



HAL
open science

Prime Numbers Networks: Visualizing and Characterizing Relationships Between Prime Numbers

Luciano da Fontoura Costa

► **To cite this version:**

Luciano da Fontoura Costa. Prime Numbers Networks: Visualizing and Characterizing Relationships Between Prime Numbers. 2023. <hal-04011595>

HAL Id: hal-04011595

<https://hal.science/hal-04011595v1>

Preprint submitted on 2 Mar 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY-NC-ND 4.0 - Attribution - Non-commercial use - No Derivative Works - International License

Prime Numbers Networks: Visualizing and Characterizing Relationships Between Prime Numbers

Luciano da F. Costa
luciano@ifsc.usp.br

São Carlos Institute of Physics – DFCM/USP

20th Feb. 2023

Abstract

Prime numbers constitute mathematical concepts of particular interest as a consequence of their relative paucity among integer values, as well as to the relative unpredictability of their appearance and density along the sequence of integer values. As a consequence of their own definition, prime numbers can be thought of as being relatively uncommon values along the integers. However, at the same time, prime numbers constitute the building blocks of all the other, composite, numbers, which gives rise to the interesting and important concept of prime factor decomposition. While this concept provides an interesting approach for quantifying interrelationships between composite numbers, it cannot be applied to pairs of prime numbers, as they do not share any factor distinct from 1. In this work, the sets of shortest paths between the prime numbers in respective prime factors networks are taken into account as the primary subsidy for estimating, in quantitative manner, interrelationships between pairs of prime numbers. More specifically, we take into account not only the length and number of shortest paths between each pair of prime numbers in a respective prime factor network, but also quantify the independence (which is related to the entanglement) between the shortest paths in a same given set. This is performed in terms of the concept of effective widths along the hierarchical levels of a set of shortest paths, corresponding to the exponential entropy of the respective sets of transition probabilities. An alternative approach, in which the shared links have their transition probabilities added, is also described for defining effective widths. Several interesting results are reported and discussed, including the identification of intrinsic heterogeneity of interconnections in prime numbers networks and the marked diversity of independence between the shortest paths in the sets respective to pairs of prime numbers. These interrelationships between shortest paths are quantified in terms of statistics of the respective effective width and efficiency index, resulting in relatively intricate distributions among the considered pairs of prime numbers. In particular, several of these indices reach the largest values along an “L”-shaped crest.

1 Introduction

Prime numbers have fascinated humans for a long period of time. Deriving from the Latin term *primus*, meaning *first*, prime numbers occupy a special place along the sequence of integer values especially because they constitute *primary* factors from which other (composite) subsequent numbers can be obtained by factor multiplication. More specifically speaking, prime numbers are integer numbers (such as 2, 3, 5, 7, 11, 13, 17) which happen to be divisible only by the number 1 and by themselves. For simplicity’s sake, only positive integer numbers are considered henceforth in this work, and the number 1 will not be considered as a prime number.

Possibly the main reason accounting for the continued interest that prime numbers have exerted on humans concerns not only their relative *paucity*, but also their in-

herent *unpredictability*, in the sense that it is not possible to accurately predict when the next number will appear along the integer numbers sequence (even though the Prime Number Theorem provides in approximate manner the asymptotic distribution of prime values). In a sense, prime numbers could be said to be interesting as a consequence of being irregularly distributed and uniquely built.

Though prime numbers are relatively special and rare, especially as one moves upward along the integer sequence, a prime number may actually constitute one of the factors composing other numbers. Actually, the decomposition of an integer value in its *prime factors*, which is *unique*, constitutes an especially interesting operation in number theory, called *prime factor decomposition*. For example, the number 1815 can be decomposed into the prime numbers $(3, 5, 11^2)$, where the prime factor 11 appears with *multiplicity* 2. When seen from the prime factor

perspective, prime numbers start being noticed not only for being relatively rare, but also for providing ubiquitous basic building blocks for all integer numbers.

In addition to providing a unique “recipe” for building numbers from prime components, prime factor decomposition (PFD) also paves the way to establishing *relationships* between integer numbers that take into account common prime factors and divisibility. For instance, if two integer values share some prime factors, both these numbers are necessarily *divisible* by the product of those prime factors. The respective prime factors underlying the integer values also provide a means to consider the *similarity of content* between those numbers, as far as their prime factor composition is concerned. For instance, the numbers 1815 and 165 can be said to have *similar contents*, as both these integer values share 3, 5 and 11 as common prime factors.

The systematic consideration of the common prime factors among a given set of integer numbers has motivated respective *network representations* (see Section 2). More specifically, in a *prime factors network* (e.g. [1]), each given integer value is represented as a node, while links are incorporated between pairs of nodes that share at least one prime factor.

Figure 1 illustrates this type of network respectively to a small set of integer numbers, more specifically $\{4, 6, 10, 15, 30\}$.

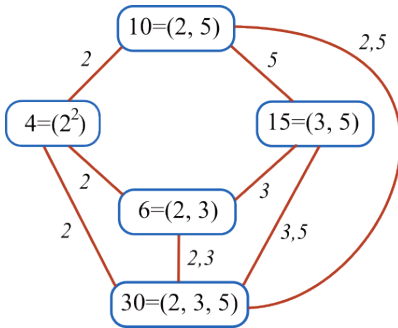


Figure 1: A simple example of a *prime factors network*, in which a given set of integer values, $\{4, 6, 10, 15, 30\}$ in this specific example, are represented as respective nodes incorporating the prime factor decomposition, and linked in pairwise manner whenever two nodes share at least one prime factor, which can be indicated as *weights* or *labels* associated to the links in the network, as illustrated in the above figure.

Pairs of nodes are interconnected through a respective link whenever they share at least one prime factor. The number of shared prime factors can then be taken as a *weight* associated to that respective link. Therefore, the interconnectivity of prime factors networks, such as that illustrated in Figure 1, reflects the sharing of common prime factors and divisibilities. For instance, a group of nodes that result to be more interconnected among

themselves than with the remainder of the network can be understood as a *module* or *cluster* of respective numbers that mutually share a significant number of common prime factors and divisibility properties.

Though the network in Figure 1 considers the multiplicity of the prime factors, in the remainder of this work we will focus only on *unique prime factors*. For instance, we will consider the number 4 as being composed by the prime factor 2 irrespectively of its multiplicity. If needed, all subsequently presented concepts and methods can be adapted to take into account the prime factors multiplicity.

In addition, while prime factors networks can be used to study relationships between generic sets of integer values, here we focus only on sets S_N of prime numbers containing the first N_p primes smaller or equal than an integer value N and which participate as a prime factor in at least one of the numbers from 2 to N . For instance, the prime numbers considered for $N = 20$ are $S_{20} = \{2, 3, 5, 7\}$, so that $N_p = 4$. Though being smaller than $N = 20$, the next prime number, namely 11, is not incorporated into S_{20} because it does not participate as a prime factor of any of the integer numbers from 2 to $N = 20$.

Each of the N_p primes in the set S_N will be henceforth understood to be sorted in increasing order and represented as $p_i, i = 1, 2, \dots, N_p$. Thus, in the case of S_{20} , we have that $p_1 = 2, p_2 = 3, p_3 = 5, \text{ and } p_4 = 7$.

In this work, we are interested in obtaining networks in which each node corresponds to a prime number in a given set S_N , while a pair of such nodes is linked provided there is a shortest path between these nodes in the complete prime factors network containing all the N nodes. Thus, each prime numbers network is henceforth understood to be associated to a respective complete prime factors network containing $N - 1$ nodes corresponding to the integer values $2, 3, \dots, N$, also including the respective set S_N of prime numbers.

Figure 2 depicts one of the shortest path, more specifically $(3, 18, 14, 7)$, between the prime numbers 3 and 7 in the set S_{20} . These shortest paths are characterized by chains of adjacent links (i.e. *paths*) in the original prime factors network which have the smallest possible length. Observe that the same set of shortest paths (though in reversed order) will be obtained while considering moving from one of the extremities to the other, or vice-versa. However, the specific flow at each involved link may vary when the set of shortest paths is reversed, so that indices such as the effective width and efficiency will not be necessarily commutative.

