Tanya Araújo, Ennes Ferreira

# The Topology of African Exports: emerging patterns on spanning trees

U

LISBOA

UNIVERSIDADE
DE LISBOA

# The Topology of African Exports: emerging patterns on spanning trees

Tanya Araújo and M. Ennes Ferreira

ISEG, University of Lisbon     Miguel Lupi 20, 1248-079 Lisbon

UECE - Research Unit on Complexity and Economics

SOCIOUS/CSG - C.I. Sociologia Económica e das Organizações

## Abstract

This paper is a contribution to interweaving two lines of research that have progressed in separate ways: network analyses of international trade and the literature on African trade and development. Gathering empirical data on African countries has important limitations and so does the space occupied by African countries in the analyses of trade networks. Here, these limitations are dealt with by the definition of two independent bipartite networks: a destination share network and a commodity share network.

These networks - together with their corresponding minimal spanning trees - allow to uncover some ordering emerging from African exports in the broader context of international trade. The emerging patterns help to understand important characteristics of African exports and its binding relations to other economic, geographic and organizational concerns as the recent literature on African trade, development and growth has shown.

# 1 Introduction

A growing literature has presented empirical findings of the persistent impact of trade activities on economic growth and poverty reduction ([1],[2],[3],[4],[5], [6]). Besides discussing on the relation between trade and development, they also report on the growth by destination hypothesis, according to which, the destination of exports can play an important role in determining the trade pattern of a country and its development path.

Simultaneously, there has been a growing interest in applying concepts and tools of network theory to the analysis of international trade ([7],[8],[9],[10], [11],[12],[13]). Trade networks are among the most cited examples of the use of network approaches. The international trade activity is an appealing example of a large-scale system whose underlying structure can be represented by a set of bilateral relations.

This paper is a contribution to interweaving two lines of research that have progressed in separate ways: network analyses of international trade and the literature on African trade and development.

The most intuitive way of defining a trade network is representing each world country by a vertex and the flow of imports/exports between them by a directed link. Such descriptions of bilateral trade relations have been used in the gravity models ([14]) where some structural and dynamical aspects of trade have been often accounted for.

While some authors have used network approaches to investigate the international trade activity, studies that apply network models to focus on specific issues of African trade are less prominent. Although African countries are usually considered in international trade network analyses, the space they occupy in these literature is often very narrow.

This must be partly due to the existence of some relevant limitations that empirical data on African countries suffer from, mostly because part of African countries does not report trade data to the United Nations. The usual solution in this case is to use partner country data, an approach referred to as **mirror statistics**. However, using mirror statistics is not a suitable source for bilateral trade in Africa as an important part of intra-African trade concerns import and exports by non-reporting countries.

A possible solution to overcome the limitations on bilateral trade data is to make use of information that, although concerning two specific trading countries, might be provided indirectly by a third and secondary source. That is what happens when we define a bipartite network and its one-mode

2

projection. In so doing, each bilateral relation between two African countries in the network is defined from the relations each of these countries hold with another entity. It can be achieved in such a way that when they are similar enough in their relation with that other entity, a link is defined between them.

Our approach is applied to a subset of 49 African countries and based on the definition of two independent bipartite networks where trade similarities between each pair of African countries are used to define the existence of a link. In the first bipartite graph, the similarities concern a mutual leading destination of exports by each pair of countries and in the second bipartite graph, countries are linked through the existence of a mutual leading export commodity between them.

Therefore, bilateral trade discrepancies are avoided and we are able to look simultaneously at network structures that emerge from two fundamental characteristics (exporting destinations and exporting commodities) of the international trade. As both networks were defined from empirical data reported for 2014, we call these networks »**destination share networks**« - $DSN_{14}$ and »**commodity share networks**« - $CSN_{14}$, respectively.

Its worth noticing that the choice of a given network representation is only one out of several other ways to look at a given system. There may be many ways in which the elementary units and the links between them are conceived and the choices may depend strongly on the available empirical data and on the questions that a network analysis aims to address ([15]).

The main question addressed in this paper is whether some relevant characteristics of African trade would emerge from the bipartite networks above described. We hypothesized that specific characteristics could come out and shape the structures of both the $DSN_{14}$ and the $CSN_{14}$. We envision that these networks will allow to uncover some ordering emerging from African exports in the broader context of international trade. If it happens, the emerging patterns may help to understand important characteristics of African exports and its relation to other economic, geographic and organizational concerns.

To this end, the paper is organized as follows: next section presents the empirical data we work with, Section three describes the methodology and some preliminary results from its application. In Section four we present further results and discuss on their interpretation in the international trade setting. Section five concludes and outlines future work.

# 2 Data

Trade Map - Trade statistics for international business development ([16]) - provides a dataset of import and export data in the form of tables, graphs and maps for a set of reporting and non-reporting countries all over the world. There are also indicators on export performance, international demand, alternative markets and competitive markets. Trade Map covers 220 countries and territories and 5300 products of the Harmonized System (HS code).

Since the Trade Map statistics capture nationally reported data of such a large amount of countries, this dataset is an appropriate source to the empirical study of temporal patterns emerging from international trade. Nevertheless, some major limitations should be indicated, as for countries that do not report trade data to the United Nations, Trade Map uses partner country data, an issue that motivated our choice for defining bipartite networks.

Our approach is applied to a subset of 49 African countries (see Table 1) and from this data source, trade similarities between each pair of countries are used to define networks of links between countries.

Table 1 shows the 49 African countries we have been working with. It also shows the regional organization of each country, accordingly to the following classification: 1 - Southern African Development Community (SADC); 2 - União do Magreb Árabe (UMA); 3 - Comunidade Económica dos Estados da Africa Central (CEEAC); 4 - Common Market for Eastern and Southern Africa (COMESA) and 5 - Comunidade Económica dos Estados da África Ocidental (CEDEAO).

For each African country in Table 1, we consider the set of countries to which at least one of the African countries had exported in the year of 2014. The specification of the destinations of exports of each country followed the International Trade Statistics database ([16]) from where just **the first and the second main destinations of exports of each country** were taken.

Similarly and also for each African country in Table 1, we took the set of commodities that at least one of the African countries had exported in 2014. The specification of the destinations of exports of each country followed the same database from where just **the first and the second main export commodities of each country** were taken.

For each country (column label »**Country**«) in Table 1, besides the regional organization (column label »**O**«) and the first and second destinations (column labels »**Destinations**«) and commodities (column labels »**Products**«), we also considered the export value in 2014 (as reported in

4

[16]) so that the size of the representation of each country in the networks herein presented is proportional to its corresponding export value in 2014.

