in the extracellular compartment. As expected, extracellular metabolites tend to have, on average, higher hide information. It is important to clarify that metabolites with zero hide information are those that remain initialized to zero because they are unreachable. However, an instructive observation could be made on *o2_c*. As commented in Section 3, a node with low but non-zero hide information is expected to be a central hub, but in reality, it has a very low degree. The explanation for this helps us to understand the implications of network directionality. The node *o2_c* is only connected to the core metabolism via the irreversible CYTBD reaction as a substrate. Consequently, there cannot be any directed path from the core metabolism to *o2_c*, only the opposite. We conclude that the node *o2_c* does not belong to the largest strongly connected component. In practice, it behaves very similarly to a source node. Nonetheless, the hide information is not zero because a pathway originates from the transport of external oxygen to the cytosol. In contrast, in cyanobacteria, algae, and plants (not investigated here), O_2 is produced via oxygenic photosynthesis. In those organisms, O_2 should be part of the strongly connected component.

4.2. Robustness and Complexity across Organisms

Our study assesses the robustness and complexity of 30 distinct metabolic hypergraphs derived from various eukaryotic and prokaryotic organisms. We selected the models to ensure a good range of diversity while avoiding having too many models for a single organism. For example, *Escherichia coli* has over 50 BiGG models. Analyzing all of them could be intriguing as well, but our primary interest lies in comparing organisms rather than models. Additionally, it is worth noting that certain BiGG models exhibit very large metabolisms, featuring thousands of metabolites and reactions. While these large models do hold potential relevance within the scope of our paper, the significant size

of the corresponding metabolic networks renders the computational cost of the search information high. It is possible to study a large metabolic network individually, but the cost of comparing many together is prohibitive. To maintain computational tractability, we restrict our analysis to metabolic networks with no more than 2000 nodes or reactions.

Assessing the robustness of metabolic networks is an important task, and many definitions exist [31,34]. Here, we use natural connectivity to evaluate the network structure’s robustness. In Figure 4, we present the computed robustness values for several organisms arranged in ascending order. The BiGG models associated with the organisms *Staphylococcus aureus subsp aureus* [58,59], *Mycobacterium tuberculosis* [60,61], *Acinetobacter baumannii* AYE [62], and *Salmonella enterica* [63] are represented in different colors because they are bacteria that have evolved resistance to antibiotics. Except for the first *Staphylococcus aureus subsp aureus* model, antibiotic-resistant bacteria tend to exhibit relatively high robustness compared to other organisms. We measured the Spearman’s rank correlation between robustness and antibiotic resistance, obtaining a value of 0.424, revealing a moderate correlation. Here, the definition of robustness is based on the network’s resilience to random or targeted node removal. The concept of natural connectivity quantifies this resilience by counting the number of closed loops in the network. If there are many alternative paths, it is less probable that node removal will disconnect the network. In the context of biology, antibiotics operate by targeting and inhibiting some specific reactions, without which the cell dies [1]. Therefore, having a structurally robust metabolism is advantageous as it allows the organism to circumvent antibiotic inhibition by utilizing alternative reactions or pathways. However, this is not the whole picture since many other factors play a role. For example, bacteria are naturally subjected to random mutations that may strengthen their response to antibiotics, and this may not necessarily be reflected in a high structural hypergraph robustness. Conversely, a very robust metabolic hypergraph, with many alternative paths, may have a few but very important reactions that are easy to target with antibiotics. Hence, high structural hypergraph robustness does not guarantee antibiotic resistance.