The total number of shortest paths between two prime numbers will be henceforth referred to as R .

Observe that every shortest path between two given prime numbers in a set S_N , having started at one ex-

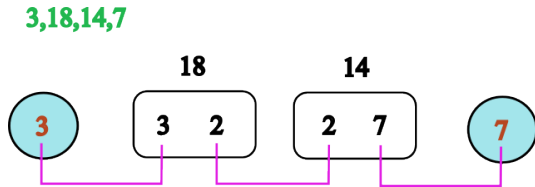


Figure 2: One of the possible shortest paths between the prime numbers 3 and 7 considering S_{20} . A path from a prime to another prime number consists of a chaining of successively shared prime factors defined in pairwise manner between any of the nodes in the complete prime factors network, such as the factor 3 being common to both the integer values 3 and 18, followed by the factor 2 shared between the integers 18 and 14, and so on. Observe that, as adopted henceforth, only 2 and 3 are identified as the unique prime factors of the number 18, which actually has the prime factor 3 with multiplicity 2, i.e. $18 = (2)(3)^2$. As implied by their name, a set of shortest paths between a pair of prime numbers will correspond to those paths that are the shortest among all possibilities. In case more than one such shortest paths can be found, they will necessarily all have the same length L , corresponding to the involved number of links. For instance, the shortest path in the above example has length $L = 3$.

tremity A , will necessarily proceed to a larger integer non-prime value, because in the considered set of integer numbers there will be (by the own definition of primes) no prime number smaller than the initial prime number p_1 that mutually share a prime factor. The last link along the shortest path will also proceed, for the same reason, from a larger non-prime number to the prime number B at the other extremity of the shortest path. We also have that no prime numbers can be found in the middle of a shortest path between two prime numbers.

For the above reasons, no shortest paths of length 1 will ever be found between two distinct prime numbers. In addition, it has been experimentally observed that all identified shortest paths have lengths 2 or 3.

The type of network obtained by considering the shortest paths between a set S_N of prime numbers is henceforth referred to simply as a *prime numbers network*. Each link can be associated to the respective set of shortest paths in the complete network, as illustrated in Figure 3 respectively to the prime numbers network obtained for S_{20} .

Prime numbers networks such as that illustrated in Figure 3 provide an effective means for representing the interrelationships between a set of prime numbers as discussed above, indicating the chained common factors necessary to proceed from one of the prime values to the other.

More specifically, the properties of the respective links, namely the number of shortest paths and their lengths, indicate how closely related each pair of prime numbers can be understood to be. For instance, in the case of the small network in Figure 3, we have that the pairs of prime numbers $\{2, 5\}$ and $\{2, 7\}$ are each interrelated by a single shortest path of length 2, as can be observed in the

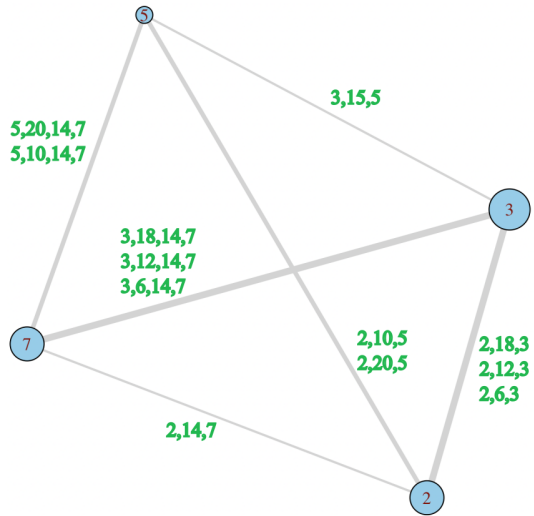


Figure 3: A small *prime numbers network*, formed by the prime numbers in the set S_{20} that are smaller or equal to N and that participate as a prime factor in at least one of the numbers from 2 to $N = 20$, i.e. $S_{20} = \{2, 3, 5, 7\}$. The links between pairs of prime numbers indicate the presence of one or more shortest paths (of the same size), in the respectively associated complete prime factors networks, with $N = 19$ nodes (i.e. $2, 3, \dots, 20$). These shortest paths are here presented in green next to the respective links. The number of shortest paths associated to a given link can be understood as its respective *weight*, which is here reflected in the links width. The sizes of the nodes are shown as being proportional to their strength, i.e. the respective total sum of weights.

labels associated to the respective links in the figure. On the other hand, the primes in the pair $\{3, 7\}$ are related by three shortest paths of length 3, indicating a relatively stronger relationship than in the previous two cases. In this sense, the closest related pair of primes in this example is $\{2, 3\}$, which involves three shortest paths of length 2.

As a summary of the type of study of interrelationships between prime numbers as discussed above, a pair of prime numbers will be considered as being more closely interrelated when there are many distinct shortest paths with a relatively small length between the two respective prime numbers in that pair. Given a shortest path, each of its links in the respective original prime factors network corresponds to a shared prime divisor, which is understood as underlying the relationships between primes as revealed by the suggested prime numbers networks.

Therefore, the longest the shortest paths, the more changes (products/divisions) will be needed along the shortest path in order to get from the prime number in one of the extremities to that in the other extremity, which can be understood as these two prime numbers being less closely related. At the same time, the larger the number of shortest paths, the more related that pair can be deemed to be. The interrelationship between two given

prime numbers is therefore henceforth understood considering both the number of shortest paths and their respective length.

Though the above discussed rational and approach present good potential for expressing and visualizing interrelationships between prime numbers, there is an intrinsic caveat in the sense that the number of distinct shortest paths may not correspond so directly to the effective number of interrelations between the two prime values. That is because the paths in a given set of shortest paths can present diverse levels of *entanglement* implying variations of the respective mutual *independence* between the paths (e.g. [2]).

Figure 4 illustrates the concept of entanglement between sets of shortest paths respectively to the pair of primes 19 and 43 in the set S_{150} . Also shown in this figure are the subsequence *hierarchical levels* h respectively to the nodes and links.

A constant flow of probability equal to 1 is assumed to emanate from the *source node*, i.e. the prime number at the left-hand side of the set of shortest paths (19 in this specific example), flowing along all hierarchical levels until reaching the *destination node* (43 in this example). Therefore, the total flow along all the links at each hierarchical level will also add to 1, ensuring normalization for the entropy calculation which is necessary for estimating the respective effective width at each level h .

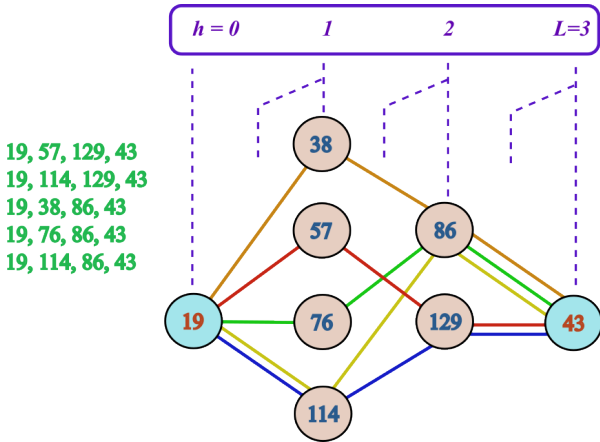


Figure 4: The set of *five* distinct shortest paths between the prime numbers 19 and 43 in the set S_{150} , all of which having length 3. The *hierarchical levels* h of the nodes and links in these shortest paths are also identified in the figure. Though there are five shortest paths, several nodes and links are shared amongst themselves (especially along the second hierarchy), resulting in a respective entanglement leading to an effective number of independent paths smaller than 5. In the present work, we consider the concept of *effective width*, described in [2], as a mean to take into account the entanglement between the shortest paths characterizing prime numbers networks.