| Country | O | Destinations | | Products | | Country | O | Destinations | | Products | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Seychelles | SEY | 1 | FRA | UK | 16 | 3 | Gabon | GAB | 3 | CHI | JAP | 27 | 44 |
| Angola | AGO | 1 | CHI | USA | 27 | 71 | Burundi | BUR | 3 | PAK | RWA | 9 | 71 |
| Mozambique | MOZ | 1 | CHI | ZAF | 27 | 76 | St.Tome | STP | 3 | BEL | TUR | 18 | 10 |
| D.R.Congo | DRC | 1 | CHI | ZMB | 74 | 26 | Kenya | KEN | 4 | ZMB | TZA | 9 | 6 |
| Botswana | BWT | 1 | BEL | IND | 71 | 75 | Egypto | EGY | 4 | ITA | GER | 27 | 85 |
| South Africa | ZAF | 1 | CHI | USA | 71 | 26 | Ethiopia | ETH | 4 | CHI | SWI | 27 | 9 |
| Zambia | ZAM | 1 | CHI | KOR | 74 | 28 | Uganda | UGA | 4 | RWA | NET | 9 | 27 |
| Tanzania | TZA | 1 | IND | CHI | 71 | 26 | Eritrea | ERI | 4 | CHI | IND | 26 | 9 |
| Namibia | NAM | 1 | BWA | ZAF | 71 | 3 | Comoros | NGA | 4 | IND | GER | 9 | 89 |
| Zimbabwe | ZWE | 1 | CHI | ZAF | 71 | 24 | Rwanda | RWA | 4 | CHI | MAS | 26 | 9 |
| Mauritius | MUS | 1 | USA | FRA | 61 | 62 | Guine Bissau | GuB | 5 | IND | CHI | 8 | 44 |
| Lesotho | LES | 1 | USA | BEL | 71 | 61 | Ghana | GHA | 5 | ZAF | EMI | 27 | 18 |
| Malawi | MWI | 1 | BEL | GER | 24 | 12 | Cote d'Ivoire | CIV | 5 | USA | GER | 18 | 27 |
| Swaziland | SWA | 1 | ZAF | IND | 33 | 17 | Nigeria | NGA | 5 | IND | BRA | 27 | 18 |
| Madagascar | MDG | 1 | FRA | USA | 75 | 9 | Burkina Faso | BFA | 5 | SWI | CHI | 71 | 52 |
| Algeria | MDG | 2 | ESP | ITA | 27 | 28 | Senegal | SEN | 5 | SWI | IND | 27 | 3 |
| Lybia | LYB | 2 | ITA | FRA | 27 | 72 | Benin | BEN | 5 | BFA | CHI | 52 | 27 |
| Morocco | MAR | 2 | ESP | FRA | 85 | 87 | Liberia | LIB | 5 | CHI | POL | 26 | 89 |
| Tunisia | TUN | 2 | FRA | ITA | 85 | 62 | Mali | MAL | 5 | SWI | CHI | 52 | 71 |
| Mauritania | MRT | 2 | CHI | SWI | 26 | 3 | Niger | NIG | 5 | FRA | BFA | 26 | 27 |
| Cameroon | CMR | 3 | ESP | CHI | 27 | 18 | Togo | TOG | 5 | BFA | LEB | 52 | 39 |
| Chad | CHA | 3 | USA | JAP | 27 | 52 | Sierra Leone | SLe | 5 | CHI | BEL | 26 | 71 |
| C.African R. | CAR | 3 | CHI | IDN | 44 | 52 | Cabo Verde | CaV | 5 | ESP | POR | 3 | 16 |
| Congo | COG | 3 | CHI | ITA | 27 | 89 | Guinea | GIN | 5 | KOR | IND | 27 | 26 |
| Eq.Guine | EqG | 3 | CHI | UK | 27 | 29 | | | | | | | |

Table 1: African countries and their classification into regional organizations, their main exporting commodities and their leading destinations of exports in 2014. Source: International Trade Map (http://www.trademap.org) ([16]).

## 2.1 The Destinations of Exports

The following list of 28 countries (Countries$_{14}$) that imported from Africa in 2014 on a first and second destination basis (as just the first and the second

main destinations of exports of each country were taken) are grouped in five partition clusters: »**African Countries**«, »**USA, »China**«, »**Europe**« and »**Other**«.

1. African Countries

| Zambia(ZMB) | Tanzania(TZA) | Botswana(BWA) | South Africa(ZAF) | Rwanda(RWA) | B.Faso(BFA) |
|---|---|---|---|---|---|

2. USA

3. China

4. Europe

| France(FRA) | Switzerland(SWI) | Netherlands(NET) | Italy(ITA) | Poland(POL) |
|---|---|---|---|---|
| United Kindom(UK) | Spain(ESP) | Portugal(POR) | Belgium(BEL) | Germany(GER) |

5. Other

| Malaysia(MAS) | India(IND) | Emirates(EMI) | Turkey(TUR) | Brazil(BRA) |
|---|---|---|---|---|
| Korea(KOR) | Japan(JAP) | Indonesia(IDN) | Lebanon(LEB) | Pakistan(PAK) |

Figure 1 shows the distribution of the frequencies of the two leading destinations of exports of each country in Table 1. The first histogram (right) in Figure 1 allows for the observation of the leading destinations of exports from Africa in 2014 and to the way they are distributed by countries. It also shows the distribution of the frequencies (left plot) of the first and second destinations when frequencies are aggregated in the five partition clusters above described.

The first plot shows that the top-five destinations of African exports in 2014 were China, South Africa, Switzerland, France and India. China holds the highest frequency, being followed by India and by two EU countries (Switzerland and France). The second histogram shows that when frequencies are aggregated in five partition clusters, »Europe« holds the highest frequency, being followed by »China«.
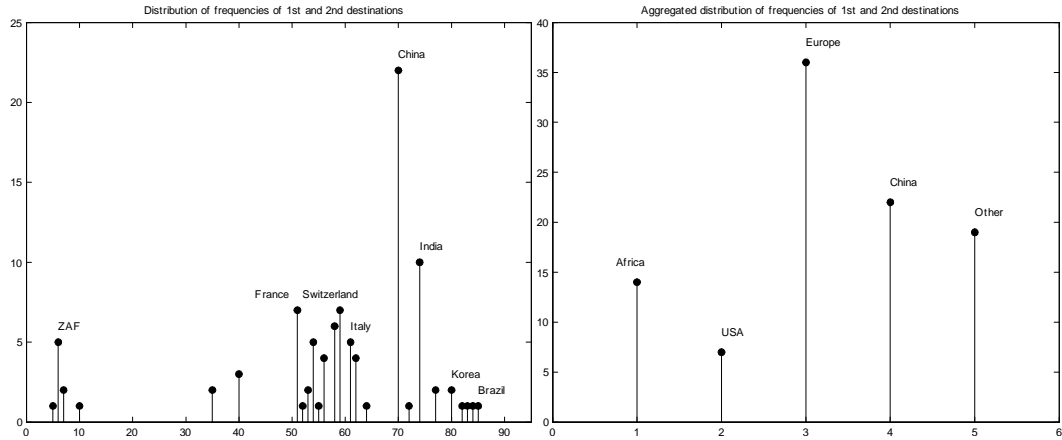
Figure 1: The distribution of the frequencies of the two leading destinations by country and the same distribution when destinations are aggregated in five partition clusters.

## 2.2 The Exporting Commodities

The following list of 27 commodities (Commodities$_{14}$) imported from Africa in 2014 on a first and second product basis (as just the first and the second main export commodities of each country were taken) are aggregated in five partition clusters: »**Petroleum**«, »**Raw Materials**«, »**Diamonds**«, »**Manufactured Products**« and »**Other Raw Materials**«.

1. Petroleum: HS code: 27(Oil Fuels)

2. Raw Materials (HS code)

   | 03 (fish)   | 06(trees)  | 08(fruit)   | 09(coffee) | 10(cereals) | 16(meat)   |
   |-------------|------------|-------------|------------|-------------|------------|
   | 17(sugars)  | 18(cocoa)  | 24(tobacco) | 33(oils)   | 44(wood)    | 52(cotton) |

3. Diamonds: HS code: 71(Pearls)

4. Manufactured Products (HS code)

   | 28(inorg.chemic.) | 29(org.chemic.) | 39(plastics)  | 61(art.apparel) | 62(art.apparel) | 72(iron-steel) |
   |-------------------|-----------------|---------------|-----------------|-----------------|----------------|
   | 74(copper)        | 75(nickel)      | 76(Aluminium) | 85(electricals) | 87(vehicles)    | 89(boats)      |

5. Other Raw Materials: HS code: 26(Ores)

7

Figure 2 shows the distribution of the frequencies of the two leading export products of each country in Table 1. The second plot (left) shows the same distribution when the products are aggregated according to the five partitions of commodities above presented.
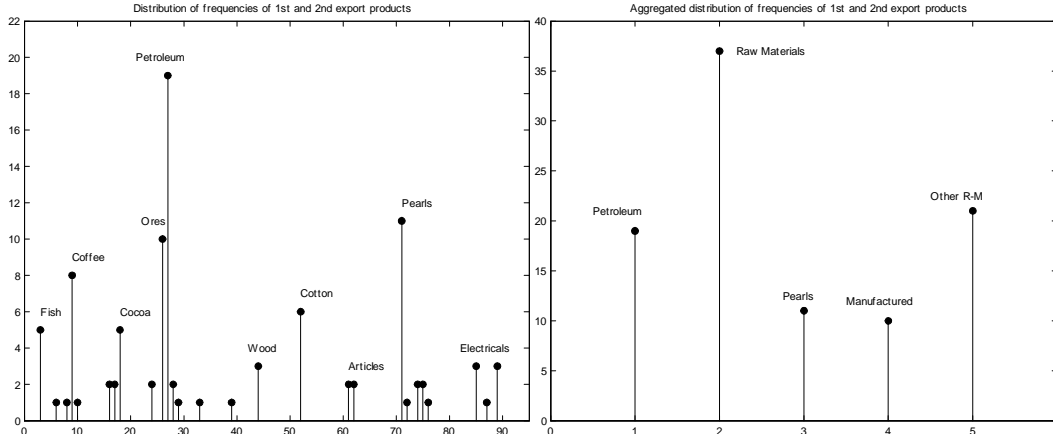


Figure 2: The distribution of the frequencies of the two leading export commodities by country and the same distribution when commodities are aggregated in five partition clusters.

