

# Reconciling contrasting views on economic complexity

Carla Sciarra <sup>1</sup>✉, Guido Chiarotti<sup>1</sup>, Luca Ridolfi<sup>1</sup> & Francesco Laio<sup>1</sup>

Summarising the complexity of a country's economy in a single number is the holy grail for scholars engaging in data-based economics. In a field where the Gross Domestic Product remains the preferred indicator for many, economic complexity measures, aiming at uncovering the productive knowledge of countries, have been stirring the pot in the past few years. The commonly used methodologies to measure economic complexity produce contrasting results, undermining their acceptance and applications. Here we show that these methodologies - apparently conflicting on fundamental aspects - can be reconciled by adopting a neat mathematical perspective based on linear-algebra tools within a bipartite-networks framework. The obtained results shed new light on the potential of economic complexity to trace and forecast countries' innovation potential and to interpret the temporal dynamics of economic growth, possibly paving the way to a micro-foundation of the field.

<sup>1</sup>DIATI, Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Torino, Italy. ✉email: [carla.sciarra@polito.it](mailto:carla.sciarra@polito.it)

The French politician and gastronome Jean Anthelme Brillat-Savarin in his book ‘Physiologie du goût’ wrote: “Tell me what you eat, and I will tell you who you are”<sup>1</sup>, aphorising on the fact that people’s food baskets reflect their wealth status. In the same vein, metrics of economic complexity (EC) aim at defining the socio-economic status of countries grounded on their export baskets<sup>2</sup>. Within economics, these approaches mainly serve as an alternative to more traditional economic growth theories<sup>3–8</sup> which are often blamed for shrinking the intricacy of countries’ socio-economic dynamics through simplistic assumptions<sup>9,10</sup>.

Within this new class of metrics, the productive knowledge owned by each country – which embeds capabilities, finance, technology, human capital and resources, and determines the country’s potential for economic growth – can be extracted from easy-to-find data<sup>11</sup>. Not surprisingly, a first proxy of countries’ productive knowledge is the number of products in their export baskets, i.e., their production diversity<sup>11–13</sup>. Although insufficient, since it does not account for baskets’ composition and complexity, the diversity is a necessary and relevant information to understand the trading competitiveness of countries<sup>11,14</sup>. The methodologies of economic complexity aim at improving this most obvious measure of competitiveness exploiting the information related to the sophistication of the exports and the capabilities required to produce and export a given good: countries with low productive knowledge only produce and export fewer and less sophisticated products, resulting in lower stages of competition<sup>11,15</sup>, while more competitive countries exploit their know-how and resources to diversify their export baskets<sup>11,15</sup>. By reversing this reasoning, it is thus expected that the diversification and composition of the export basket can be used to measure the countries’ and products’ economic complexity, thus posing the bases for a data-based (bottom-up) ranking of countries and products. This rationale lies at the base of the commonly used methodologies to measure the economic complexity of countries and products, namely the Method of Reflections (MR)<sup>12</sup> and the Fitness and Complexity algorithm (FC)<sup>15</sup>. In spite of their common root, these two methods radically differ in the conceptual approach to the problem and, as a consequence, in the obtained outcomes.

The MR approach measures a country’s economic complexity as the average of the complexities of the products in its export basket. In a specular manner – from which the name “reflections” –, a product’s complexity is obtained as the average of the complexities of the countries exporting it. The equations defining the two averages are coupled to obtain the Economic Complexity Index (ECI) and the Product Complexity Index (PCI), which have been shown to be the result of a linear algebra exercise<sup>11,16,17</sup>. As an effect of taking the averages, the obtained measures turn out to lose information about countries’ diversification and products’ ubiquity<sup>18</sup> (ubiquity is defined as the number of exporters for a given product<sup>11</sup>).

In contrast, Tacchella et al.<sup>15</sup> counter on the assumption of a linear relation between the products’ and countries’ complexities. In their view, the fact that a less competitive country exports a given product should unavoidably downgrade the product’s complexity, an effect that the Authors argue could only be obtained through the use of a non-linear relation. As a consequence, these authors introduce two metrics, the Fitness of countries  $F_c$  and the Quality of products  $Q_p$ , where products’ Quality non-linearly depends on the Fitness of the exporting countries (see Methods section, Eq. (11)); in contrast, the Fitness is obtained as the sum of the Qualities of the exported products. In this approach, contrarily to MR, the countries’ Fitness preserves the information on the diversification of the export baskets<sup>14,19</sup>.

It is not only the mathematics of the two approaches which is different, but also the obtained outcomes significantly diverge: as shown in Supplementary Fig. 1, the countries’ rankings obtained with  $ECI_c$  and  $F_c$  widely scatter (see Supplementary Note 1 for details on the implementation of the two algorithms). This poses an issue of practical use of the economic complexity measures, potentially undermining the very essence of the economic complexity theory. We argue that the role played by EC measures in economics and policy making (see, e.g., refs. 20–24) requires more consistency in the outcomes of different methods.

In this paper, we reconcile the MR and FC approaches by recasting them into a mathematically-sound, multidimensional framework, which allows us to recover and combine the strengths of both methods, still maintaining the relevant feature of providing countries’ and products’ rankings.

## Results

**A general framework for economic complexity.** Economic complexity approaches are grounded on the trade data collected into a bipartite network, defining exporters and products, and detailing whether and how much (in monetary value) a country exports a given product. The bipartite network is interpreted as the compact representation of the tripartite network constituted by countries-capabilities-products<sup>12,15</sup>; most applications<sup>12,15</sup> take into account only the relevant exporters in the network, where the relevance is computed according to the Relative Comparative Advantage (RCA)<sup>25</sup>. Moreover, to highlight the role played by network’s topology, the weights in the bipartite network are typically neglected, turning to a binary incidence matrix  $\mathbf{M}$  where  $M_{cp} = 1$  implies that the country  $c$  is a relevant exporter of the product  $p$  (see Methods section, Eq. (6)).