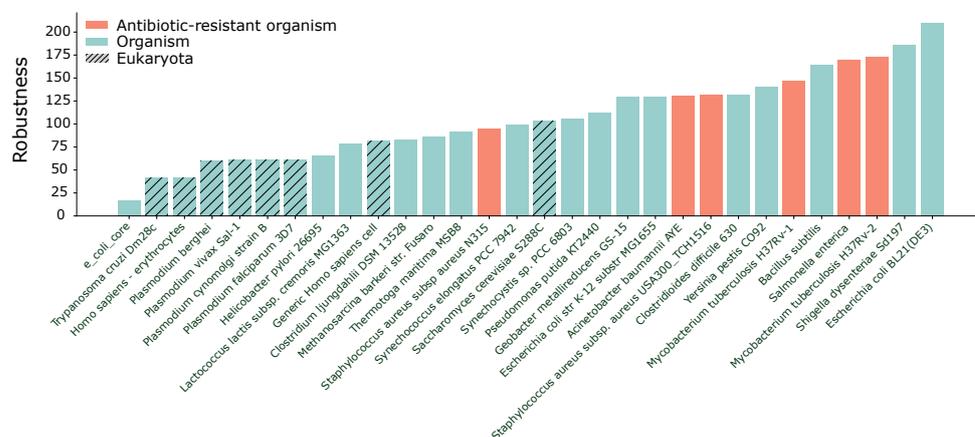


Figure 4. The robustness measured as the natural connectivity λ^V of 30 different BiGG models. The organisms resistant to antibiotics are shown in different colors. The models are ordered with increasing robustness.

The complexity of metabolic networks is anticipated to be quite similar across organisms since they share many common reactions and metabolic pathways. Nevertheless, some differences are expected in the metabolism of aerobic and anaerobic organisms, as well as between eukaryotes and prokaryotes. Aerobic and anaerobic organisms should have a different metabolism because of the different ways they produce energy, while eukaryotes and prokaryotes have significantly different cell structures. With this in mind, we measure the average search information of the 30 different metabolic hypergraphs and report the results in Figure 5. We notice a clear separation between eukaryotes and some aerobic organisms, showing a high complexity, and prokaryotes, which have a lower

complexity. A few outliers exist, including *Staphylococcus aureus subsp aureus* N315, which exhibits high complexity, potentially due to unusually large weights associated with certain reactions compared to other organisms. Setting all the weights to 1 would indeed lead to a much lower complexity, ranked slightly below the average, indicating a possible bias. In addition, one can also notice that the other model for *Staphylococcus* has a low complexity. Another outlier is the first model we analyzed for *Homo sapiens—erythrocytes* [64] that may be expected to be complex. However, it is important to note that this model refers just to the erythrocyte metabolism (blood cells) rather than the entire human metabolism. Erythrocytes lack mitochondria and produce ATP through anaerobic glycolysis, so their metabolism could be closer to that of anaerobic organisms. Conversely, the low complexity of the aerobic organisms *Acinetobacter baumannii* AYE, *Pseudomonas putida*, and *Helicobacter pylori* is curious, and we do not have a clear motivation. Note that a generic human (*Homo sapiens*) cell has a similar complexity to a yeast cell (*Saccharomyces cerevisiae*). That is expected. Eukaryote cells have similar metabolic pathways. The additional complexity in human metabolism is due to multi-cellularity, which is not accounted for in this study.

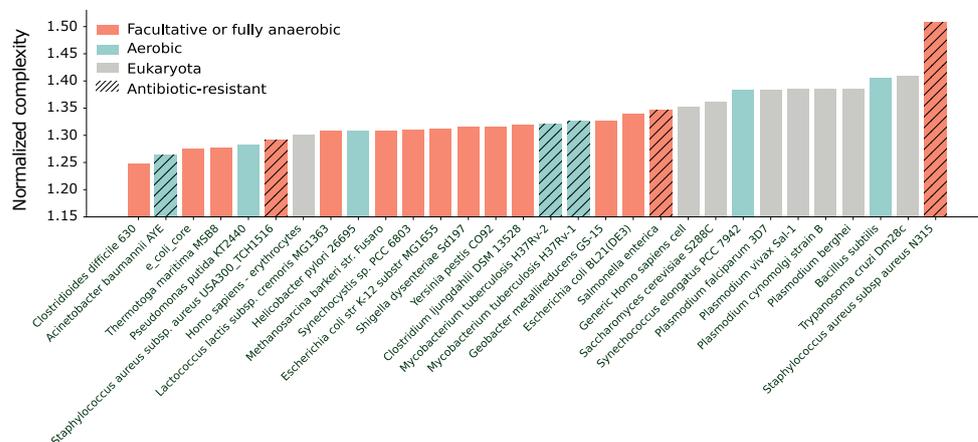


Figure 5. The complexity measured as the average search information $\sigma^V = \frac{S^V}{\log_2 N}$ of 30 different BiGG models. The models are ordered with increasing complexity, and the y axis is zoomed in for visualization purposes.