From the first histogram in Figure 2, one is able to observe the most requested products exported from Africa in 2014 and to the way their frequencies are distributed. Not surprisingly, Petroleum crude holds the highest frequency, being followed by Pearls, Ores and by Coffee. The top-five exporting commodities in 2014 when frequencies are aggregated are »Raw Materials«, and »Other Raw Materials«, with which »Petroleum« shares a similar frequency.

## 3    Methodology

Network-based approaches are nowadays quite common in the analysis of systems where a network representation intuitively emerges. It often happens in the study of international trade networks.

As earlier mentioned, the choice of a given network representation is only one out of several other ways to look at a given system. There may be many

8

ways in which the elementary units and the links between them are defined. Here we define two independent bipartite networks where trade similarities between each pair of countries are used to define the existence of every link in each of those networks.

In this section, we analyze the projections of those bipartite networks, the earlier described $\text{DSN}_{14}$ and $\text{CSN}_{14}$ networks are weighted graphs and the weight of each link corresponds to the intensity of the similarity between the linked pair of countries. In the next section, the weighted networks are further analyzed through the construction of their corresponding minimal spanning tress (MST). In so doing, we are able to emphasize the main topological patterns that emerge from the network representations and to discuss their interpretation in the international trade setting.

## 3.1   Bipartite Graphs

A bipartite network $B$ consists of two partitions of nodes $V$ and $W$, such that edges connect nodes from different partitions, but never those in the same partition. A one-mode projection of such a bipartite network onto $V$ is a network consisting of the nodes in $V$; two nodes $v$ and $v\prime$ are connected in the one-mode projection, if and only if there exist a node $w \in W$ such that $(v, w)$ and $(v\prime, w)$ are edges in the corresponding bipartite network $(B)$.

In the following, we explore two bipartite networks and their corresponding one-mode projections, the earlier described $\text{DSN}_{14}$ and $\text{CSN}_{14}$ networks.

### 3.1.1   Topological Coefficients

The adoption of a network approach provides well-known notions of graph theory to fully characterize the structure of the projections $\text{DSN}_{14}$ and $\text{CSN}_{14}$. These notions are formally defined as topological coefficients. Here, we concentrate on the calculation of five coefficients. Three of them are quantities related to averages values of one topological coefficient defined at the node level, as the network degree $\langle k \rangle$, the betweenness centrality $\langle B \rangle$ and the average clustering coefficient $\langle C \rangle$. The other two coefficients are measured at the network level, they are the density $(d)$ of the network and the network diameter $(D)$.

1. the average degree $(\langle k \rangle)$ of a network measures the average number of links connecting each element of the network.

2. the betweenness centrality ($\langle B \rangle$) measured as the fraction of paths connecting all pairs of nodes and containing the node of interest ($i$).

3. the clustering coefficient ($\langle C \rangle$) measures the average probability that two nodes having a common neighbor are themselves connected

$$C_i = \frac{E(v_i)}{v_i(v_i - 1)} \tag{1}$$

where $E(v_i)$ is the size of the neighbourhood ($v_i$) of the node $i$ and the neighbourhood of $i$ consisting of all nodes adjacent to $i$.

4. the diameter of the network ($\langle D \rangle$) measuring the shortest distance between the two most distant nodes in the network.

5. the density ($0 \leq d \leq 1$) of the network is the ratio of the number of links in the network to the number of possible links

$$d = \frac{2L}{n(n-1)} \tag{2}$$

where $L$ is the number of links and $n$ is the number of nodes.

Here, these coefficients are computed for different sub-graphs of both the $DSN_{14}$ and the $CSN_{14}$. The nodes (countries) in each sub-graph are grouped accordingly to the partition clusters of main destinations (»African Countries«, »USA, »China«, »Europe« and »Other«) and by partition of main exporting commodity (»Petroleum«, »Raw Materials«, »Diamonds«, »Manufactured Products« and »Other Raw Materials«). We also apply these measures to the partition clusters defined by the regional organizations to which the countries belong (SADC, UMA, CEEAC, COMESA and CEDEAO). In so doing, it is possible to compare in terms of topological coefficients the different structures of the $DSN_{14}$ and the $CSN_{14}$ networks. For each of them, the topological coefficients are computed at the node level and then averaged by partitions of interest (main destination, main commodity or regional organization).

## 3.2 First Results

As a first approach and from the 49 African countries in Table 1 we start by developing the $DSN_{14}$ of 2014. Then, and from the same set of countries and the same year, we develop the $CSN_{14}$.

### 3.2.1 Connecting countries by a mutual export destination

The bipartite network $DSN_{14}$ consists of the following partitions:

- the set $A$ of 49 African countries presented in Table 1 and

- the set of Countries$_{14}$ (Section 2.1) to which at least one of the countries in $A$ had exported in 2014 on a first and second main destination basis.

As such, in the $DSN_{14}$ two countries are linked if and only if they shared a mutual destination of exports in 2014 among their two main export destinations. We have considered the two main destinations of exports of each country in Table 1 (columns »Destinations«). Otherwise, if just the main destination was considered, the resulting $DSN_{14}$ would comprise a set of disconnected and complete sub-graphs as, by definition, each country has just one main destination. Therefore, links in the $DSN_{14}$ are weighted by the number of coincident destinations a pair of countries share (among the two main destinations of each country in the pair), consequently, every link $L_{(i,j)}$ in the $DSN_{14}$ takes value in the set $\{0, 1, 2\}$.

As an example, $L_{(AGO,ZAF)} = 2$ since AGO and ZAF share two main destinations of exports in 2014: China and USA.

Another example is $L_{(CMR,CAV)} = 1$ due to CMR and CaV mutual destination of exports to ESP in 2014.

Among the many examples of missing links there are the cases of AGO and KEN ($L_{(AGO,KEN)} = 0$) and AGO and CaV ($L_{(AGO,CAV)} = 0$) since neither AGO and KEN nor AGO and CaV share any mutual leading destination of exports in 2014.

Figure 3 presents the $DSN_{14}$, a network of 49 African countries linked by mutual leading destinations of exports in 2014. Nodes are colored according to the partition cluster to which their main export destination belongs: red nodes identify countries whose main export destination is »China«, yellow for »Europe«, green for »USA«, blue for »African countries« and purple for the cluster of »Other«.

Such a partition of the set of countries into five clusters - defined by the country main destination of exports - allows for computing the average values of some topological coefficients by partition cluster.