It can be readily observed that the five distinct shortest paths identified for this specific pair of prime numbers

in the set S_{150} , which are indicated in green in Figure 4 and in distinct colors in the respective network defined by these shortest paths, though being formally distinct one another, share several nodes and links along the hierarchical levels $h = 1, 2$ and 3.

These several intersections observed between the distinct shortest paths constitute what we will refer to as *entanglements* of those paths. For instance, only two nodes, instead of the maximum $R = 5$, can be found at the hierarchical level $h = 2$. The larger the number of these node and link intersections, the less independent and more entangled the set of paths can be said to be.

The present work adopts a recently described approach [2] to quantifying the entanglement/independence between the sets of shortest paths defined by each pair of prime numbers in sets S_N . That methodology, jointly with a respective modification, are considered here as possible means for quantifying the sets of shortest paths interconnecting the nodes in a prime numbers network.

The approach suggested in [2] assumes that a flow of probability, constant and equal to 1, is established from the *source* to the *destination* nodes, which are taken as the prime numbers at the left- and right-hand sides of Figure 5. As the flow is assumed to be conserved along the R shortest paths, the sum of flow probabilities at each hierarchical level results as being normalized to add 1, therefore allowing a stable probabilistic representation from which the respective entropy, and its exponential, can then be obtained. The equilibrium flow at the links is obtained from the respective transition probabilities as described in [2]. Then, the entropy of the equilibrium flow probabilities at each hierarchical level is calculated, and its exponential is understood to correspond to the *effective width* E_h of the paths at link level h .

For instance, in the case of the shortest paths illustrated in Figure 5, the effective width at level $h = 3$ can be readily calculated as follows:

$$E_3 = \exp \{ -5/8 \log(5/8) - 3/8 \log(3/8) \} = 1.9378\dots$$

This value, here taken as an indication of the effective number of paths at the respective hierarchical level, is much smaller than the total number of shortest paths $R = 5$. Because there are effectively two links at level $h = 3$, the respectively obtained effective width can be understood to be sub-optimal, reflecting the presence of shared links at this specific level.

It is also possible to consider the relative efficiency index as follows:

$$I_3 = \frac{E_3}{5} = 0.38756\dots$$

with $0 \leq I_h \leq 1$.

The overall effective width of the given set of shortest

paths can be quantified in terms of the following average:

$$W_e = \langle E_h \rangle = 4$$

or in terms of the minimum effective width, i.e.:

$$W_m = \min_{h=1,L} E_h = 1.94$$

Observe that sets of shortest paths with length $L = 2$ will necessarily have optimal effective width, as it is not possible to share nodes or links in this particular situation, implying all shortest paths to be mutually independent.

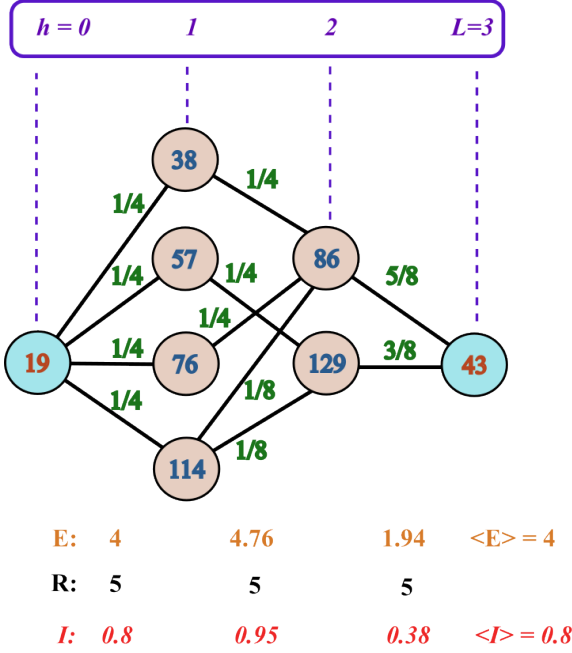


Figure 5: Illustration of the concepts of *equilibrium link flow* and *effective width*, relative to the simple prime numbers network in Fig. 4. The equilibrium flow of probabilities is shown as weights (dark green) associated to each link along the $L = 3$ hierarchies of this network. Observe that the total of probability flow at each hierarchical level h is intrinsically normalized to 1, corresponding to the maximum flow emanating from node 19. The effective widths E of each level corresponds to the exponential entropy of the respective equilibrium probability flows, being shown along the first row at the bottom of the figure. The relative effective width, henceforth called the *efficiency index* I at each level can be obtained by dividing the effective width at that level by the total $R = 5$ of distinct shortest paths in this set. The average values of the effective widths and efficiency index are also shown in the figure as $\langle E \rangle$ and $\langle I \rangle$, respectively.

All in all, we are left with a number of possible measurements reflecting the independence (or entanglement) between a set of shortest paths, including the effective width per hierarchical level E_h , the efficiency index per level I_h , as well as the respective averages and minimum values. The choice among these indices will depend on the specific interests and questions implied by distinct respective applications. In particular, the consideration

of the whole set of indices will provide a more complete characterization of the relationship between the considered prime numbers. At the same time, in case bottlenecks are of particular relevance for a given application, it would also be possible to take into account the minimum effective width as a sole quantification of the prime number relationships.

The above illustrated approach to characterizing the interrelationship between the shortest paths in a given set does not take into account the specific paths in the respective calculations, relying instead on the topology of the interconnections as in the respective prime factors network and considering a randomly uniform random path along the links.

It is also possible to take an alternative, non-probabilistic approach, in which all the identified shortest paths are taken into account as illustrated in Figure 6. In this case, the flows at each link are determined simply by adding the transition probability $1/R$ of each link, so that the effective widths are calculated by taking into account these probabilities instead of the probability link flow as in the previous approach described above. Figure 6 presents the estimation of the effective widths for S_{20} considering this alternative methodology.

At least for the set S_{20} , similar indices quantifying the independence / entanglement between the respective sets of shortest paths have been obtained by using any of the two described methodologies. This second approach to estimating effective widths, however, has the property that the same values will be obtained (though in reversed order) when applied to a set of shortest paths from A to B as well as from B to A .

It should be observed that the two approaches for obtaining effective widths of the sets of shortest paths described above are not alternatives leading to the same estimation, but actually two different (but related) approaches to defining the own concept of effective width.

It follows from the above assumptions and discussions (see also [2]) that an optimal set of R shortest paths would be characterized by R completely independent shortest paths, all with the same length L , which do not share any node or link. As a consequence, a total of $(L - 1)R + 2$ nodes and LR links will be required for the an optimal set of R shortest paths of length L . In the present work, the independence between the shortest paths averages in prime numbers networks will be characterized in terms of the average and minimum values of effective widths and efficiency index.

The present work therefore describes an approach to interrelate prime numbers in a set S_N in terms of respective *prime numbers networks* derived from respective sets of shortest paths existing in the related *prime factors network*. Once this network has been obtained, it can be not

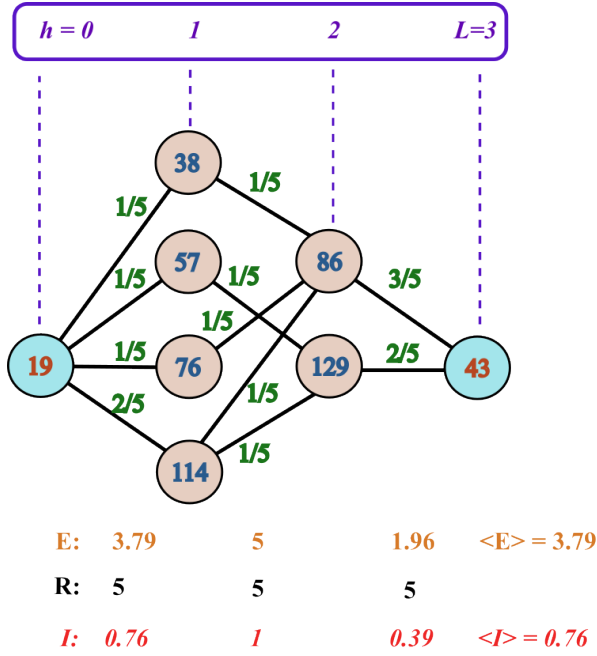


Figure 6: An alternative approach to the estimation of the effective width of a set of shortest paths along its subsequent hierarchical levels. Instead of considering the probability flow as in Fig. 5, each individual link along each of the $R = 5$ identified shortest paths between the prime numbers 19 and 43 is associated to a respective constant transition probability $p = 1/R$, which corresponds to $p = 0.2$ in the specific case shown in the figure. This basic probability is added among shared links, yielding the shown distribution of probability flows, from which the effective widths at each hierarchical level are determined in terms of the respective exponential entropy in a manner analogous to that adopted in the previous described approach.

only visualized, but characterized in terms of several measurements including statistics of the number and length of shortest paths between each considered prime number, as well as by statistics of the respective effective widths and efficiency indices.