5. Conclusions

Metabolic networks are very large and complex systems. For this reason, it is important to build a framework able to unite biology and network theory. Many successful studies have represented metabolic networks as graphs with metabolites as nodes, reactions as nodes, or both. Taking a step further, with the employment of hypergraphs, we are able to capture what all of these previous graph representations were missing, the higher-order interactions of reactions. In this paper, we show how metabolic networks are naturally mapped into hypergraphs. In particular, the stoichiometry matrix can be viewed as a weighted incidence matrix of a directed hypergraph with edge-dependent vertex weight. No information is lost representing metabolic networks as hypergraphs: the higher-order interactions between metabolites, the directionalities of reactions, and the stoichiometric weights are all included.

Within this novel framework, we propose two measurements to characterize a hypergraph's robustness and complexity. We apply them to directed hypergraphs with EDVW, but the generalization to undirected and unweighted hypergraphs is straightforward. This approach allows for analysis at the local scale, with communicability and access and hide information, and at the global scale, with natural connectivity as a measure of robustness and average search information as a measure of complexity. We comment on the complications introduced by directionality and how they can be reflected in the measures. To illustrate the practical application of our framework and metrics, we present an example using the *e_coli_core* model. This small-scale metabolism demonstrates how our metrics

operate locally, and it offers valuable insights into the behavior of metabolic hypergraphs. At the global scale, we compare 30 different BiGG models in robustness and complexity, leading to some interesting results. We show that the metabolisms of organisms that have evolved resistance to antibiotics are associated with hypergraphs that display high robustness. Furthermore, we observe that eukaryotic and prokaryotic organisms have different complexity values.

In our analysis of complexity, we excluded the source and sink reactions because they create problems when computing the search information (they are unreachable hyperedges). Another possibility could be to add a boundary node, representing the environment around the cell, that links the sinks with the sources. In this way, the search information is no longer ill defined and the external reactions could be included in the analysis. However, the introduction of such externally may have undesired effects on the measures, like introducing new and biologically unmotivated shortest paths. It is worth mentioning that an additional boundary node could be crucial when incorporating hypergraph dynamics in the model. A possibility for future works could be modifying the definition of the average search information and the probability of taking a step in the hypergraph. Here, we consider a walk biased by the stoichiometric weights, but more options could be explored. One possibility is to define the probabilities based on the communicability measure or on the rates computed via flux balance analysis [38,43]. Indeed, from flux balance analysis, we obtain rates that could be interpreted as edge-dependent vertex weights, substituting the stoichiometric coefficients. The union of FBA with hypergraph theory, to the best of our knowledge, has not been studied yet and could be an original contribution to the field. Also, we did not consider the information regarding genes that are contained in the BiGG models. Genomics plays a crucial role, especially in resistance to antibiotics, and for this reason, it could be interesting to integrate it into this framework. Another possibility is to apply our measures to other contexts, like social or technological hypergraphs.

We believe that this framework represents a promising approach to bridging network theory and biology. We hope that it may serve as a starting point, potentially reaching experts in the field who could further refine and utilize these findings to obtain more biological insights.

Author Contributions: Conceptualization, G.F.d.A. and Y.M.; Methodology, P.T.; Investigation, P.T., G.F.d.A., A.V. and Y.M.; Writing—original draft, P.T.; Writing—review & editing, P.T., G.F.d.A., A.V. and Y.M.; Supervision, G.F.d.A. and Y.M.; Project administration, Y.M. All authors have read and agreed to the published version of the manuscript.