In so doing, and as we shall see later in the paper, it is possible to compare important patterns coming out from the $DSN_{14}$ and to evaluate the extent to which the emerging patterns relate to geographic, regional or economic concerns.
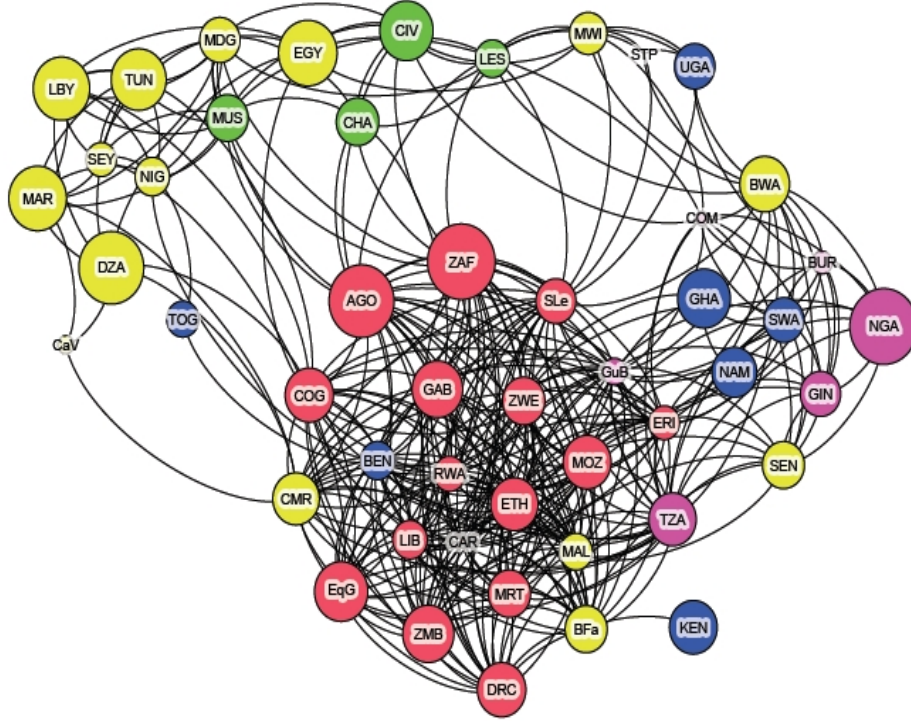
Figure 3: The DSN$_{14}$ colored by partition of main destination.

In all networks presented in this paper, the size of each node is proportional to the export value of the country in 2014. Therefore, the largest nodes are ZAF, AGO, DZA and NGA since these countries hold the largest amounts of export values in 2014. Figure 3 shows that the highest connected nodes are those whose main export destination is »China« (colored red), showing that these countries are those that share (with other countries in the whole network) the highest numbers of mutual destinations of exports. On the other hand, there are countries like TOG and STP that share very few mutual destinations with any other countries in the network.

The remarkable (red) bulk of highly connected countries in the »China« destination cluster is the very first pattern coming out from our approach. It is followed by another interesting result that clusters together the exporters to »Europe« (colored yellow), in the left upper side of Figure 3. There, another (small) cluster of exporters to »USA« (colored green) can also be seen. On the other hand, countries that have »African countries« (purple) as the first destination of exports seem to be the less clustered in the DSN$_{14}$,

being followed by those that export mainly to countries in the partition of
»Other« (colored blue) like TOG and STP.

Another evidence coming out from the network in Figure 3 is that, excluding NGA, the strongest connected countries coincide with the countries with the highest amounts of export values in 2014 (the larger nodes). Not surprisingly, it shows a positive non-negligible correlation between the amount of exports of a country and its weighted degree in the $DSN_{14}$: the countries with the highest amounts of exports in 2014 tend to be those that cluster as exporters to the most frequent African export destinations: »China« and »Europe« (most of the large nodes are yellow and red nodes).

Table 2 shows some topological coefficients computed for each node of the $DSN_{14}$ and averaged by partition of main destination. The second column (»**Size**«) shows the number of countries in each partition. The averages of the weighted degree ($\langle k \rangle$), betweenness ($\langle B \rangle$) and clustering ($\langle C \rangle$) of each partition show that there is a remarkable clustering in the »China« partition and that although the »Europe« cluster has the same number of nodes, its values of $\langle k \rangle$, $\langle B \rangle$ and mostly $\langle C \rangle$ are far below those of the »China« partition, confirming the relevance of the partition of »China« exporters as the very first pattern coming out from our approach.

| Partition | Size | $\langle k \rangle$ | $\langle B \rangle$ | $\langle C \rangle$ |
|---|---|---|---|---|
| $Africa$ | 7 | 15.8 | 22.3 | 0.78 |
| $USA$ | 4 | 11.6 | 12.4 | 0.73 |
| $Europe$ | 16 | 19.4 | 27.4 | 0.73 |
| $China$ | 16 | 25.5 | 32.1 | 0.96 |
| $Other$ | 6 | 14.5 | 17.5 | 0.68 |

Table 2: The $DSN_{14}$ topological coefficients averaged by partition of main destination.

To have an idea on the influence of regional concerns in the patterns that come out in the $DSN_{14}$, Figure 4 shows the same network of Figure 3 but now the colors of the nodes are defined by their regional organizations: blue for SADC countries, green for UMA, yellow for countries in the CEEAC, red for those in COMESA and purple for the countries in CEDEAO.

At a first glance, the degree (number of links) of each country and the general pattern of their connections do not seem to be conditioned by any regional concern since the bulk of countries in the most connected part of the network comprises countries of different regional organizations. However, there is a slightly negative difference regarding COMESA (purple) and

13

CEDEAO (red) countries. Together with UMA countries (green) they are the fewest connected countries in the entire DSN$_{14}$.
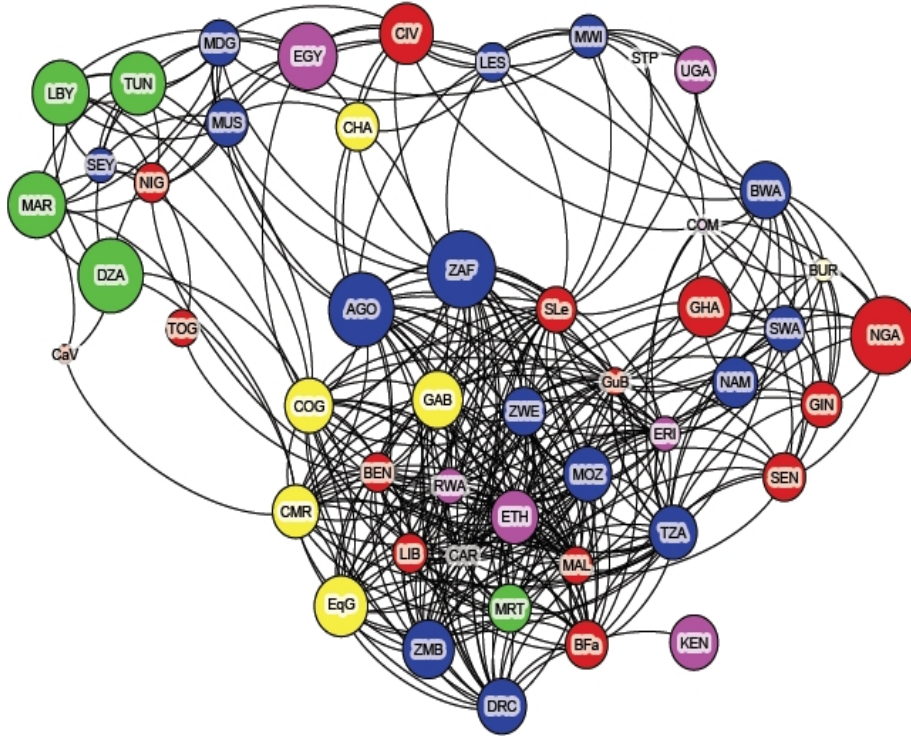


Figure 4: The DSN$_{14}$ colored by partition of regional organization.

Table 3 shows some topological coefficients computed for each node of the DSN$_{14}$ and averaged by partition of regional organization. The averages of the weighted degree ($\langle k \rangle$), betweenness ($\langle B \rangle$) and clustering ($\langle C \rangle$) of each partition show that although CEEAC partition (yellow) has just 8 countries, these countries have the greatest centrality ($\langle k \rangle$ and $\langle B \rangle$) and clustering in the whole DSN$_{14}$. This is certainly due to the fact that half of CEEAC countries has »China« as their main export destination and that besides exporting to China, the exports of CEEAC countries are concentrated in a small number of other destinations.