More specifically, the main motivation for the concepts, methods and results described in the current work can be summarized as:

- Suggest an approach for interrelating prime numbers in terms of shortest paths between composite numbers in which they participate within respective prime factors networks;
- Obtain complex networks where each prime number is a node, while links between pairs of nodes indicate the presence of respective shortest paths between those values;
- Characterize these interrelations in terms of the number of and length of involved shortest paths, as well as statistics their effective width and efficiency indices;

- Use prime numbers networks as a benchmark for illustrating and evaluating concepts and methods related to the characterization of entanglement between shortest paths in complex networks.

The present work is organized as follows. First, a brief review of some of the main related works is presented, followed by a summary of the adopted concepts and methods. Several methods and results are then presented, including the determination of prime numbers networks and respective characterization in terms of the properties of the respective sets of shortest paths.

2 Related Works

The systematic study of prime numbers constitute one of the subjects of the area known as *Number Theory* (e.g. [3, 4], see also [5, 6, 7]).

Several interesting works have addressed the properties and interrelationship between prime numbers, including respective representation as *graphs* and/or *networks*. Among these works, Pomerance [8] described a *prime number graph*, consisting of the set of lattice points (n, p_n) , $n = 1, 2, \dots$, where p_n is the n -th prime number, indicating that there are k collinear points in this type of graph. Graphs have also been defined [9] by including edges between each pair of prime numbers whose absolute difference is a non-negative power of 2.

In another approach [10] (see also [11, 12]), integers are interconnected whenever one divides the other, so that an unweighted network is obtained, which are found to be scale free, but with non-stationary degree distribution. Congruence relationships between natural numbers in terms of multiplex networks have been addressed in [13].

Networks of prime numbers have also been approached [14] from the perspective of the Goldbach's conjecture, stating that an even number can be decomposed as the sum of two prime numbers, and the obtained networks have been found to be small world.

A bipartite approach to composite and prime numbers has been described in [15] in the context of stochastic processes theory, and a method is also described for generating random numbers and their relationships with composite numbers.

Networks based on shared prime factor decomposition of integer numbers have been described in [1], networks of integer numbers have been described in which each number is associated to a node, and pairs of these nodes are interconnected whenever the two respective integer values share a prime factor. Weights corresponding to the total number of shared factors, or the coincidence similarity between the respective vector of prime factors, can then be

associated to each link.

3 Basic Concepts and Methods

In this section we provide a brief presentation of the main concepts and methods adopted in the present work, starting with prime numbers (e.g. [16]) and following with complex networks e.g. [17, 18, 19]).

An integer value m is said to be *divisible* by another value p iff m/p is an integer number. In this work, we assume all integer values to be positive. A *prime number* is an integer number that is divisible only by 1 and by itself. Numbers that are not primes, are said to be *composite*. Any given integer number m can be *decomposed* into respective integer components, or *factors*, which are found to be *divisors* of m . In particular, the *prime factor decomposition* of an integer number corresponds to all its factors that are prime numbers. Prime factors may repeat themselves with a respective *multiplicity*.

A *complex network* is a graph with noticeable interconnection complexity, in the sense of being substantially greater than the complexity of a deterministic regular network or a stochastic uniformly random network. Each network is henceforth understood to have n nodes and e edges or links. The number of links attached to a given node corresponds to the node *degree*. A *path* along a network consists of a sequence of adjacent links. The number of links in a given path corresponds to its respective *length*. Two nodes are said to be *connected* iff there is a path between them. The *shortest path* between two given nodes corresponds to the path of minimum length among the possible paths interconnecting those two nodes. More than one shortest path can be obtained for the same pair of nodes, but all them will necessarily have the same length.

Given a set of R probabilities P_i , $i = 1, 2, \dots, R$, *normalized* so that:

$$\sum_{i=1}^R P_i = 1 \quad (1)$$

the *entropy* of this set is calculated as:

$$\varepsilon = - \sum_{i=1}^R P_i \log(P_i) \quad (2)$$

The *exponential entropy* of this set of probabilities can then be expressed as:

$$X = e^\varepsilon \quad (3)$$

This quantity can be understood as corresponding to the *effective* number of choices in the original set of probabilities. The largest number of choices, corresponding

to R , is obtained when all probabilities are identical to $P_i = 1/R$. Otherwise, values smaller than R will be obtained, especially when the probability values are particularly unbalanced.

4 Examples of Prime Factors and Prime Numbers Networks

All experimental results presented in this work refer to the sets of primes S_{300} or S_{1000} .

Figure 11 illustrates the *prime factors network* obtained for the integer values $2, 3, \dots, 300$. Each node in this network corresponds to an integer value varying from 2 to 300, while the edges have weights corresponding to the number of prime factors shared by the integer values associated to the respective pair of nodes.

A noticeably modular interconnectivity structure can be observed, presenting *modules* (or communities, clusters) interconnected one another through varying densities of edges. The isolated satellite nodes correspond to the prime numbers in the range $2, 3, \dots, 300$ that do not participate as a prime factor in any of the composite numbers in that same range.

The prime numbers network obtained for the sets S_{300} and S_{1000} are depicted in Figures 8 (a) and (b), respectively. As with the respectively prime factors networks from which these networks derive, a marked heterogeneity of interconnections can be observed, including two main modules including a smaller star-like cluster around the number 3, and another module involving the other prime numbers.

In addition to the strong heterogeneity of interconnections, the prime numbers networks also indicate a particularly low number of connections involving the prime number 2, which a substantially larger number of connections involve the number 3. The largest degrees can be observed for intermediate prime numbers, particularly those resulting at the center of the obtained networks. Another particularly interesting result concerns the fact that the largest prime numbers in the prime number networks tend to be connected to the number 3.

5 Length and Number of Shortest Paths

Given a prime numbers network, the distributions of number of shortest paths between pairs of nodes and respective lengths, as discussed in Section 1, constitute a particularly interesting characterization of the interrelationship between the prime numbers in those networks.