Funding: P.T., G.F.A. and Y.M. acknowledge the financial support of Soremartec S.A. and Soremartec Italia, Ferrero Group. Y.M. acknowledges partial support from the Government of Aragon and FEDER funds, Spain through grant E36-20R (FENOL), and the EU program Horizon 2020/H2020-SCI-FA-DTS-2020-1 (KATY project, contract number 101017453). We acknowledge the use of the computational resources of COSNET Lab at Institute BIFI, funded by Banco Santander (grant Santander-UZ 2020/0274) and by the Government of Aragón (grant UZ-164255).

Institutional Review Board Statement: Not applicable.

Data Availability Statement: All data are publicly available on the BiGG models [49] web page in different formats. In this analysis, the *.json* format is used.

Conflicts of Interest: The funders had no role in study design, data collection, analysis, decision to publish, or preparation of the manuscript. Nodes & Links Ltd provided support in the form of salary for Alexei Vazquez, but did not have any additional role in the conceptualization of the study, analysis, decision to publish, or preparation of the manuscript.

Appendix A. BiGG Models

In Table A1, we provide the number of nodes, reactions, and hyperedges for each analyzed hypergraph. We also report the associated BiGG model ID to facilitate its identification, reproduction, and further studies. For more information, see the BiGG models webpage [49].

Table A1. BiGG models and their numbers of metabolites and reactions. The directed hypergraph constructed from the model has a number of nodes equal to the number of metabolites and a number of hyperedges bigger than the number of reactions because of the presence of reversible reactions.

Organism	BiGG Model	Metabolites	Reactions	Hyperedges
Saccharomyces cerevisiae S288C	iND750	1059	1266	1702
Pseudomonas putida KT2440	iJN746	907	1054	1415
Plasmodium cynomolgi strain B	iAM_Pc455	907	1074	1563
e_coli_core	e_coli_core	72	95	141
Staphylococcus aureus subsp. aureus USA300_TCH1516	iYS854	1335	1453	1872
Mycobacterium tuberculosis H37Rv-1	iNJ661	825	1022	1293
Mycobacterium tuberculosis H37Rv-2	iEK1008	998	1224	1500
Clostridium ljungdahlii DSM 13528	iHN637	698	773	988
Yersinia pestis CO92	iPC815	1552	1960	2507
Shigella dysenteriae Sd197	iSDY_1059	1888	2529	3172
Escherichia coli str. K-12 substr. MG1655	iJR904	761	1075	1329
Lactococcus lactis subsp. cremoris MG1363	iNF517	650	730	979
Helicobacter pylori 26695	iIT341	485	554	737
Homo sapiens	iAB_RBC_283	342	469	645
Homo sapiens2	iAT_PLT_636	738	1008	1455
Plasmodium falciparum 3D7	iAM_Pf480	909	1083	1576
Escherichia coli BL21(DE3)	iEC1356_BI21DE3	1918	2730	3376
Synechococcus elongatus PCC 7942	iJB785	768	843	1064
Plasmodium berghei	iAM_Pb448	903	1067	1554
Trypanosoma cruzi Dm28c	iIS312	606	519	806
Staphylococcus aureus subsp aureus N315	iSB619	655	729	945
Thermotoga maritima MSB8	iLJ478	570	652	852
Methanosarcina barkeri str. Fusaro	iAF692	628	690	900
Clostridioides difficile 630	iCN900	885	1222	1455
Plasmodium vivax Sal-1	iAM_Pv461	909	1078	1570
Bacillus subtilis	iYO844	990	1250	1589
Synechocystis sp. PCC 6803	iJN678	795	862	1086
Geobacter metallireducens GS-15	iAF987	1109	1281	1642
Acinetobacter baumannii AYE	iCN718	888	1013	1436
Salmonella enterica	STM_v1_0	1802	2528	3133