In a general framework, economic complexity theories aim at determining two properties  $X_c$  and  $Y_p$  – describing the complexity of country  $c$  and product  $p$ , respectively – by a system of coupled equations

$$\begin{cases} X_c = f(Y_1, Y_2, \dots, Y_p, M_{cp}), & p = [1, \dots, P], \\ Y_p = g(X_1, X_2, \dots, X_c, M_{cp}), & c = [1, \dots, C], \end{cases} \quad (1)$$

where  $f$  and  $g$  are linear functions and  $C$  and  $P$  are the number of countries and products considered in the analysis, respectively. To consider  $f$  and  $g$  as linear functions allows one to recast the determination of  $X_c$  and  $Y_p$  as the solutions of an eigen-problem of a suitable (approach dependent) transformation matrix  $\mathbf{W}$ , whose elements  $W_{cp}$  are derived from  $\mathbf{M}$ . In this case, these properties’ values are obtained from the following coupled linear equations:

$$\begin{cases} X_c = \frac{1}{\sqrt{\lambda}} \sum_p W_{cp} Y_p, \\ Y_p = \frac{1}{\sqrt{\lambda}} \sum_c W_{cp} X_c, \end{cases} \quad (2)$$

being  $\lambda$  the eigenvalue of the equivalent eigen-problem, such that the following relations hold

$$X_c = \frac{1}{\lambda} \sum_c \sum_{c^*} W_{cp} W_{c^*p} X_{c^*} = \frac{1}{\lambda} \sum_{c^*} N_{cc^*} X_{c^*}, \quad (3)$$

and

$$Y_p = \frac{1}{\lambda} \sum_p \sum_{p^*} W_{cp} W_{cp^*} Y_{p^*} = \frac{1}{\lambda} \sum_{p^*} G_{pp^*} Y_{p^*}. \quad (4)$$

A by-product of Eqs. (3)–(4) is that the squared, symmetric matrices  $\mathbf{N} = \mathbf{W}\mathbf{W}^T$  and  $\mathbf{G} = \mathbf{W}^T\mathbf{W}$  can be interpreted as proximity matrices for nations and products, respectively, where proximity defines similarity (for example,  $N_{cc^*} = N_{c^*c}$  describes the similarity in the export baskets between countries  $c$  and  $c^*$ , see Discussion section). Note that the set of equations in Eq. (2) involves the same transformation matrix  $\mathbf{W}$ . This entails that: the

matrix  $\mathbf{W}$  represents a weighted incidence matrix of an undirected bipartite network uniquely describing the relations between countries and products – this would no longer be true if two different matrices were used for the transformation; moreover, the feature of symmetry for the matrices  $\mathbf{N}$  and  $\mathbf{G}$  is essential to interpret them as proximity matrices, thus defining a bijective function.

The eigen-problems in Eqs. (3)–(4) have multiple solutions, provided by the eigenvalues  $\lambda_i$  and the corresponding eigenvectors of the matrices  $\mathbf{N}$  and  $\mathbf{G}$ , respectively<sup>26</sup>. In most situations, the eigenvector corresponding to the largest eigenvalue  $\lambda_1$  carries the maximum amount of information<sup>27</sup> and it is thus taken as solution (although we will demonstrate the potential of combining more eigenvectors). In complex network jargon,  $X_c$  and  $Y_p$  are (eigen-)centrality metrics in the bipartite network of countries and products<sup>19,28</sup>.

We now provide two examples of application of this general framework pertaining with the two aforementioned EC metrics, MR and FC, referring to these examples by using the superscripts  $A$  and  $B$ , respectively. The MR method is simply recast by setting  $W_{cp}^A = M_{cp}/\sqrt{k_c k_p}$  in Eq. (2), which provides the indices  $ECI_c^A$  and  $PCI_p^A$  from the transformations  $X_{c,2}^A = ECI_c^A \sqrt{k_c}$  and  $Y_{p,2}^A = PCI_p^A \sqrt{k_p}$  – being  $k_c = \sum_p M_{cp}$  the degree of the countries, i.e., their diversity, and  $k_p = \sum_c M_{cp}$  the degree of the products, i.e., their ubiquity. In this case, the first eigenvectors  $X_{c,1}^A$  and  $Y_{p,1}^A$  carry a trivial information, since they equal the square roots of the degrees,  $k_c$  and  $k_p$  (see Methods section, Eq. (8)), thus leading to unitary  $ECI_c^A$  and  $PCI_p^A$  values, discarded in the original works for being uninformative<sup>11,12</sup>. For this reason the eigenvectors  $X_{c,2}^A$  and  $Y_{p,2}^A$ , corresponding to the second largest eigenvalue, are taken by the authors as the solution of Eqs. (3)–(4)<sup>18</sup>. The mapping  $\{ECI_c, PCI_p\} \Leftrightarrow \{X_{c,2}, Y_{p,2}\}$  completely preserves the MR outcoming information.

Instead, the FC method is recast by setting  $W_{cp}^B = M_{cp}/k_c k'_p$ ,  $X_{c,1}^B = F_c/k_c$ , and  $Y_{p,1}^B = Q_p k'_p$  in Eq. (2), where  $k'_p = \sum_c M_{cp}/k_c$  (see Methods section, Eq. (13)). Differently from the MR mapping, in the case of FC, this mapping is not merely the results of algebraic manipulation, but implies a non-trivial linearisation of the relation between Quality and Fitness values (see Methods section, Eqs. (11)–(13)). Surprisingly enough, comparing the terms  $X_{c,1}^B$  and  $F_c/k_c$ , or  $X_{c,1}^B k_c$  and  $F_c$ , for the Fitness values – analogously  $Y_{p,1}^B$  and  $Q_p k'_p$  (or  $Y_{p,1}^B/k'_p$  and  $Q_p$ ) for the Quality values – this linearisation almost entirely preserves the information of the non-linearly computed values (independently of the kind of indicator used to measure correlation, Supplementary Fig. 2). Notice also that our linear formulation does not suffer from the well-known convergence problems of the iterative FC algorithm<sup>29</sup> and provides more regular solutions.