Figures 9 (a) and (b) present the number of shortest

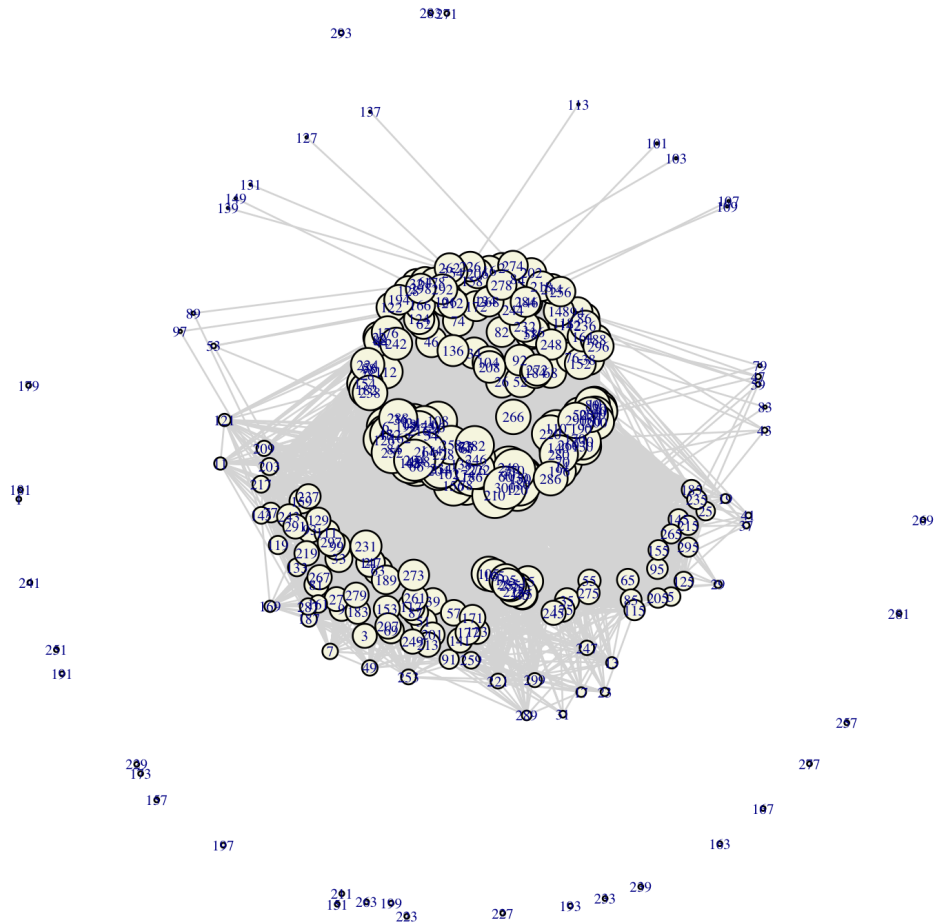


Figure 7: The *prime factors network* obtained for the integer values $2, 3, \dots, 300$, each of which being represented as a node in this network. The link between each pair of numbers corresponds to the existence of shared prime factors. The number of shared prime factors is understood as the *weight* of the respective link. The size of the nodes have been shown so as to indicate the respective *strength*, i.e. the sum of the weights of links attached to that node. A noticeably modular organization can be observed, involving modules that are attached to the remainder of the network through varying densities of edges. The prime numbers which do not participate as factors in other numbers result as surrounding isolated nodes. The other prime numbers in the range from 2 to 300 are connected to other composite numbers as a consequence of participating as a prime factor in those numbers.

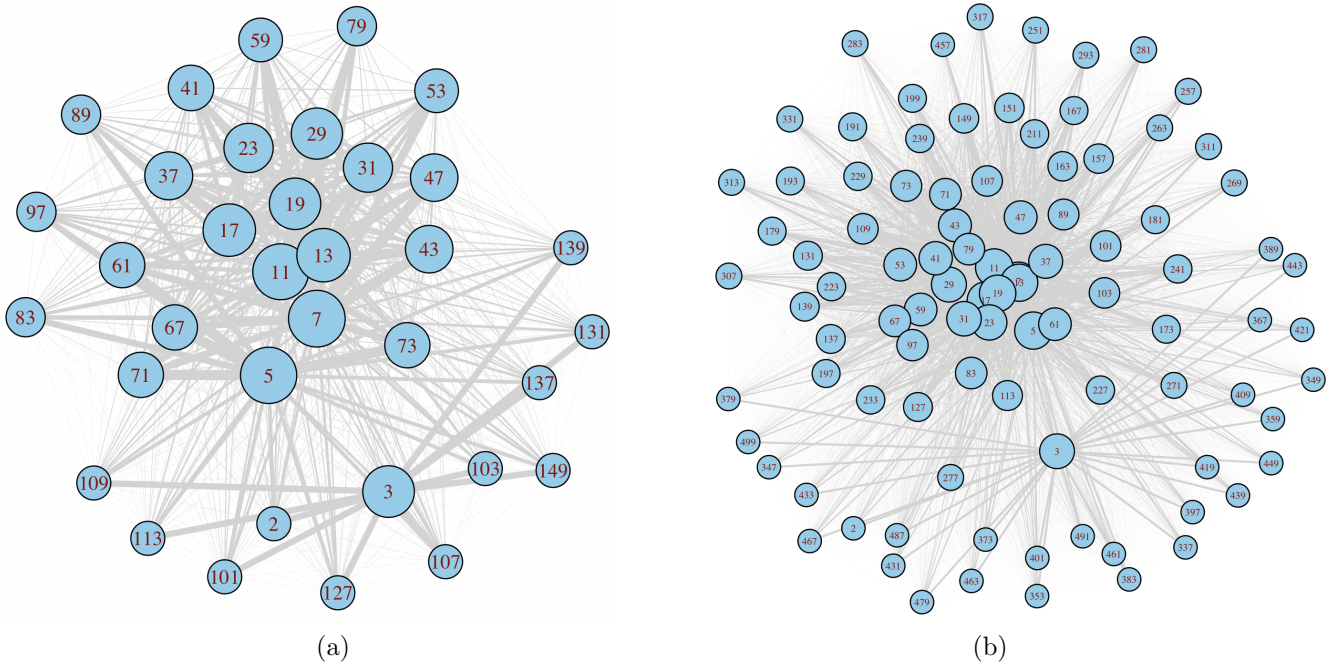


Figure 8: The *prime numbers networks* obtained for the sets S_{300} (a) and S_{1000} (b), respectively, present similar organizations of interconnections. Heterogeneous interconnectivity can be observed in both cases, including two modules. The width of the links indicate the number of shortest paths among each respective pair of nodes, and the size of the nodes reflects the respective strength, corresponding to the sum of the weights associated to respectively attached edges.

paths and respective lengths between each pair of nodes in the prime numbers network obtained for S_{1000} .

A markedly heterogeneous distribution of numbers of shortest paths between pairs of prime numbers can be observed from Figure 9 (a), with the highest numbers defining an “L”-like crest. Interestingly, each of the sides of this crest have distinct properties, with the upper portion presenting a graded distribution of number of shortest paths, which tends to decrease with the distance from the crest. At the same time, the lower region adjacent to the crest is characterized by an abrupt transition to particularly low values of numbers of shortest paths. A three-dimensional visualization of the number of shortest paths between pairs of primes in S_{1000} can be seen in Figure 10.

The lengths of shortest paths in S_N has been identified to be either $L = 2$ or 3 , defining two respective connected regions, as illustrated in Figure 9 (b). The region with $L = 2$ involves pairs of prime numbers so that at least one of these two numbers is relatively small. The second region, characterized by $L = 3$ involves relatively larger prime numbers. This indicates that the shortest paths between prime numbers relatively larger require 3, instead of 2 links before a path can be obtained between those two numbers.

The border between the two regions above can be found to correspond to the crest identified for the number of

shortest paths in Figure 9 (a). Interestingly, a large majority of pairs of primes are characterized by respective sets of shortest paths having length equal to 3 links.

6 Effective Width in Prime Numbers Networks

Having discussed prime numbers networks in terms of the respective number of shortest paths between its pairs of nodes, as well as from the perspective of the length of these paths, in this section we focus on a more systematic characterization of the independence (related to entanglement) between the shortest paths in these sets. More specifically, we will estimate the average effective width and efficient indices of the sets of shortest paths between every pair of prime in the respective networks, as well as their minimum effective width. The effective width will be calculated by using the first, stochastic approach described in Section 1, characterized by the equilibrium link flow, from source to destination nodes, in each set of shortest paths.

However, before proceeding to the presentation of the above mentioned indices, it is interesting to visualize some of the sets of shortest paths between considered (*source, destination*) pairs of nodes in the set S_{300} , as illustrated in Figure 11.