Likely CEEAC countries, SADC members also display a high value of betweenness centrality ($\langle B \rangle$), which is certainly related to the fact that half of the countries in these regional organizations occupy positions in the bulk of highly connected countries in the »China« partition at the bottom of Figures

3 and 4. In contrast, UMA countries display low betweenness centrality, showing that besides having »Europe« as their main export destination, the second destination of exports of UMA countries is spread over several countries.

| Partition | Size | $\langle k \rangle$ | $\langle B \rangle$ | $\langle C \rangle$ |
|---|---|---|---|---|
| $SADC$ | 15 | 17.2 | 26.7 | 0.7 |
| $UMA$ | 5 | 12.2 | 6.5 | 0.65 |
| $CEEAC$ | 8 | 17 | 25.9 | 0.86 |
| $COMESA$ | 7 | 14.5 | 13 | 0.62 |
| $CEDEAO$ | 14 | 14.8 | 16.6 | 0.83 |

Table 3: The DSN$_{14}$ topological coefficients averaged by partition of regional organization.

### 3.2.2 Connecting countries by a mutual export commodity

Here we develop the commodity share network (CSN$_{14}$) where African countries in Table 1 are the network nodes and the intensity of a link between each pair of them depends on the number of commodities that they share as export products in 2014.

The bipartite network CSN$_{14}$ consists of the following partitions:

- the set $A$ of 49 African countries presented in Table 1 and

- The set of commodities Commodities$_{14}$ (Section 2.2) that at least one of the countries in the first partition have exported in 2014 on a first and second commodity basis.

Therefore, in the CSN$_{14}$ two countries are linked if and only if they shared a mutual leading export commodity in 2014. We have considered the two main export products of each country in Table 1 (columns »Products«). Otherwise, if just the main product was considered, the resulting CSN$_{14}$ would comprise a set of disconnected sub-graphs as each country has just one main exporting product. Links in the CSN$_{14}$ are weighted by the number of coincident products a pair of countries share (among the two main products), consequently, every link $L_{(i,j)}$ in CSN$_{14}$ takes value in the set $\{0, 1, 2\}$.

As an example, the intensity of the link between KEN and UGA equals two ($L_{(KEN,UGA)} = 2$) since KEN and UGA share two mutual leading export products in 2014, they are Petroleum and Coffee. As another example,

$L_{(ERI,RWA)} = 1$ due to ERI and RWA mutual leading exports of Ores in 2014. Among the many examples of missing links there are the cases of MOZ and KEN ($L_{(MOZ,KEN)} = 0$) since MOZ and KEN did not share any mutual leading export product in 2014.

Figure 5 presents the $CSN_{14}$, a network of 49 African countries linked by at least one mutual leading export commodity in 2014. Nodes are colored accordingly to the partition cluster to which their main exporting product belongs: blue nodes have »Petroleum« as the main export commodity in 2014, red nodes identify countries whose main export products are »Manufactured«, yellow for »Diamonds«, green for »Raw Materials« and purple for the cluster of »Other Raw Materials«.

Similarly to what was done within the $DSN_{14}$, the partition of the set of countries into five clusters by main export commodity allows for computing the average values of some topological coefficients by partition cluster. In so doing it is possible to compare important patterns coming out from the $CSN_{14}$. Like in the $DSN_{14}$ representations, the size of each node is proportional to the export value of the country in 2014.

Likewise observed in the $DSN_{14}$, the graph in Figure 5 suggests the existence of a positive and strong correlation between the amount of exports of a country and its weighted degree in the $CSN_{14}$: the countries with the highest amounts of exports in 2014 (the larger nodes) tend to be those that cluster as petroleum exporters (blue), being followed by those that export »Diamonds« (yellow) and »Manufactured« (red).

The highest degrees belong to AGO, NGA and ZAF, showing that these countries are those that share (with other countries in the whole network) the highest numbers of mutual export commodities. On the other hand, there are countries like MWI and SWA that share just one exporting commodity with the all other countries in the network.

In the $CSN_{14}$ of Figure 5, the bulk of highly connected countries are placed at the right side. It corresponds to the cluster of »Petroleum« exporters (blue), a highly clustered and almost fully connected set of nodes. It is followed by the cluster of »Diamonds« exporters (yellow) at the left side. The very first pattern coming out from our $CSN_{14}$ is the remarkable centrality of AGO.
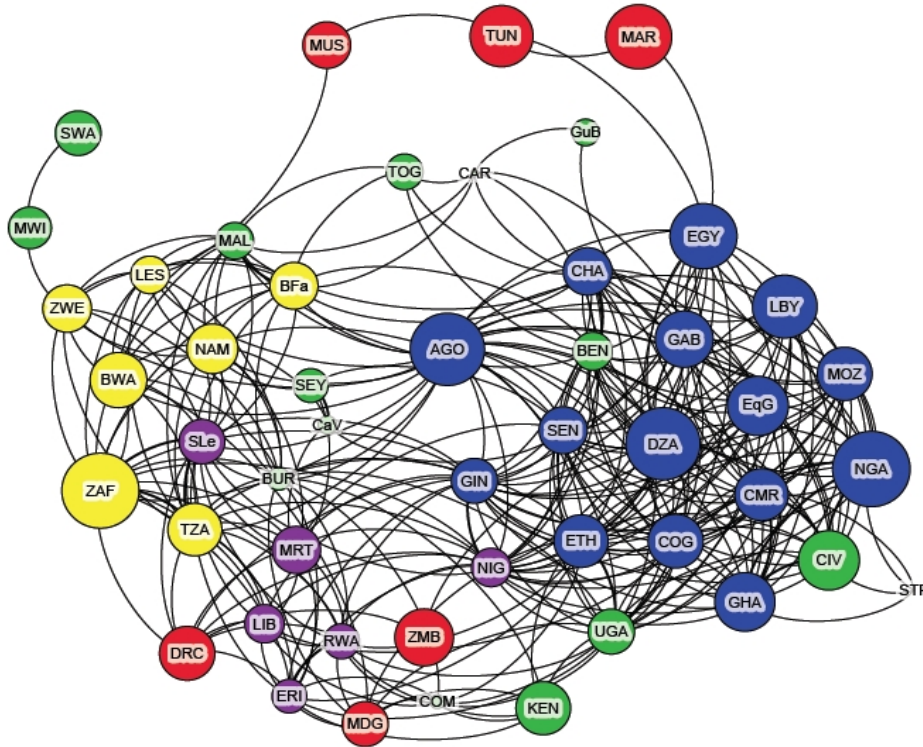
Figure 5: The $CSN_{14}$ colored by partition of main export commodity.

Table 4 shows some topological coefficients computed for each node of the $CSN_{14}$ and averaged by partition of main exporting product. The averages of the weighted degree ($\langle k \rangle$), betweenness ($\langle B \rangle$) and clustering ($\langle C \rangle$) of each partition show that the strongest connected countries are those whose main exporting product is »Petroleum«. In terms of connectivity, they are followed by the cluster of Manufactured Products, even though this cluster has just six countries.

There is a very high value of betweenness centrality ($\langle B \rangle$) characterizing countries in the (small) »Diamonds« cluster, being certainly led by the centrality of ZAF. Although this cluster has just seven countries, it displays the second highest betweenness in the ranking of partitions. In so doing, we are informed that the countries whose main exporting commodity is »Diamonds« have a large sharing of common export products with other countries in the whole network. The same applies to the cluster of »Petroleum« exporters.

17

| Partition | Size | $\langle k \rangle$ | $\langle B \rangle$ | $\langle C \rangle$ |
|---|---|---|---|---|
| $Petroleum$ | 15 | 21.6 | 35.5 | 0.82 |
| $Raw Materials$ | 15 | 9.9 | 14.1 | 0.72 |
| $Diamonds$ | 7 | 14.2 | 33.7 | 0.70 |
| $Manufactured$ | 6 | 18.6 | 22.3 | 0.77 |
| $Other RM$ | 6 | 4.5 | 5.8 | 0.48 |

Table 4: The $CSN_{14}$ topological coefficients averaged by partition of main exporting product.