Some comments on the obtained results are due to the reader. First, the original  $ECI_c$ ,  $PCI_p$ ,  $F_c$  and  $Q_p$  variables are recovered within our general framework through simple (but non-trivial) mappings from  $X_c$  and  $Y_p$ . The use of the variables  $X_c$  and  $Y_p$  allows one to gain neatness in the mathematics, reflected by the fact that the matrices  $\mathbf{N}$  and  $\mathbf{G}$  can be considered as suitable proximity matrices containing information about the similarities among countries and products, respectively. This aspect may have important consequences on the interpretation of the economic significance of these metrics, as outlined in the Discussion section. Second, the matrices  $\mathbf{W}^A$  and  $\mathbf{W}^B$  differ for the specific scaling factors adopted on the matrix  $\mathbf{M}$ . It is hard to recognise an economic (or a mathematical) basis on how the factors are determined, and this leaves no solid ground for a potential user to

decide which approach, between MR and FC, to follow. Third, notwithstanding the differences among  $\mathbf{W}^A$  and  $\mathbf{W}^B$ , the eigenvectors  $X_{c,1}^A$  and  $X_{c,1}^B$  carry similar information (Supplementary Fig. 3), as also  $X_{c,2}^A$  and  $X_{c,2}^B$  (this is also partially true for  $Y_p$ , Supplementary Fig. 4). Therefore, the divergences between  $F_c$  and  $ECI_c$  – and corresponding outcomes – shown in Supplementary Fig. 1 should be mainly attributed to the fact that eigenvectors of different order are considered in the two approaches. Hence, the two metrics bring different information; albeit different, this information is relevant for both metrics, as demonstrated by numerous practical applications of the two approaches<sup>20–22,24,30–33</sup>.

Grounded on these considerations, we promote here an integrated measure of economic complexity, which exploits the neatness of the proposed framework. By employing the recently introduced framework to deal with multidimensional centrality<sup>28</sup>, we combine the two existing measures into unique centrality metrics unveiling the multidimensional complexity of countries and products. Either using  $\mathbf{W}^A$  or  $\mathbf{W}^B$  to develop the new integrated measure of complexity would lead to reliable and comparable results. We lean toward the use of  $\mathbf{W}^B$ , the one related to the FC method, for the following reasons: on the one hand, the first eigenvector  $X_{c,1}^A$  – from which, using Eq. (8), the unitary first eigenvector of MR is recovered – equals  $\sqrt{k_c}$ , thus carrying no added information beyond diversity (and the same holds for products); on the other hand, the last update on the MR method, named ECI+<sup>34</sup>, has been shown to be equivalent to the non-linear FC algorithm<sup>35</sup>, thus implicitly supporting the idea that FC carries more information than MR. The grounding hypotheses about the hidden capabilities of countries and on how these can be deduced by looking at the export baskets of countries upon which the EC algorithms are built – are preserved in our framework. From here on, we will thus use the matrix  $\mathbf{W}^B$  in Eq. (2) and drop the superscript  $B$  in the mathematical notation. In the following, we will focus on the analysis of countries' complexity. A similar reasoning also applies to the sophistication of products, whose details for the computation are given in the Methods section, Eqs. (18)–(19), while results are shown in Supplementary Figs. 4–5 and commented in Supplementary Note 2.

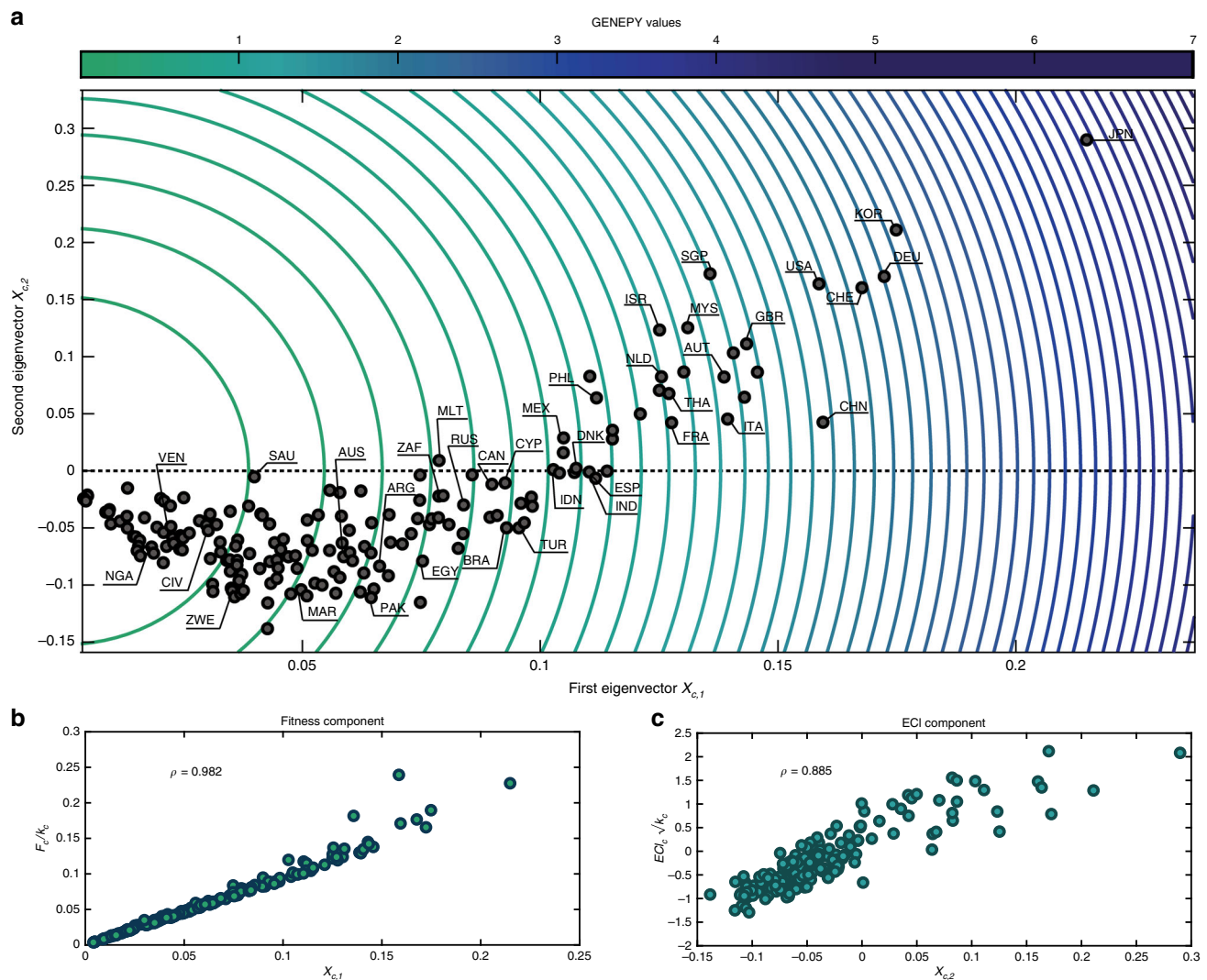
**The generalised economic complexity index.** We propose to distil the information on economic complexity into a GENERALised Economic complexitY index, GENEPY (the Genepy is a herb-based distillate typical of the north-western part of Italy). The GENEPY index for countries is defined as follows:

$$GENEPY_c = \left( \sum_{i=1}^2 \lambda_i X_{c,i}^2 \right)^2 + 2 \sum_{i=1}^2 \lambda_i^2 X_{c,i}^2, \quad (5)$$

where  $X_{c,1}$  and  $X_{c,2}$  are the eigenvectors corresponding to the first two largest eigenvalues  $\lambda_1$  and  $\lambda_2$  of the proximity matrix

$$\begin{cases} N_{cc^*} = \sum_p W_{cp} W_{c^*p} = \sum_p \frac{M_{cp} M_{c^*p}}{k_c k_{c^*} (k'_p)^2}, & \text{if } c \neq c^*, \\ N_{cc^*} = 0, & \text{if } c = c^*, \end{cases}$$

in which the redundant information of the self-proximity is deleted setting all diagonal elements of an arbitrary constant value (we set this value to zero). The rationale to compute the GENEPY index grounds on two key points: firstly, to interpret the symmetric squared matrix  $\mathbf{N}$  as the mathematical description of the weighted topology of an undirected network<sup>26</sup> – such that the countries are the nodes and the similarities between the export baskets are the links connecting them – and, secondly, to interpret the eigenvectors of  $\mathbf{N}$  as the (multidimensional) eigenvector centrality of the nodes in the network. Using this approach, the



**Fig. 1** The GENEPY index and its components. **a**  $\{X_{c,1}, X_{c,2}\}$  plane and  $GENEPY_c$  from the data of 2017 international products' trade. The x-axis reports the components of the first eigenvector  $X_{c,1}$ , whilst the y-axis the components of the second eigenvector  $X_{c,2}$ . The eigenvectors are normalised such that their Frobenius norm is unitary, i.e.,  $\sum_c X_{c,1}^2 = \sum_c X_{c,2}^2 = 1$ . Contours range from lower  $GENEPY_c$  values (green) to higher ones (blue). **b** Fitness component. Scatter plot of the first component  $X_{c,1}$  compared with the values of the Fitness values  $F_c$  rescaled by the countries degree  $k_c$  (see Methods section, Eq. (14)). **c** ECI component. Scatter plot of the second component  $X_{c,2}$  compared with  $ECI_c$  values rescaled by the term  $\sqrt{k_c}$  (see Methods section, Eq. (8)). The correlation coefficient in the plots **b** and **c** is of the Pearson's kind. Figures have been produced with MATLAB 2019b.

eigenvectors are combined into a unique metrics (the GENEPY one), following a statistically grounded framework where the same eigenvectors are obtained as the result of a least-squares estimation exercise<sup>28</sup> (Methods section, Eqs. (18)–(19); for more details we refer the reader to Sciarra et al.<sup>28</sup>).

We exemplify the use of the GENEPY index by considering the international trade of goods during the years 1995–2017<sup>36</sup>. In Fig. 1, the results are processed for the 2017 trade. Figure 1a displays the position of countries on the  $\{X_{c,1}, X_{c,2}\}$  plane. Most economies with a high drive for innovation and technology<sup>37</sup> – such as the UE-28 countries, Switzerland (CHE), China (CHN), Japan (JPN), Singapore (SGP) and the United States of America (USA) – are found far from the origin. This entails the presence of top-quality products among their exports and, therefore, of relevant productive knowledge. Less economically stable economies, such as those of many African and South-American countries, are located in the bottom-left part of the graph. The GENEPY index also identifies potentially top-competitive countries, such as Australia (AUS) and Canada (CAN), struggling to

boost their complexity due to remoteness and resources-dependency, well-known factors for affecting trade and economic growth<sup>38–40</sup>. The information distilled through the GENEPY index can be better understood by considering the meaning of its components, i.e., the two eigenvectors  $X_{c,1}$  and  $X_{c,2}$ , as contextualised in complex network theory<sup>26</sup>. In fact, the elements of the first eigenvector represent the eigenvector centrality of the countries as obtained from the proximity matrix  $N$ , interpreting the matrix as the weighted, adjacency matrix of an undirected network connecting the countries for the similarities in their export baskets (see Discussion section). Instead, the values of  $X_{c,2}$  cluster countries according to the similarities in their export baskets. In fact, the strict nexus between  $X_{c,2}$  and  $ECI_c$  recalls the results provided in Mealy et al.<sup>41</sup>, where the Authors proved that ECI perfectly solves a spectral clustering algorithm. Interpreting this result within the network of similarities designed by  $N$ , the GENEPY centrality identifies that set of capabilities (contributing to the productive knowledge) a country owns and shares with others. In this sense, more central nodes are found within a



cluster of highly competitive countries, while less complex countries are found moving towards the borders of the graph. This result is confirmed by the reordering of the matrix  $\mathbf{N}$  according to the GENE $PY$  values (Supplementary Fig. 6, Supplementary Table 1), showing that countries with higher complexity share similar sets of capabilities, as their export baskets are similar.