References

1. Campbell, N.A.; Urry, L.A.; Cain, M.L.; Wasserman, S.A.; Minorsky, P.V.; Reece, J.B. *Biology: A Global Approach, Eleventh Edition, Global Edition*; Pearson: New York, NY, USA, 2018; ISBN 9781292170435.
2. Dräger, A.; Planatscher, H. *Encyclopedia of Systems Biology*; Dubitzky, W., Wolkenhauer, O., Cho, K.-H., Yokota, H., Eds.; Springer: New York, NY, USA, 2013; pp. 1249–1251, ISBN 9781441998620/9781441998637.
3. Ma, H.; Zeng, A.-P. Reconstruction of metabolic networks from genome data and analysis of their global structure for various organisms. *Bioinformatics* **2003**, *19*, 270–277. [[CrossRef](#)] [[PubMed](#)]

4. Stitt, M.; Sulpice, R.; Keurentjes, J. Metabolic networks: How to identify key components in the regulation of metabolism and growth. *Plant Physiol.* **2010**, *152*, 428–444. [[CrossRef](#)] [[PubMed](#)]
5. Renz, A.; Mostolizadeh, R.; Dräger, A. *Systems Medicine*; Wolkenhauer, O., Ed.; Academic Press: Oxford, UK, 2021; pp. 362–371, ISBN 978-0-12-816078-7.
6. Guimerà, R.; Nunes Amaral, L.A. Functional cartography of complex metabolic networks. *Nature* **2005**, *433*, 895–900. [[CrossRef](#)]
7. Niu, X.; Doyle, C.; Korniss, G.; Szymanski, B.K. The impact of variable commitment in the naming game on consensus formation. *Sci. Rep.* **2017**, *7*, 41750. [[CrossRef](#)]
8. Centola, D.; Becker, J.; Brackbill, D.; Baronchelli, A. Experimental evidence for tip** points in social convention. *Science* **2018**, *360*, 1116–1119. [[CrossRef](#)] [[PubMed](#)]
9. Baronchelli, A. The emergence of consensus: A primer. *R. Soc. Open Sci.* **2018**, *5*, 172189 [[CrossRef](#)] [[PubMed](#)]
10. Benson, A.R.; Abebe, R.; Schaub, M.T.; Jadbabaie, A.; Kleinberg, J. Simplicial closure and higher-order link prediction. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, E11221–E11230. [[CrossRef](#)]
11. Iacopini, I.; Petri, G.; Barrat, A.; Latora, V. Simplicial models of social contagion. *Nat. Commun.* **2019**, *10*, 2485. [[CrossRef](#)]
12. de Arruda, G.F.; Petri, G.; Moreno, Y. Social contagion models on hypergraphs. *Phys. Rev. Res.* **2020**, *2*, 023032. [[CrossRef](#)]
13. Landry, N.W.; Restrepo, J.G. The effect of heterogeneity on hypergraph contagion models. *Chaos: An Interdisciplinary. J. Nonlinear Sci.* **2020**, *30*, 103117.
14. Barrat, A.; Ferraz de Arruda, G.; Iacopini, I.; Moreno, Y. Social contagion on higher-order structures. In *Higher-Order Systems*; Springer: Cham, Switzerland, 2022; pp. 329–346.
15. Ferraz de Arruda, G.; Tizzani, M.; Moreno, Y. Phase transitions and stability of dynamical processes on hypergraphs. *Commun. Phys.* **2021**, *4*, 24. [[CrossRef](#)]
16. Alvarez-Rodriguez, U.; Battiston, F.; de Arruda, G.F.; Moreno, Y.; Perc, M.; Latora, V. Evolutionary dynamics of higher-order interactions in social networks. *Nat. Hum. Behav.* **2021**, *5*, 586–595 [[CrossRef](#)] [[PubMed](#)]
17. Neuhäuser, L.; Mellor, A.; Lambiotte, R. Multibody interactions and nonlinear consensus dynamics on networked systems. *Phys. Rev. E* **2020**, *101*, 032310. [[CrossRef](#)] [[PubMed](#)]
18. Bodó, Á; Katona, G.Y.; Simon, P.L. SIS epidemic propagation on hypergraphs. *Bull. Math. Biol.* **2016**, *78*, 713–735. [[CrossRef](#)] [[PubMed](#)]