Another interesting characteristic is the poor connectivity pattern of countries in the »Raw Materials« partition cluster. Even being a large partition in size, its average weighed degree ($\langle k \rangle$) is the second smallest in the $CSN_{14}$. It means that countries that mainly export raw materials have a small share of mutual export products with other countries.
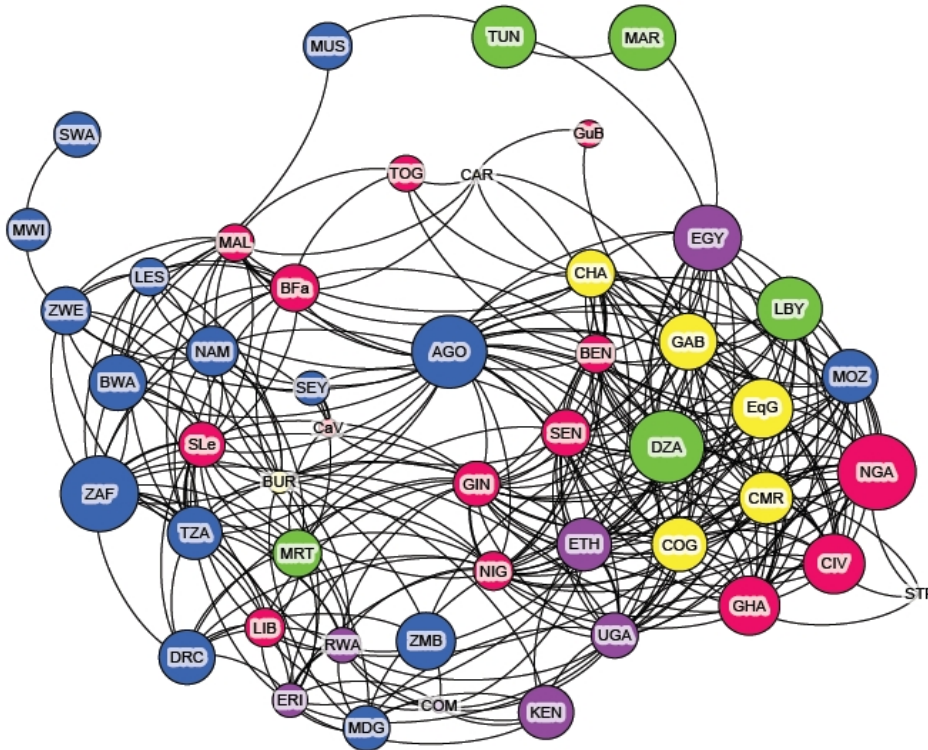


Figure 6: The $CSN_{14}$ colored by partition of regional organization.

When a regional perspective is taken, the $CSN_{14}$ in Figure 6 is the same network presented in Figure 5 but nodes are colored according to their

18

regional organizations: blue for SADC countries, green for UMA, yellow for countries in the CEEAC, red for those in COMESA and purple for the countries in CEDEAO.

Table 5 shows some topological coefficients computed for each node of the $CSN_{14}$ and averaged by partition of regional organization. There, the CEEAC partition (yellow) although having just eight countries, has the greatest clustering ($\langle C \rangle$) in the whole $CSN_{14}$. On the contrary, the SADC cluster, although comprising a large number of countries, is the one of the poorest clustering value in the ranking of five partitions presented in Table 5.

Another evidence coming out from the regional perspective is that without AGO and MOZ, the large set of fifteen SADC countries (blue) would be excluded from the large cluster of »Petroleum« exporters. This is certainly related to the small clustering that SADC countries have in the $CSN_{14}$, being most of SADC countries mainly diamond and ore exporters.

| Partition | Size | $\langle k \rangle$ | $\langle B \rangle$ | $\langle C \rangle$ |
|---|---|---|---|---|
| $SADC$ | 15 | 10.7 | 27.7 | 0.56 |
| $UMA$ | 5 | 11 | 11.7 | 0.76 |
| $CEEAC$ | 8 | 16.3 | 19 | 0.8 |
| $COMESA$ | 7 | 16.8 | 29.6 | 0.72 |
| $CEDEAO$ | 14 | 16.9 | 21.4 | 0.78 |

Table 5: The $CSN_{14}$ topological coefficients averaged by partition of regional organization.

### 3.2.3 Comparing the destination share and the commodity share networks

Table 6 shows some network coefficients computed for the $DSN_{14}$ and the $CSN_{14}$. The coefficients $\langle k \rangle$, $\langle B \rangle$ and $\langle C \rangle$ were computed at the node level and averaged by network. The averages of the weighted degree ($\langle k \rangle$), betweenness ($\langle B \rangle$) and clustering ($\langle C \rangle$) of these networks show that for the 49 African countries in 2014, sharing a mutual leading export product happens less often than sharing a mutual leading destination of exports, since the degree of the $DSN_{14}$ is greater than the degree of the $CSN_{14}$.

The columns (»**density**«) and (»**diameter**«) provide values for the most typical coefficients in network analysis. These coefficients are not computed at the node level but for each (entire) network. Since the diameter of

the $CSN_{14}$ is larger than the diameter of the $DSN_{14}$, the 49 African countries are on average closer to each other when connected by a mutual leading export destination than when connected by a mutual exporting product.

| Graph | Size | $\langle k \rangle$ | $\langle B \rangle$ | $\langle C \rangle$ | $density$ | $diameter$ |
|-------|------|------|------|------|---------|----------|
| $DSN_{14}$ | 49 | 15.6 | 19.7 | 0.76 | 0.31 | 3 |
| $CSN_{14}$ | 49 | 14.3 | 23.1 | 0.72 | 0.28 | 5 |

Table 6: Comparing topological coefficients obtained for $DSN_{14}$ and $CSN_{14}$.

The $DSN_{14}$ also has a larger clustering coefficient than the $CSN_{14}$, showing that when a country shares a mutual export destination with other two countries, these two other countries also tend to share a mutual export destination between them. The densities of $DSN_{14}$ and $CSN_{14}$ confirm that topological distances in the $DSN_{14}$ are shorter than in the $CSN_{14}$ and that on average going from one country in the $DSN_{14}$ to any other country in the same graph takes less intermediate nodes than in the $CSN_{14}$.

Although the networks $DSN_{14}$ and $CSN_{14}$ inform about the degree of the nodes, their densely-connected nature does not help to discover any dominant topological pattern besides the distribution of the node's degree. Moving away from a dense to a sparse representation of a network, one shall ensure that the degree of sparseness is determined endogenously, instead of by an a priory specification. It has been often accomplished ([17],[18]) through the construction of a Minimal Spanning Tree (MST), in so doing one is able to develop the corresponding representation of the network where sparseness replaces denseness in a suitable way.

## 3.3 The Minimum Spanning Tree Approach

In the construction of a MST by the *nearest neighbor* method, one defines the 49 countries (in Table 1) as the nodes ($N_i$) of a weighted network where the distance $d_{ij}$ between each pair of countries $i$ and $j$ corresponds to the inverse of weight of the link ($d_{ij} = \frac{1}{L_{ij}}$) between $i$ and $j$.

From the $NxN$ distance matrix $D$, a hierarchical clustering is then performed using the *nearest neighbor* method. Initially $N$ clusters corresponding to the $N$ countries are considered. Then, at each step, two clusters $c_i$ and $c_j$ are clumped into a single cluster if

$$d\{c_i, c_j\} = \min\{d\{c_i, c_j\}\} \tag{3}$$

with the distance between clusters being defined by

$$d\{c_i, c_j\} = \min\{d_{pq}\} \text{ with } p \in c_i \text{ and } q \in c_j \qquad (4)$$

This process is continued until there is a single cluster. This clustering process is also known as the *single link method*, being the method by which one obtains the minimal spanning tree (MST) of a graph.