As mentioned, our framework combines the advantages – and information – of the two existing metrics of economic complexity, ECI and Fitness. On the one hand, the countries' Fitness values obtained with the iterative FC method are recovered, with great accuracy, from the product of the first eigenvector  $X_{c,1}$  with  $k_c$  (see Fig. 1b and Methods section, Eq. (14)). The very small deviations from the 1:1 line shown in Fig. 1b are not induced by the linearisation procedure. In fact, they disappear when the equation  $N_{cc^*} = \sum_p M_{cp} M_{c^*p} / k_c k_{c^*} (k'_p)^2$  is used also for  $c = c^*$ , i.e., when the matrix  $\mathbf{N}$  is not interpreted as a proximity matrix (Methods section, Eq. (15) and Supplementary Fig. 7). However, this would imply inflating the  $F_c$  (or  $X_{c,1}$ ) values for countries with large self-interactions, which, in our opinion, induces an undesired bias in the results. On the other hand, a good proxy of the  $ECI_c$  values is obtained by dividing the values of the second eigenvector  $X_{c,2}$  by  $\sqrt{k_c}$ , as shown in Fig. 1c (Methods section, Eq. (8)). In this case, the scatter of the plot is due to the differences in the matrices  $\mathbf{N}^A$  and  $\mathbf{N}^B$  (Methods section, Eqs. (9) and (15)), respectively.

Being the GENE $PY$  framework grounded on both existing indicators of economic complexity (the FC and MR algorithms), it inherits the intuitions and rationales upon which these two metrics are built: the capabilities of countries to export diversely complex goods are hidden within the bipartite network of countries and exports, under which they combine to maximise the complexity of the goods. Also, since  $X_{c,1}$  maintains a very high correlation with  $k_c$  (Supplementary Fig. 8), our framework preserves the information on diversification, which is a relevant one to understand how export capabilities are exploited by countries.

### The trajectories of economic growth

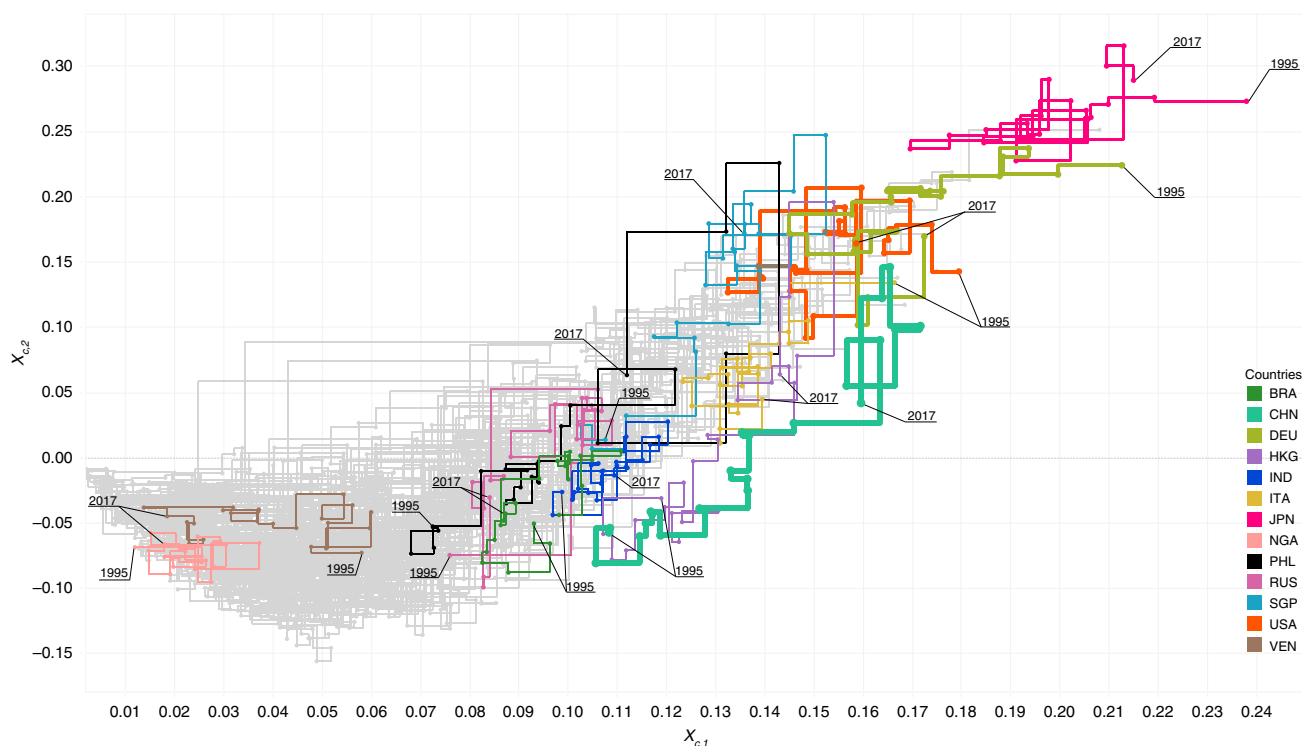
The ability of the proposed multidimensional index to assess the sophistication of countries' export-baskets and, simultaneously, define clusters of economic growth can be exploited to track the path toward prosperity of countries as driven by economic complexity. In fact, according to the economic complexity theory, a country's acquisition of capabilities, employed in the production – and hence export – of goods<sup>2,11,42</sup> is a determining factor for its economic growth. Any country at a lower stage of growth uses its increasing capabilities to fill its export basket with higher-quality goods, possibly similar to those traded by countries at higher stages of growth. This entails the creation of a wider export basket allowing the country to gain momentum in the market. Also, in order to boost its economic complexity – and growth – such a country may enlarge its offer including products for which it can be considered the only relevant exporter, hence gaining advantage<sup>4</sup>. Connecting the GENE $PY_c$  values of countries in time allows one to draw the path along this growth process, as shown in Fig. 2, in which we show some economic complexity growth paths such as the ones of China (CHN), Germany (DEU), Japan (JPN), Nigeria (NGA) and Philippines (PHL). One recognises that also the ensemble of the trajectories is knee-shaped: in fact, in each year of analysis the positions of countries in the plane  $X_{c,1}$ – $X_{c,2}$  arrange in a knee-like shape as shown in Fig. 1a for the year 2017. The presence of this shape is related to linear algebra and network science (Supplementary Note 3). By analysing the aggregated displacements of countries in time from 1995 to 2017 (for details