A substantial variation of network size and topology

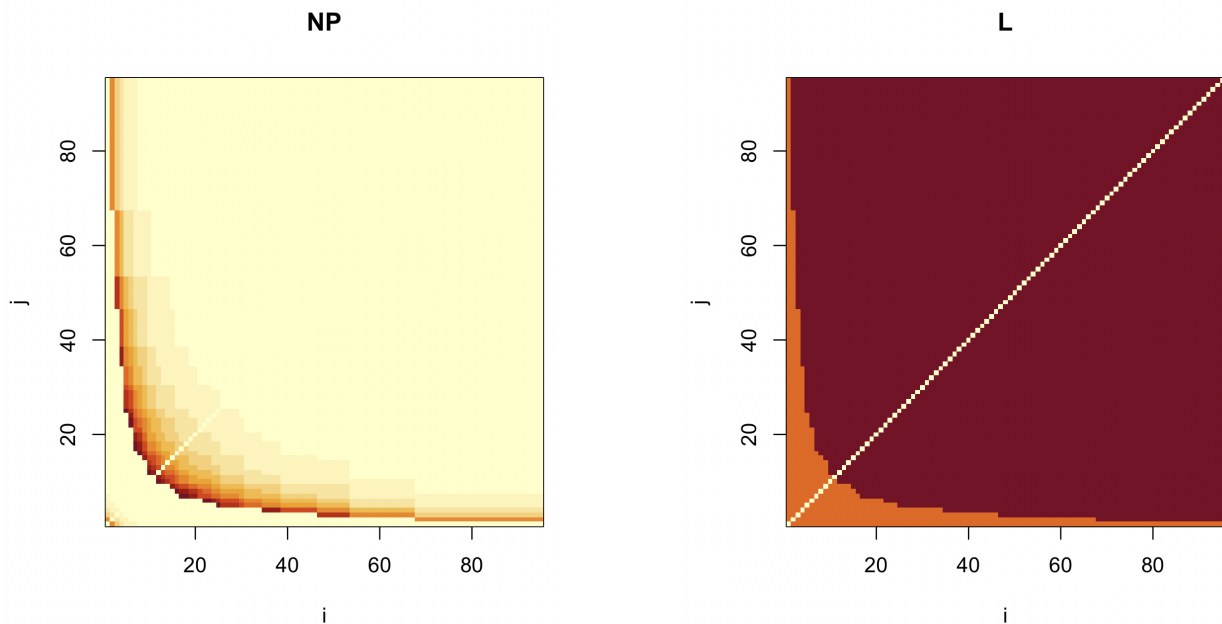


Figure 9: Number of shortest paths between nodes (a) and respective lengths (b) in the prime numbers network obtained for S_{1000} . The adopted heatmap proceeds from yellow to brown. An interesting distribution of values can be observed, with the largest values tending to follow a “L”-like curve. Also interesting is the fact that larger numbers of shortest paths tend to be observed above the aforementioned curve than below. The number of shortest paths in this set of primes have been identified to be either 2 (shown in orange) or 3 (shown in brown). A well-defined border can be observed between the two regions respectively defined by these two values of shortest path length, which coincides with the crest of the “L”-like curve.

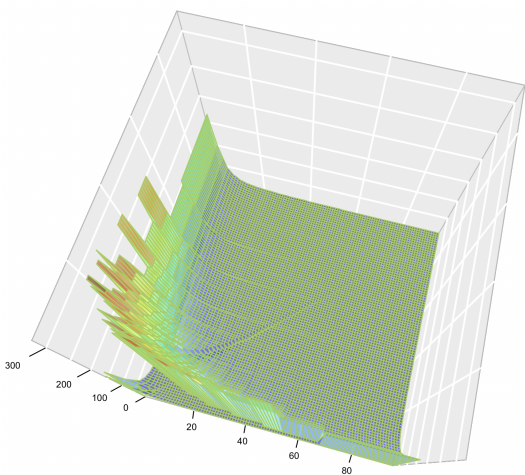


Figure 10: Visualization of the number of shortest paths between the prime numbers in S_{1000} visualized three-dimensionally, from which the markedly distinct distributions of values at each side of the “L”-crest, with abrupt and smooth respective transitions, can be readily observed.

can be observed for the examples of sets of shortest paths, indicating that markedly diverse interactions exist between prime numbers. While all sets of shortest paths with $L = 2$ will have optimal independence, leading to maximum effective widths, the observed diversity of interconnectivity between prime numbers presented in Fig-

ure 11 indicates that great variations of effective width and efficiency indices can be expected.

In this work, we will approach the quantification of the independence between the shortest paths in each set defined by respective pairs of primes in terms of the following four measurements: (i) the average effective width for each pair of nodes in the prime factors networks; (ii) the respective standard deviation; (iii) the minimum effective width; and (iv) the average efficiency index, namely the division of the average effective widths by the total number of shortest paths in each set.

Figure 12 illustrates the values of the above measurements for each pair of the $N_p = 95$ prime nodes in the set S_{1000} .

A strong correlation can be observed between the average effective width W_e of the paths in the prime numbers networks for S_{1000} (Fig. 12(a)) and the respective number of paths in Figure 9(a).

As can be noticed from Figure 12(b), the standard deviation of the effective width resulted correlated with the respective average, indicating that the larger sets of shortest paths also tend to have greater dispersion of effective width and, therefore, greater heterogeneity of mutual independence. Recall that each set of shortest paths is considered between every pair of prime numbers in the set S_N .

A particularly interesting distribution of the minimum

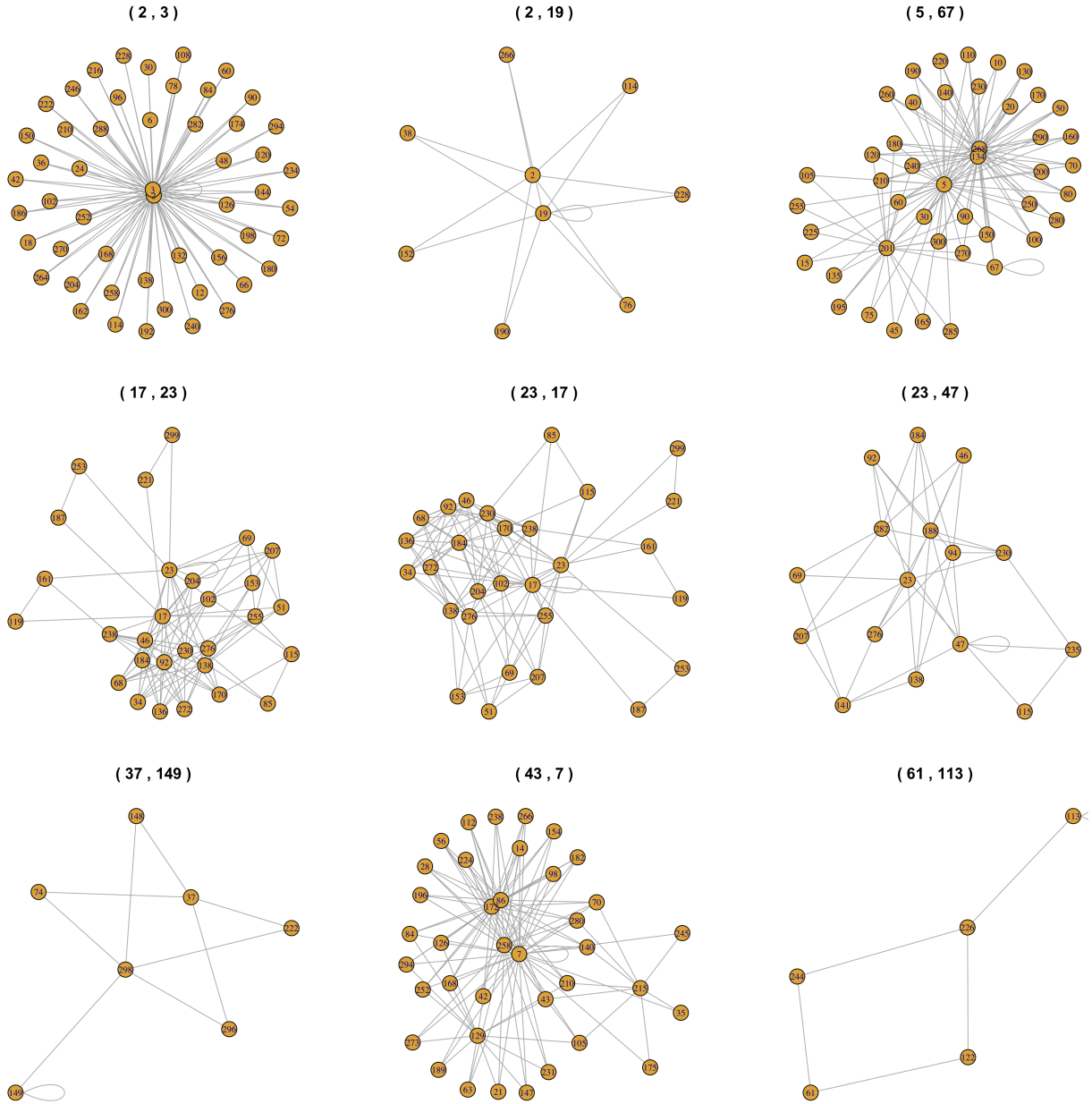


Figure 11: Examples of sets of shortest paths, with lengths $L = 2$ or 3 , between pairs of prime numbers considering S_{300} . The respective identification of the $(source, destination)$ pairs of nodes is provided respectively to each presented case. A marked diversity of interconnectivity can be observed even among these few examples. All destination nodes are marked with a self-loop.