In a connected graph, the MST is a tree of $N-1$ edges that minimizes the sum of the edge distances. In a network with $N$ nodes, the hierarchical clustering process takes $N-1$ steps to be completed, and uses, at each step, a particular distance $d_{i,j} \in D$ to clump two clusters into a single one.

# 4    Results

In this section we discuss the results obtained from the MST of each one-mode projected graphs $DSN_{14}$ and $CSN_{14}$. As earlier mentioned, the MST of a graph may allow for discovering relevant topological patterns that are not easily observed in the dense original networks. As in the last section, we begin with the analysis of the $DSN_{14}$ and then proceed to the $CSN_{14}$.

We look for eventual topological structures coming out from empirical data of African exports, in order to see whether some relevant characteristics of African trade have any bearing on the network structures that emerge from the application of our approach. In the last section, we observed some slight influence of the regional position of each country in its connectivity. With the construction of the minimum spanning trees we envision that some stronger structural patterns would come to be observed on the trees.

## 4.1    The MST of the destination share network

Figure 7 shows the MST obtained from the $DSN_{14}$ and colored according to each country main destination of exports in 2014.

The first evidence coming out from the MST in Figure 7 is the central position of AGO clustering together the entire set of »China« exporters (red) in 2014. Another important pattern that emerges in the MST is the branch of UMA countries (yellow) in the right side of the tree, being »Europe« their most frequent destination of exports. Similarly, part of the countries that export mostly to »Other« seems to cluster on the left branch (purple).

Interestingly, the countries that exports to »African countries« (blue) occupy the less central positions on the tree. This result illustrates the suitability of the MST to separate groups of African countries according to their main export destinations and the show how opposite are the situations of those that export to »China« from the countries that have Africa itself as their main export destinations.
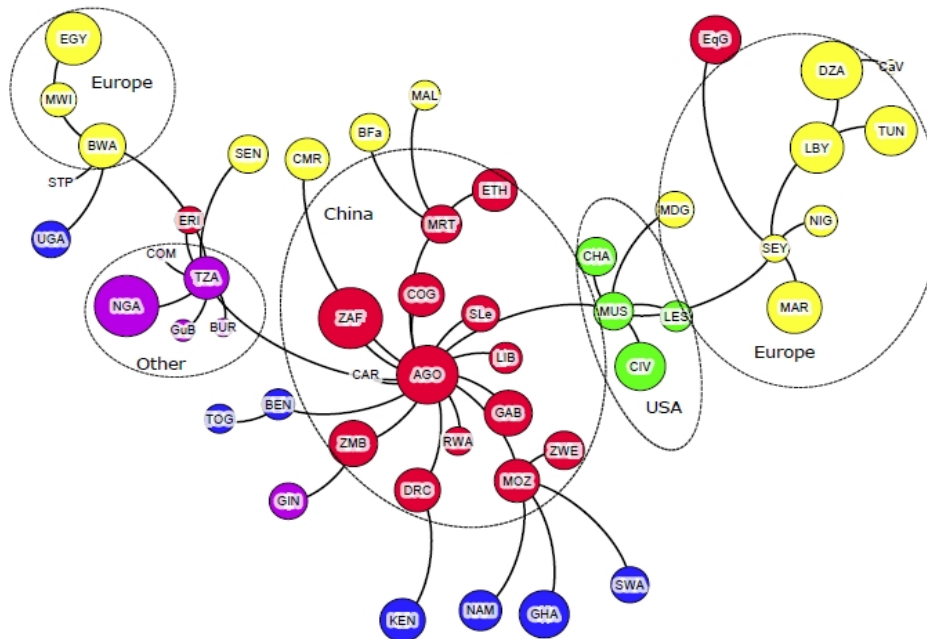


Figure 7: The MST of the $\text{CSN}_{14}$ colored by partition of main destination.

Regarding centrality, AGO occupies the most central position of the network since this country exports to the top most African export destinations (»China« and »Europe«) being therefore, and by this means, easily connected to a large amount of other countries. Indeed, AGO is the center of the most central cluster of »China« exporters. On the other hand, many leaf positions are occupied by countries that exports to other African countries as they have the smallest centrality in the whole network, they are KEN, NAM and SWA. Their weak centrality is due to the fact that their leading export destinations are spread over several countries (ZMB, TZA, ZAF, BWA and IND).

## 4.2 The MST of the commodity share network

Figure 8 shows the MST obtained from the $CSN_{14}$ and colored according to the main export commodity of each country in 2014. The first observation on the MST presented in Figure 8 is that, centrality is concentrated in a fewer number of countries (when compared to the MST of the $DSN_{14}$).
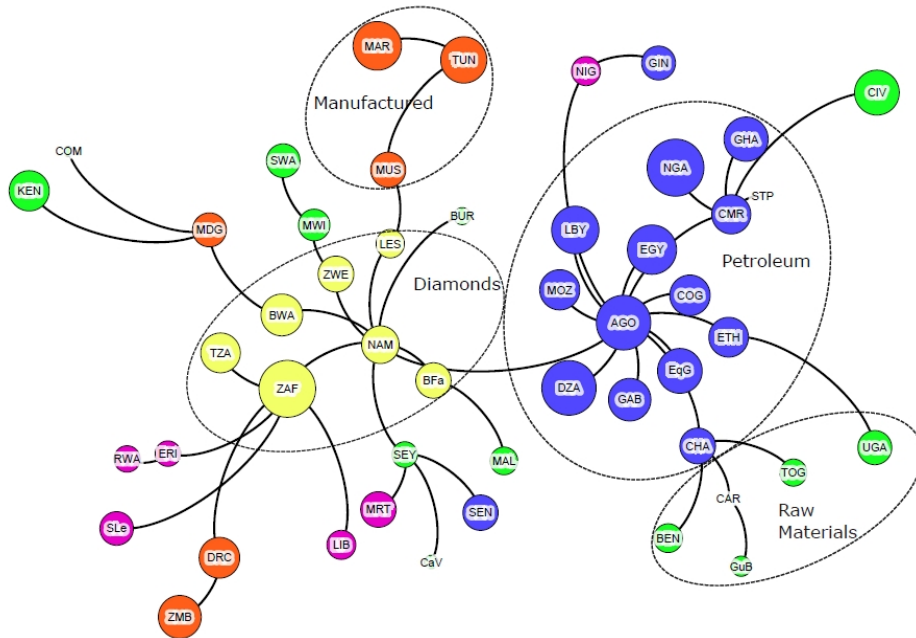


Figure 8: The MST of the $CSN_{14}$ colored by partition of main exporting commodity.

The top most central and connected positions are shared by countries belonging to two regional organizations: SADC and CEDEAO, being mainly represented by ZAF and AGO and clustering countries whose main export commodities are »Diamonds« and »Petroleum«, respectively. Unsurprisingly, centrality and connectivity advantages seem to be concentrated in these two leading commodity partitions (»Diamonds« and »Petroleum«) and organization groups (SADC and CEDEAO).

Indeed, half of CEDEAO countries occupy the upper branch rooted in ZAF and having »Diamonds« (yellow) as their main export commodity. Another regional cluster is rooted in AGO and tie together several UMA countries whose main export commodity is »Petroleum« (blue). On the other hand, half of UMA countries are far from each other on the tree, they occupy

23

the leaf positions, being weakly connected to the other African countries to which, the few connections they establish rely on having »Manufactured« as their main export commodity. Likewise, there is a branch clustering exporters of »Raw Materials« (green) being also placed at the leaf positions on the tree. Such a lack of centrality of »Raw Materials« exporters in the $CSN_{14}$ seems to be due to the fact that their leading export products are spread over many different commodities (the »Raw Materials« partition comprises 12 different commodities).

# 5  Concluding remarks

In the last decade, a debate has taken place in the network literature about the application of network approaches to model international trade. In this context, and even though recent research suggests that African countries are among those to which exports can be a vehicle for poverty reduction, these countries have been insufficiently analyzed.