see Supplementary Fig. 9, Supplementary Note 4), it is possible to identify in the graph three regimes of growth. The first one is the “Impasse”. The countries that lie within this area averagely exhibit a horizontal displacement within the borders delimited by low values of  $X_{c,1}$  and negative values of  $X_{c,2}$ . Countries whose dynamics of growth lie in this area may suffer from lack of skills, human and capital investments and resources, thus resulting in low productive knowledge and, consequently, reduced diversification and complexity<sup>4</sup>. These countries hence face an impasse condition, resulting in a saddle point of growth and poor growth potential. In Fig. 2, Nigeria (NGA) and Venezuela (VEN) are tangled in this bottom-left part of the graph. The second regime is the “Bounce”. It is marked by the crossing of the zero value of the  $y$ -axis and this area defines the increment in quantity and quality of the exports. Here, the average dynamics of the countries is uplifting toward higher stages of growth. Countries such as China (CHN), India (IND) and Singapore (SGP) have clearly boosted their complexity to higher levels during the last years, joining the rich countries cluster ( $X_{c,2} > 0$ ) during the period of observation (1995–2017). The third regime is the “Arena”. Once in the economically advanced cluster, countries can play in the arena of competitiveness, where the GENE $PY$  index of some countries increases in time, that of others follows a decreasing path, instead. In fact, in this area countries aim at increasing the sophistication and the quantity of their exports which contribute to the increase of the  $X_{c,1}$  values; at the same time, countries compete to become leaders in the economically grown group, hence earning scores on  $X_{c,2}$ . However, the entrance of new countries in the competitive market is likely to affect other countries' growth. This area includes Japan (JPN), USA, Germany (DEU) and Switzerland (CHE) as paradigmatic examples.

Therefore, during their economic growth process, countries tend to move from lower stages of complexity, delimited within the bottom-left quadrant, to higher ones, framed into the top-right quadrant. The former stage is associated with low productive knowledge and, consequently, low diversity in the exports. Contrarily, the latter is characterised by gain in skills and capital's investments, for which competition and growth are determined.

In Fig. 2, the interactions among countries are also evident. The rapid growth of a country, such as the dynamics shown by China<sup>37,43</sup>, naturally impacts other economies, whose GENE $PY_c$  values change according to the increased complexity of the competitor. An example is given by the nested trajectories of the arena-countries, such as Germany, Japan and USA, concurrent with the raise of China and Singapore. Some steadiness points in the trajectories can also be explained by the economic history of the countries. For example, the reduced trade capacity of countries, as a consequence of the 2008 financial crisis<sup>44</sup>, produces a drop in complexity as shown by Germany, Italy and USA among the others. Instead, the Chinese last downgrading points of 2016–2017 may be explained as spillover effects of the 2015 stock market crash<sup>45</sup> and could also be related to the largely debated hard landing of the Chinese economy of the last years<sup>46</sup>.

To collapse the information on how countries' rankings evolve in time we compute, for each year in the period 1995–2017, the world's centre of GENE $PY$  by weighting all countries' geographical barycentres by their GENE $PY_c$  values. This computation has been executed according to the procedure defined by the McKinsey Global Institute in Dobbs et al.<sup>47</sup> to compute the shift during history of the Gross Domestic Product (GDP); the outcomes are shown in Fig. 3 in yellow. For comparison, in Fig. 3 we replicate the same procedure to compute the trajectories of the world's barycentre by weighting the countries' barycentres by their GDP at Purchasing Power Parity (GDP PPP in blue) – and, alternatively, their population (in purple). Since the economic complexity metrics are intensive ones (i.e., their values are “per



**Fig. 2 Countries' trajectories in the GENEPLY plane.** The values of the first eigenvector  $X_{c,1}$  are on the x-axis, whilst on the y-axis the values of the second eigenvector  $X_{c,2}$  are found. The eigenvectors are normalised such that their Frobenius norm is unitary, i.e.,  $\sum_c X_{c,1}^2 = \sum_c X_{c,2}^2 = 1$ . We highlight the trajectories of Brasil (BRA), China (CHN), Germany (DEU), Hong Kong (HKG), India (IND), Italy (ITA), Japan (JPN), Nigeria (NGA), Philippines (PHL), Russia (RUS), Singapore (SGP), United States of America (USA) and Venezuela (VEN), against the background created by trajectories of all other countries in grey. Line width reflects the countries' share of world exports in monetary value during 2017. To improve the readability of the plot, the paths from one point to another were forced to follow right-angled movements. The figure has been produced with Tableau Public 2019.4.