effective width along the hierarchical levels can be verified in Figure 12(c), with a blocked structure observed for pairs of primes above the “L”-shaped crest. The sets of shortest paths at this crest tend to have the largest minimum effective width, though these values decrease progressively as one moves away from the middle of the crest. Particularly large minimum effective widths can also be found near the origin of the axis in the plot.

The distribution of efficiency index values is presented in Figure 12(d). Of particular interest is the fact that the smallest efficiencies can be found along the “L”-shaped crest, increasing progressively with the distance from that

crest. Recall that the efficiency index is not necessarily commutative, given the stochastic methodology adopted for the estimation of the effective widths. Interestingly, the region with the smallest efficiency index values corresponds to the center of the “L”-shaped crest, which is at the same time characterized by large average and minimum effective widths.

Figure 13 illustrates the prime numbers network for S_{1000} considering the average efficiency index $\langle I \rangle$ for the weights of respective links between the involved prime numbers. It can be observed from this network that the highest efficiency index values have been ob-

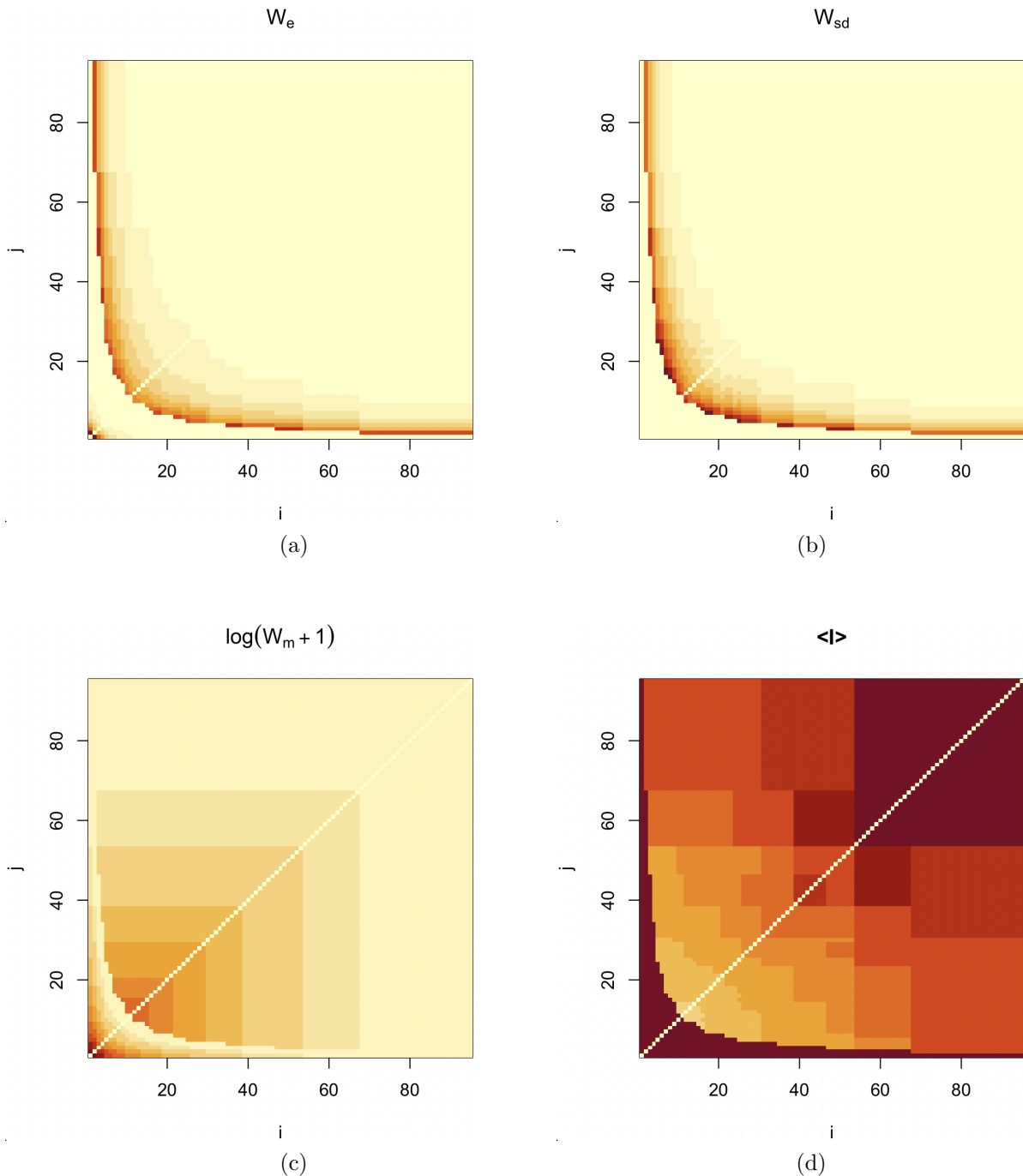


Figure 12: The characterization of the independence between the shortest paths in the respective sets defined by each pairwise combinations of the primes in S_{1000} : (a) the average effective width; (b) the standard deviation of the effective width; (c) the minimum effective width; and (d) the average of the efficiency index. Observe that, for the sake of enhanced visualization, the minimum effective width is shown as $\log(W_{sd} + 1)$. The heatmap progresses from yellow (smallest value) to brown (largest value).

tained between the intermediate valued prime numbers (e.g. 5, 7, 13, 17, ..., 31) and other prime numbers. The two previously identified modules can also be discerned in the resulting network.

All in all, the consideration of the several measurements in this section allowed the complementation of the respective analysis and interpretation, leading to a better understanding of the interrelationships, in terms of sets of

shortest paths, between pairs of prime numbers in sets S_N .

7 Concluding Remarks

Prime numbers have been the subject of intense human interest from the earliest days of mathematic development,

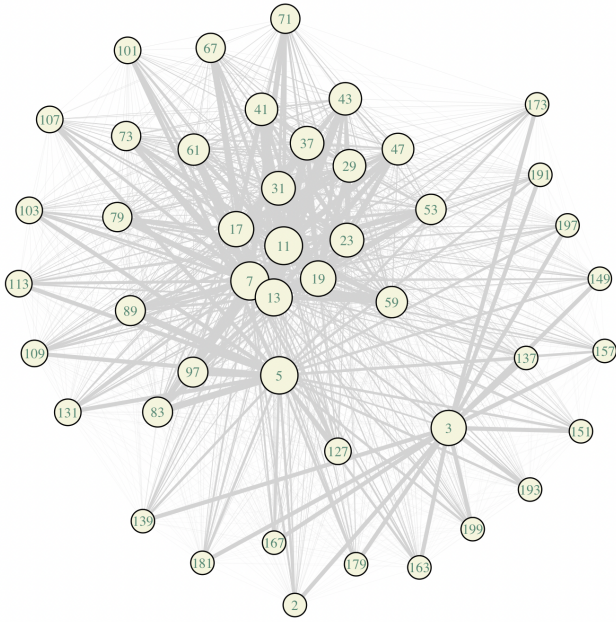


Figure 13: The prime numbers network for S_{1000} taking the average efficiency index $\langle I \rangle$ as weights for the respective edges, reflected in the widths of the links in the figure. The sizes of the nodes indicate the number of shortest paths between each pair of primes in the network.

which is possibly a consequence of being unpredictably and irregularly distributed along the other integer numbers. As a consequence of this continuing interest, a large number of concepts and methods, especially in the area of Number Theory, have been developed and systematically applied.