We have proposed the definition of trade networks where each bilateral relation between two African countries is defined from the relations each of these countries hold with another entity. Both networks were defined from empirical data reported for 2014.They are independent bipartite networks: a destination share network ($DSN_{14}$) and a commodity share network ($CSN_{14}$). In the former, two African countries are linked if they share a mutual leading destination of exports, and in the latter, countries are linked through the existence of a mutual leading export commodity between them.

Our conclusions can be summarized in the following.

1. Sharing a mutual export destination happens more often: The very first remark coming out from the observation of both the $DSN_{14}$ and the $CSN_{14}$ is that, in 2014 and for the 49 African countries, sharing a mutual exporting product happens less often than sharing a mutual destination of exports.

2. Great exporting countries tend to be more linked: There is a positive correlation between strong connected countries in both the $DSN_{14}$ and the $CSN_{14}$ and those with high amounts of export values in 2014. It is in line with recent research placed in two different branches of the literature on international trade: the World Trade Web (WTW) empirical

exploration ([7],[8],[9],[10], [11],[12],[13]) and the one that specifically focus on African trade ([1],[2],[3],[4],[5],[6]). References ([2],[4]) reports on the role of export performance to economic growth. They also discuss on the relation between trade and development, and on the growth by destination hypothesis, according to which, the destination of exports can play an important role in determining the trade pattern of a country and its development path.

3. Destination matters: The idea that destination matters is in line with our finding that in the $DSN_{14}$, the highest connected nodes are those whose main export destination is China. According to Baliamoune-Lutz ([4]) export concentration enhances the growth effects of exporting to China, implying that countries which export one major commodity to China benefit more (in terms of growth) than do countries that have more diversified exports.

   - The China effect: One of the patterns that came out from our $DSN_{14}$ shows that half of CEEAC and SADC countries belongs to the bulk of »China« destination cluster, having high betweenness centrality. Additionally, the »China« destination group of countries displays the highest clustering coefficient (0.96), meaning that, besides having China as their main exports destination, the second destination of exports of the countries in this group is highly concentrated on a few countries.

   - The role of Intra-African trade: Another important pattern coming out from the MST of the $DSN_{14}$ shows how opposite are the situations of the countries that export to »China« from the countries that have Africa itself as their main export destinations. In the MST of the destination share network, many leaf positions are occupied by intra-African exporters as they have the smallest centrality in the whole network. Their weak centrality is due to the fact that their leading export destinations are spread over several countries. This result is in line with the results reported by reference ([2]) where the growing importance of intra-African trade is discussed and proven to be a crucial channel for the expansion of African exports. Moreover, Kamuganga found significant correlation between the participation in intra-African trade and the diversification of exports.

- The Angola cluster: our results highlighted the remarkable centrality of AGO as the center of the most central cluster of »China« exporters. Indeed, AGO is the country that holds the most central position when both $DSN_{14}$ and $CSN_{14}$ are considered. This country occupies in both cases the center of the largest central clusters: »China« exporters in $DSN_{14}$ and exporters of »Petroleum« in $CSN_{14}$.

- UMA countries anti-diversification: In the opposite situation, we found that UMA countries display very low centrality, showing that besides having »Europe« as their first export destination, the second destination of exports of UMA countries is spread over several countries. This result is in line with reference ([6]) report on European unilateral trade preferences and anti-diversification effects. We showed that UMA countries occupy a separate branch in the MST of the $DSN_{14}$, being »Europe« their most frequent destination of exports.

4. In the $CSN_{14}$, the highest connected nodes are those that cluster as »Petroleum« exporters, being followed by those that export »Diamonds«. Unsurprisingly, »Raw Materials« exporters display very low connectivity as their second main exporting product is spread over several different commodities.

5. Organizations matter: Regional and organizational concerns seem to have some impact in the $CSN_{14}$.

- SADC and Petroleum: The group of SADC countries, although comprising a large number of elements, is the one with the poorest connectivity and clustering in the $CSN_{14}$. It is certainly due to the fact that without AGO and MOZ, this large group of countries does not comprise »Petroleum« exporters.

- UMA countries anti-diversification: Again in the $CSN_{14}$, its MST shows that UMA countries are placed on a separate branch. Although they are countries with high amounts of export values in 2014, UMA countries display low connectivity and low centrality. The leaf positions in the MST of either $DSN_{14}$ or $CSN_{14}$ - while occupied by countries with very low centrality and connectivity -

were shown to characterize countries that export mainly to »Europe« and whose main exporting product is »Raw Materials«.

Future work is planned to be twofold. We plan to further improve the definition of networks of African countries, enlarging the set of similarities that define the links between countries in order to include aspects like mother language, currencies, demography and participation in trade agreements. On the other hand, we also plan to apply our approach to different time periods. As soon as we can relate the structural similarities (or differences) and their evolution in time to certain trade characteristics, the resulting knowledge shall open new and interesting questions for future research on African trade.

# References

[1] Portugal-Perez, A. and Wilson, J.S. (2008) Trade Costs in Africa: Barriers and Opportunities for Reform, *in* Policy Research Working Paper 4619.

[2] Kamuganga, Dick N. (2012) What drives Africa's export diversification?, *in* Graduate Institute of International and Development Studies Working Paper, No. 15/2012.

[3] Ackah, C. and Morrissey O. (2007) Trade Protection as Income Protection in Poor Countries *in* GEP-Murphy Institute Conference on New Political Economy of Globalization, New Orleans.

[4] Baliamoune-Lutz, M. (2011) Growth by Destination (Where You Export Matters): Trade with China and Growth in African Countries in African Development Review, Special Issue: Special Issue on the 2010 African Economic Conference on "Setting the Agenda for Africa's Economic Recovery and Long-term Growth" V.23, 2.

[5] Ackah, C. and Morrissey O. (2005) Trade policy and performance in Sub-Saharan Africa since the 1980s, *in* School of Economics, University of Nottingham: CREDIT Research Paper 07/01.

[6] Gamberoni, E. (2007) Do Unilateral Trade Preferences help Export Diversification? an Investigation of the Impact of the European Unilateral

Trade Preferences on Intensive and Extensive Margin of Trade, IHEID working paper 17.

[7] Serrano, M.; Boguná, M. and Vespignani, A. (2007) Patterns of dominant flows in the trade web, *in* J. Econ. Interact. Coord., V.2.

[8] Almog, A.; Squartini, T. and Garlaschelli, D. (2015) The Double Role of GDP in Shaping the Structure of the International Trade Network, arXiv:1512.02454.

[9] Fagiolo, G.; Reyes J. and Schiavo S. (2009) World-trade web: Topological properties, dynamics, and evolution, *in* Phys. Rev. E, V.79.

[10] Saracco, F.; Di Clemente, R.; Gabrielli, A. and Squartini, T. (2015) Detecting the bipartite World Trade Web evolution across 2007: a motifs-based analysis, arXiv:1508.03533.

[11] De Benedictis, L.; Tajoli, L. (2011) The world trade network, *in* The World Economy, V.34.

[12] Picciolo, F.; Squartini, T.; Ruzzenenti, F.; Basosi, R. and Garlaschelli, D. (2012) The role of distances in the World Trade Web, *in* Proceedings of the Eighth International Conference on Signal-Image Technology & Internet-Based Systems (SITIS 2012).

[13] Yang, Y; Poonb, J., Liua, Y. and Bagchi-Senb, S. (2015) Small and flat worlds: A complex network analysis of international trade in crude oil, *in* Energy (Elsevier) V.93.

[14] Bergstrand, J. (1985) The Gravity Equation in International Trade: some Microeconomic, Foundations and Empirical Evidence, in Review of Economics and Statistics, V. 67, 3.

[15] Araújo. T. and Banisch, S. (2014), Multidimensional Analysis of Linguistic Networks, *in* Towards a Theory of Complex Linguistic Networks, Springer, Berlin.

[16] International Trade Map (http://www.trademap.org).

[17] Araújo, T., Spelta, A. (2012), The Topology of Cross-border Exposures: beyond the minimal spanning tree approach, *in* Physica A, V.391.

[18] Dias, J. (2012) Sovereign debt crisis in the European Union: A minimum spanning tree approach, *in* Physica A, V. 391.