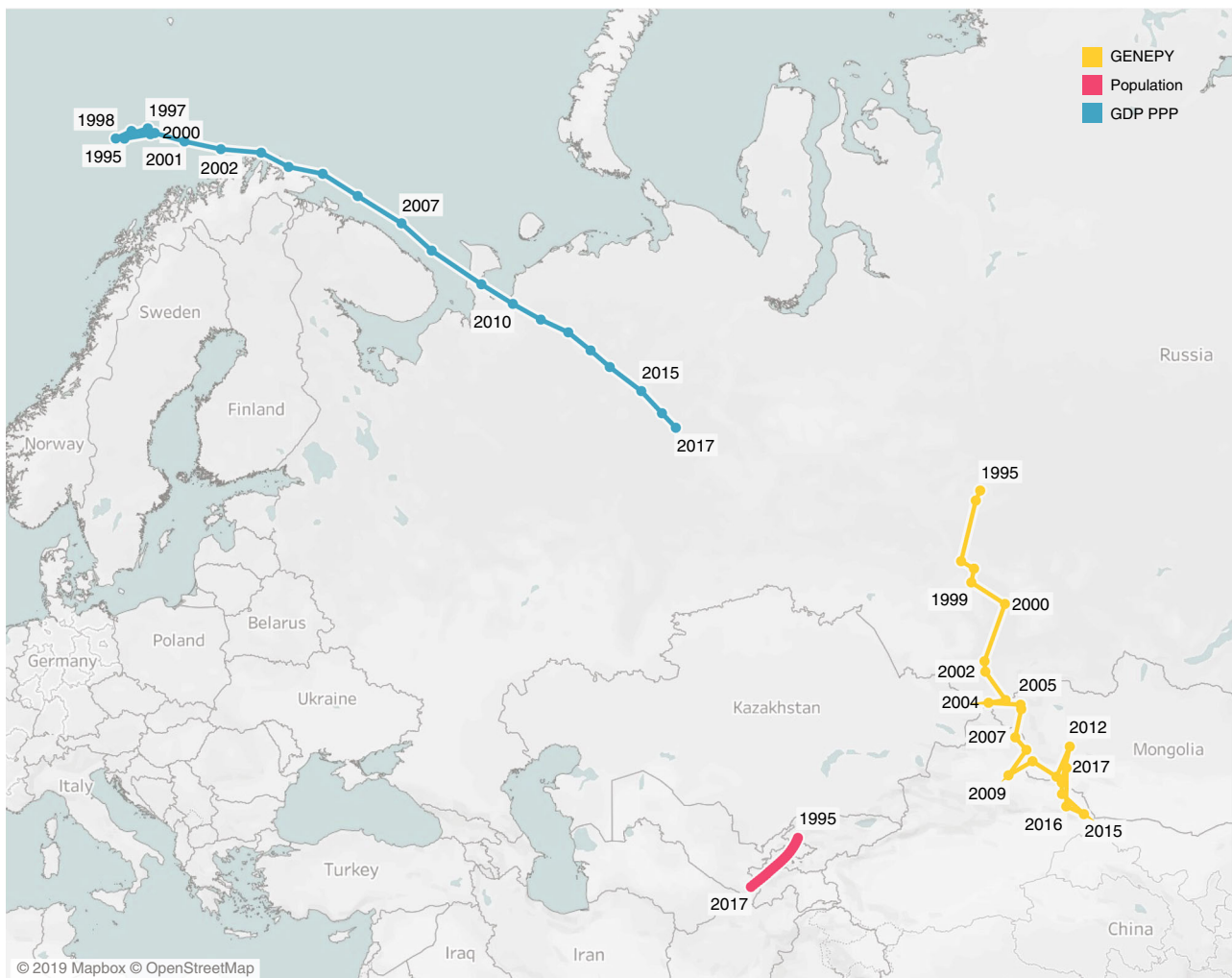
capita" ones<sup>12,15,18,41</sup>), the shifting in the world's centre of GENEPLY has been computed by multiplying each country's GENEPLY index for its population value in time, thus allowing for a fair comparison with the path followed by the GDP (in absolute value) in time. As the figure shows, the trajectories of the GDP and GENEPLY index, differently from the population path, move towards East. The world's centre according to population, although clearly centred in the middle of Asia (as it would have been expected due to the high density of population this area has always recorded<sup>48</sup>), curves toward West as provoked by the increasing population in Africa<sup>48</sup>. The differences in the world's GENEPLY, GDP and population paths confirm that, year by year, the economy is more centred in the East and that increasing population poorly impacts the ability of countries to economically grow. The distance between the current position of the barycenter of GDP and GENEPLY may also imply that Asian countries (China included) still have a strong potential for economic growth, as also stated in Cristelli et al.<sup>23</sup>. Also, the trajectory drawn using GDP differs from the one drawn using the GENEPLY index of countries as weights, because of the ability of the latter to capture both the productive knowledge of countries and the aforementioned dynamics of growth and competition between the actors in the trade.

## Discussion

We have introduced the GENeralised Economic comPlexityY index (GENEPLY), which provides a multidimensional metrics of countries' (and products') complexity. GENEPLY arises from the eigenvectors of a symmetric proximity matrix, describing the similarities in the export baskets of countries. These eigenvectors

combine in a multidimensional fashion, the information obtained from MR and FC metrics, thanks to a mapping (and linearisation for FC) of the original metrics to reduce the problem of finding these metrics to an eigen-centrality problem. GENEPLY ranks countries for their multidimensional complexity, squeezing the eigenvectors through the adoption of a statistical framework on centrality metrics<sup>28</sup>. Moreover, the multidimensionality of our approach can be exploited to trace the economic growth process of countries in time. The richness of the proposed framework demands a deeper focus on some of its aspects.