19. Ferraz De Arruda, G.; Petri, G.; Rodriguez, P.M.; Moreno, Y. Multistability, intermittency, and hybrid transitions in social contagion models on hypergraphs. *Nat. Commun.* **2023**, *14*, 1375. [[CrossRef](#)]
20. Higham, D.J.; de Kergorlay, H.L. Mean field analysis of hypergraph contagion models. *SIAM J. Appl. Math.* **2022**, *82*, 1987–2007. [[CrossRef](#)]
21. Higham, D.J.; De Kergorlay, H.L. Epidemics on hypergraphs: Spectral thresholds for extinction. *Proc. R. Soc. A* **2021**, *477*, 20210232. [[CrossRef](#)]
22. Stewart, I.; Golubitsky, M.; Pivato, M. Symmetry groupoids and patterns of synchrony in coupled cell networks. *SIAM J. Appl. Dyn. Syst.* **2003**, *2*, 609–646. [[CrossRef](#)]
23. Golubitsky, M.; Stewart, I.; Török, A. Patterns of synchrony in coupled cell networks with multiple arrows. *SIAM J. Appl. Dyn. Syst.* **2005**, *4*, 78–100. [[CrossRef](#)]
24. Golubitsky, M.; Stewart, I. Nonlinear dynamics of networks: The groupoid formalism. *Bull. Am. Math. Soc.* **2006**, *43*, 305–364. [[CrossRef](#)]
25. Yu, S.; Yang, H.; Nakahara, H.; Santos, G.S.; Nikolić, D.; Pleniz, D. Higher-order interactions characterized in cortical activity. *J. Neurosci.* **2011**, *31*, 17514–17526. [[CrossRef](#)] [[PubMed](#)]
26. Bairey, E.; Kelsic, E.D.; Kishony, R. High-order species interactions shape ecosystem diversity. *Nat. Commun.* **2016**, *7*, 12285. [[CrossRef](#)] [[PubMed](#)]
27. Battiston, F.; Amico, E.; Barrat, A.; Bianconi, G.; de Arruda, G.F.; Franceschiello, B.; Iacopini, I.; Kéfi, S.; Latora, V.; Moreno, Y.; et al. The physics of higher-order interactions in complex systems. *Nat. Phys.* **2021**, *17*, 1093 [[CrossRef](#)]
28. Cervantes-Loreto, A.; Ayers, C.A.; Dobbs, E.K.; Brosi, B.J.; Stouffer, D.B. The context dependency of pollinator interference: How environmental conditions and co-foraging species impact floral visitation. *Ecol. Lett.* **2021**, *24*, 1443–1454. [[CrossRef](#)] [[PubMed](#)]
29. Franzese, N.; Groce, A.; Murali, T.M.; Ritz, A. Hypergraph-based connectivity measures for signaling pathway topologies. *PLoS Comput. Biol.* **2019**, *15*, e1007384. [[CrossRef](#)]
30. Klamt, S.; Haus, U.U.; Theis, F. Hypergraphs and cellular networks. *PLoS Comput. Biol.* **2009**, *5*, e1000385. [[CrossRef](#)]
31. Larhlimi, A.; Blachon, S.; Selbig, J.; Nikoloski, Z. Robustness of metabolic networks: A review of existing definitions. *Biosystems* **2011**, *106*, 1–8. [[CrossRef](#)]
32. Yeung, M.; Thiele, I.; Palsson, B.Ø. Estimation of the number of extreme pathways for metabolic networks. *BMC Bioinform.* **2007**, *8*, 363. [[CrossRef](#)]
33. Ghaderi, S.; Haraldsdóttir, H.S.; Ahookhosh, M.; Arreckx, S.; Fleming, R.M. Structural conserved moiety splitting of a stoichiometric matrix. *J. Theor. Biol.* **2020**, *499*, 110276 [[CrossRef](#)]
34. Pearcy, N.; Chuzhanova, N.; Crofts, J.J. Complexity and robustness in hypernetwork models of metabolism. *J. Theor. Biol.* **2016**, *406*, 99–104. [[CrossRef](#)]
35. Mulas, R.; Zhang, D. Spectral theory of Laplace operators on oriented hypergraphs. *Discret. Math.* **2021**, *344*, 112372. [[CrossRef](#)]
36. Jost, J.; Mulas, R. Hypergraph Laplace operators for chemical reaction networks. *Adv. Math.* **2019**, *351*, 870–896. [[CrossRef](#)]