With advances in graph theory and network science, respective concepts and methods have been developed and applied to the visualization and characterization of relationship between integer values, including their respective prime factor decompositions. The present work developed along this perspective. More specifically, concepts and methods related to graph theory including shortest paths, as well as the exponential entropy, have been suggested in order to obtain prime numbers networks, with links defined from the shortest paths between prime numbers, and to characterize these networks in terms of several measurements and indices capable of expressing the length, number, and effective width of the sets of shortest paths associated to respective pairs of prime numbers.

Several interesting results have been obtained and discussed, including those listed as follows:

- Relationships between prime numbers can be defined in the sense of factors shared with other composite numbers along respective shortest paths in the associated prime factor network, yielding *prime numbers*

networks;

- The concept of *sets of prime numbers* S_N up to a maximum integer number N , have been suggested so as to restrict the possible interactions between the prime numbers, which would otherwise imply infinite sets of numbers;
- The sets of shortest paths obtained for the considered sets of prime numbers revealed only paths with lengths $L = 2$ or 3 , with the latter typically being found between two relatively larger prime numbers in the respective sets;
- The obtained prime numbers networks presented a strongly heterogeneous pattern of interconnectivity, with at least two well-defined modules;
- The smallest between these two modules has star-like connectivity centered at the value 3 , which tends to establish shortest paths with the largest prime numbers in the given set S_N ;
- The largest module has at its most central region the intermediate valued prime numbers, which present a relatively larger number of shortest paths with other prime numbers in the respective set;
- The pairwise distribution of numbers of shortest paths in set defined by a respective pair of prime numbers revealed a well-defined “L”-shaped crest of peak value;
- Completely distinct distributions of lengths and numbers of shortest paths were observed at each of the two sides around the obtained crest, with the upper portion containing particularly large numbers of sets of shortest paths that decrease progressively with the distance from the crest, while the portion at the other side was characterized by smaller numbers of shortest paths;
- The two identified regions around the “L”-shaped crest were composed, respectively, of sets of shortest paths with lengths $L = 2$ and 3 ;
- The sets of shortest paths of length $L = 3$ between the pairs of prime numbers presented a marked topological diversity, implying in varying levels of entanglement / independence between the involved shortest paths;
- The concept of effective width suggested recently [2] was adopted in this work together with a respective modification in which the weights of the links are obtained directly from the respective paths in a set, instead of from the equilibrium. link flow established

between a source and destination node. The latter approach has the characteristic of leading to the same set of effective widths along the respective hierarchical levels, though in reversed order, when the source node is exchanged with the destination node;

- The relationship between the obtained sets of shortest paths were characterized in terms of several indices related to the concept of *effective width* of a set of shortest paths, especially the average, standard deviation and minimum values of the number of shortest paths in a set, as well as in terms of the efficiency index;
- The average and standard deviation of the effective widths resulted strongly correlated with the respective number of shortest paths in the obtained sets, indicating that the former two properties tend to follow the total number of shortest paths in the obtained sets;
- The minimum value of effective widths along the shortest paths hierarchical levels was found to present a particularly intricate distribution which, though tending to take the largest values at or near the identified crest, also established a blocked structure at one of the sides of the crest;
- The distribution of efficiency indices resulted particularly intricate, with the smallest efficiency being observed mostly at the same cases in which the average number of effective paths is high, then increasing progressively with the distance from the crest in the upper respectively defined region.

All in all, the proposed approach unveiled a particularly intricate interrelationship, in terms of shortest paths, between the prime number in bound sets S_N of integer values.

Possible future related works include extending the reported analysis to larger prime numbers networks, considering the second approach for defining effective width suggested in the present work and comparing it with the currently reported results, confirming (or not) that the shortest paths in the described networks necessarily have lengths equal to either 2 or 3, and applying the suggested concepts and methods to the characterization of systems involving interrelationship between prime numbers (e.g. as in some signal and pattern analysis approaches). In addition, the described concepts and methods can also be applied to the characterization of sets of shortest paths in networks other than those defined by prime number relationships.

Acknowledgments

Luciano da F. Costa thanks CNPq (grant no. 307085/2018-0) and FAPESP (grant 15/22308-2).

Observations

As all other preprints by the author, the present work contains preliminary work subject to further revision and validation. Respective modification, commercial use, or distribution of any of its parts are not possible, as this work has author copyright. Many of the preprints by the author are also available in Hal and/or arXiv. This work can also be cited by using the DOI number or article identification link. Thanks for reading.

References

- [1] L. da F. Costa. Prime factors networks. https://www.researchgate.net/publication/361687024_Prime_Factors_Networks, 2022.
- [2] L. da F. Costa. Quantifying shortest paths entanglement in terms of hierarchies and similarities. https://www.researchgate.net/publication/368442645_Quantifying_Shortest_Paths_Entanglement_in_terms_of_Hierarchies_and_Similarities, 2023.
- [3] T. M. Apostol. *Introduction to Analytic Number Theory*. Springer, 1998.
- [4] I. M. Vidogradov. *Elements of Number Theory*. Dover, 2003.
- [5] A. E. Ingham. *The distribution of prime numbers*. Cambridge University Press, 1990.
- [6] D. G. Wells. *Prime numbers*. Wiley Hoboken, 2005.
- [7] R. Crandall and C. Pomerance. *Prime numbers*. Springer, 2001.
- [8] C. Pomerance. The prime number graph. *Mathematics of Computation*, 33(145):399–408, 1979.
- [9] D. E. G. Malm. A graph of primes. *Mathematics Magazine*, 66(5):317–320, 1993.
- [10] S. M. Shekatkar, C. Bhagwat, and G. Ambika. Divisibility patterns of natural numbers on a complex network. *Scientific Reports*, 5(1):1–10, 2015.

- [11] G. Corso. Families and clustering in a natural numbers network. *Physical Review E*, 69(3):036106, 2004.
- [12] D.-H. Shi and H.-J. Zhou. Natural number network and prime number theorem. *Complex Systems and Complexity Science*, 7(4):52–54, 2010.
- [13] X.-Y. Yan, W.-X. Wang, G.-R. Chen, and D.-H. Shi. Multiplex congruence network of natural numbers. *Scientific Reports*, 6(1):1–8, 2016.
- [14] A. K. Chandra and S. Dasgupta. A small world network of prime numbers. *Physica A: Statistical Mechanics and its Applications*, 357(3-4):436–446, 2005.
- [15] G. Garcia-Perez, M. A. Serrano, and M. Boguná. Complex architecture of primes and natural numbers. *Physical Review E*, 90(2):022806, 2014.
- [16] L. da F. Costa. A first glance at prime numbers. https://www.researchgate.net/publication/349466484_A_First_Glance_at_Prime_Numbers_CDT-55, 2021. [Online; accessed 02-Feb-2021].
- [17] A.L. Barabási and Pósfai M. *Network Science*. Cambridge University Press, 2016.
- [18] L. da F. Costa, F. A. Rodrigues, G. Travieso, and P. R. Villas Boas. Characterization of complex networks: A survey of measurements. *Advances in Physics*, 56(1):167–242, 2007.
- [19] L. da F. Costa, O.N. Oliveira Jr., G. Travieso, F.A. Rodrigues, P.R. Villas Boas, L. Antiqueira, M.P. Viana, and L.E.C. Rocha. Analyzing and modeling real-world phenomena with complex networks: A survey of applications. *Advances in Physics*, 60(3):329–412, 2011.