37. Chitra, U.; Raphael, B. Random Walks on Hypergraphs with Edge-Dependent Vertex Weights. In Proceedings of the 36th International Conference on Machine Learning, Long Beach, CA, USA, 10–15 June 2019; pp. 1172–1181
38. Beguerisse-Díaz, M.; Bosque, G.; Oyarzún, D.; Picó, J.; Barahona, M. Flux-dependent graphs for metabolic networks. *NPJ Syst. Biol. Appl.* **2018**, *4*, 32. [[CrossRef](#)] [[PubMed](#)]
39. Link, H.; Christodoulou, D.; Sauer, U. Advancing metabolic models with kinetic information. *Curr. Opin. Biotechnol.* **2014**, *29*, 8–14. [[CrossRef](#)]
40. Adadi, R.; Volkmer, B.; Milo, R.; Heinemann, M.; Shlomi, T. Prediction of microbial growth rate versus biomass yield by a metabolic network with kinetic parameters. *PLoS Comput. Biol.* **2012**, *8*, e1002575. [[CrossRef](#)] [[PubMed](#)]
41. Gillespie, D.T. Exact stochastic simulation of coupled chemical reactions. *J. Phys. Chem.* **1977**, *81*, 2340–2361. [[CrossRef](#)]
42. Srinivasan, S.; Cluett, W.R.; Mahadevan, R. Constructing kinetic models of metabolism at genome-scales: A review. *Biotechnol. J.* **2015**, *10*, 1345–1359. [[CrossRef](#)]
43. Orth, J.D.; Thiele, I.; Palsson, B.Ø. What is flux balance analysis? *Nat. Biotechnol.* **2010**, *28*, 245–248. [[CrossRef](#)]
44. Oberhardt, M.A.; Palsson, B.Ø.; Papin, J.A. Applications of genome-scale metabolic reconstructions. *Mol. Syst. Biol.* **2009**, *5*, 320. [[CrossRef](#)]
45. Fang, X.; Lloyd, C.J.; Palsson, B.O. Reconstructing organisms in silico: Genome-scale models and their emerging applications. *Nat. Rev. Microbiol.* **2020**, *18*, 731–743. [[CrossRef](#)]
46. Traversa, P.; de Arruda, G.F.; Moreno, Y. From unbiased to maximal entropy random walks on hypergraphs. *arXiv* **2023**, arXiv:2306.09499.
47. Banerjee, A. On the spectrum of hypergraph. *Linear Algebra Its Appl.* **2021**, *614*, 82–110. [[CrossRef](#)]
48. King, Z.A.; Lu, J.; Dräger, A.; Miller, P.; Federowicz, S.; Lerman, J.A.; Ebrahim, A.; Palsson, B.O.; Lewis, N.E. BiGG Models: A platform for integrating, standardizing and sharing genome-scale models. *Nucleic Acids Res.* **2016**, *44*, D515–D522. [[CrossRef](#)]
49. BiGG Models. Available online: <http://bigg.ucsd.edu/models> (accessed on 20 June 2022).
50. Estrada, E.; Hatano, N.; Benzi, M. The physics of communicability in complex networks. *Phys. Rep.* **2012**, *514*, 89–119. [[CrossRef](#)]
51. Estrada, E.; Hatano, N. Communicability in complex networks. *Phys. Rev. E* **2008**, *77*, 036111. [[CrossRef](#)] [[PubMed](#)]
52. Estrada, E. The many facets of the Estrada indices of graphs and networks. *SeMA J.* **2022**, *79*, 57–125. [[CrossRef](#)]
53. Jun, W.; Barahona, M.; Yue-Jin, T.; Hong-Zhong, D. Natural connectivity of complex networks. *Chin. Phys. Lett.* **2010**, *27*, 078902. [[CrossRef](#)]
54. Rosvall, M.; Trusina, A.; Minnhagen, P.; Sneppen, K. Networks and cities: An information perspective. *Phys. Rev. Lett.* **2005**, *94*, 028701. [[CrossRef](#)]
55. Sneppen, K.; Trusina, A.; Rosvall, M. Hide-and-seek on complex networks. *Europhys. Lett.* **2005**, *69*, 853. [[CrossRef](#)]
56. Orth, J.D.; Fleming, R.M.; Palsson, B.Ø. Reconstruction and use of microbial metabolic networks: The core *Escherichia coli* metabolic model as an educational guide. *Ecosal Plus* **2010**, *4*, 10–128. [[CrossRef](#)]
57. *E. Coli* Core Fba Escher Map. Available online: <https://sbrg.github.io/escher-fba/> (accessed on 21 September 2023).
58. Becker, S.A.; Palsson, B.Ø. Genome-scale reconstruction of the metabolic network in *Staphylococcus aureus* N315: An initial draft to the two-dimensional annotation. *BMC Microbiol.* **2005**, *5*, 8. [[CrossRef](#)] [[PubMed](#)]
59. Seif, Y.; Monk, J.M.; Mih, N.; Tsunemoto, H.; Poudel, S.; Zuniga, C.; Broddrick, J.; Zengler, K.; Palsson, B.O. A computational knowledge-base elucidates the response of *Staphylococcus aureus* to different media types. *PLoS Comput. Biol.* **2019**, *15*, e1006644. [[CrossRef](#)] [[PubMed](#)]
60. Jamshidi, N.; Palsson, B.Ø. Investigating the metabolic capabilities of *Mycobacterium tuberculosis* H37Rv using the in silico strain iNJ 661 and proposing alternative drug targets. *BMC Syst. Biol.* **2007**, *1*, 26. [[CrossRef](#)] [[PubMed](#)]
61. Kavvas, E.S.; Seif, Y.; Yurkovich, J.T.; Norsigian, C.; Poudel, S.; Greenwald, W.W.; Ghatak, S.; Palsson, B.O.; Monk, J.M. Updated and standardized genome-scale reconstruction of *Mycobacterium tuberculosis* H37Rv, iEK1011, simulates flux states indicative of physiological conditions. *BMC Syst. Biol.* **2018**, *12*, 25. [[CrossRef](#)]
62. Norsigian, C.J.; Kavvas, E.; Seif, Y.; Palsson, B.O.; Monk, J.M. iCN718, an updated and improved genome-scale metabolic network reconstruction of *Acinetobacter baumannii* AYE. *Front. Genet.* **2018**, *9*, 121. [[CrossRef](#)]
63. Thiele, I.; Hyduke, D.R.; Steeb, B.; Fankam, G.; Allen, D.K.; Bazzani, S.; Charusanti, P.; Chen, F.C.; Fleming, R.M.; Hsiung, C.A.; et al. A community effort towards a knowledge-base and mathematical model of the human pathogen *Salmonella Typhimurium* LT2. *BMC Syst. Biol.* **2011**, *5*, 8. [[CrossRef](#)]
64. Bordbar, A.; Jamshidi, N.; Palsso, B.Ø. iAB-RBC-283: A proteomically derived knowledge-base of erythrocyte metabolism that can be used to simulate its physiological and patho-physiological states. *BMC Syst. Biol.* **2011**, *5*, 110. [[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.