A key point is that the proximity matrix  $\mathbf{N}$  among countries is symmetric; as a consequence, the left and right eigenvectors coincide and the eigenvector centrality, whereupon our metrics are grounded, is distinctly defined<sup>26,27</sup>. In contrast, by adopting the mathematical approaches of MR or FC, asymmetric matrices are recovered to map countries' Economic Complexity (see Methods section, Eq. (7) for the MR case) – or Fitness (see Methods section, Eq. (11) for the FC case) – onto itself (a mirror argument holds for products). In this case, the eigen-problem can be formulated by considering either right or left eigenvectors, thus posing the question of how the problem should be tackled. This is not just a matter of mathematical formalism: in fact, the eigenvector centrality for directed networks – whose adjacency matrices are asymmetric – typically considers the right and left eigenvectors for determining the out and in centralities of the nodes, respectively, as caused by directionality of the edges<sup>26</sup>. In the same vein, the well-known PageRank<sup>49</sup> centrality algorithm for directed networks considers the left eigenvector to assess only the in-centrality of the nodes. For bipartite networks, the most basic and simple case to rank nodes would be to set  $M_{cp} = W_{cp}$ , thus providing two symmetric proximity matrices  $\mathbf{M} \cdot \mathbf{M}^T$  in Eq.



**Fig. 3 The world's economic and demographic barycentre, 1995–2017.** The trajectories are computed by weighting the countries' geographical centres by their GENEPY index, in yellow, the Gross Domestic Product at Purchasing Power Parity (GDP PPP), in blue, and the population size, in purple. The GDP trajectory is consistent with the one shown by the McKinsey Global Institute<sup>47</sup> taking as reference the path in there shown from 1990–2025. Data for the GDP PPP and the population of countries are provided by the World Bank. The coordinates of countries are provided by the Portland State University and defined according to the georeference system WGS 1984. The figure has been produced with Tableau Public 2019.4.

(3) and  $\mathbf{M}^T \cdot \mathbf{M}$  in Eq. (4)<sup>50</sup>. Contrarily, although set in a bipartite network framework, economic complexity methods as MR and FC generate artificial asymmetry by rescaling this symmetric matrices (using the countries' degree or some of its transforms) without taking care of preserving the feature of symmetry; thus leaving almost arbitrary choice to the solution of the eigen-problem. The symmetry of the transition matrices, also in terms of the adherence to the original symmetric structure of the problem, represents an added value of our framework. Moreover, the bilateral information of the proximity matrix can be used to understand the structure of the export baskets of countries and how these are related through shared common capabilities (Supplementary Fig. 6).

We have also shown how GENEPY can be used to track the economic growth of countries during the years as driven by their economic complexity. Even though economic complexity metrics have already been used to draw these paths<sup>42,51–53</sup>, our innovative multidimensional approach allows one to draw these trajectories without the need of embedding the exogenous information on the GDP per capita that most applications require. As such, the chance of maintaining the simplicity of a data-driven approach endows the GENEPY framework with the main founding reason

for which economic complexity theory was born, i.e., to provide the ground for a more quantitative, data-driven approach to the assessment of the potential economic growth of countries as factored by the productive knowledge<sup>54</sup>.

A further advantage of the GENEPY index is given by its robustness. In fact, when conceiving the bipartite network of countries and products, the commonly used binarisation procedure of the RCA matrix (see Methods section, Eq. (6)) is adopted, aiming at capturing the network topology. However, a different (but possibly relevant) matrix is the one obtained by directly working on the RCA matrix, without reverting the weighted network into a directed one. We show that, also if this path is followed, the GENEPY results remain coherent with respect to changes in the incidence matrix of the network (Supplementary Fig. 10). This does not hold when the MR and FC approaches are used.

The fact of having found very similar results between the linear and the non-linear versions of the FC algorithm (on average, 99.5% Pearson's correlation, Supplementary Fig. 2) cannot be systematically generalised to other cases: in fact, some bipartite systems may require a genuine non-linear approach to let their nested nature emerge (see, e.g., the results pertaining to the



pollinators-plants bipartite network in Supplementary Fig. 11, discussed in Supplementary Note 5). However, the good results obtained in this case suggest that there are also systems where non-linearity plays a minor role. We speculate that this might be related to the differences in the decision-making processes ruling these systems. On the one hand, e.g., nested ecological networks self-organise following ecological rules of non-linear population dynamics<sup>55</sup>. These systems are thus driven by more rigid decision-making processes. On the other hand, the plastic human decision-making process – which is of course at the base of the trade network self-organisation – may give rise to less nested network structures: for a given productive knowledge, trade may follow a simpler sum rule, i.e., “the more, the better”, as trade enhances growth<sup>56</sup>; thus clarifying the reason why the diversity of a country is used as a first proxy of the productive knowledge itself.

Moreover, in the FC algorithm the Quality of a product is mainly determined by the least fit country exporting it, a crucial property accomplished by the non-linearity of the FC approach. In our linear framework, this property is maintained through the term  $k'_p = \sum_c M_{cp}/k_c$ , occurring in  $W_{cp} = M_{cp}/k_c k'_p$ . This term in fact represents the degree of a product corrected by how easily it is found within the network. Its inverse  $1/k'_p$  is an anti-centrality score for the product, determining how limited is its presence within the producers’ baskets and thus suggesting the need for higher productive knowledge in its production process. Notice that, by substituting the incidence matrix  $\mathbf{M}$  with the traded monetary values, the term  $k'_p$  also recurs in the so-called EXPY rationale by Hausmann et al.<sup>2</sup>. Based on a decision-making model of firms’ investment choices, Hausmann et al.<sup>2</sup> defined an index of economic growth potential of countries, assessed through the required productive level of the exported products, i.e., EXPY. As we show, (see Methods section, Eq. (17)), the equations to compute  $X_c$  in the GENEPEY framework are similar to those defining the EXPY scores of countries<sup>2</sup>. Clearly, EXPY has been defined from a different deductive rationale, which considers the trade as described by the weighted incidence matrix of the monetary fluxes (thus providing different input information) and embeds exogenous information such as the GDP per capita. Notwithstanding these differences, the formal similarity of GENEPEY with EXPY is striking. This similarity is a result of the application of our framework, and not an “a priori” construction: in a sense, the economic concepts are self-emerging, with some significant variations with respect to the original EC framework we here reconcile<sup>12,15</sup>. In our view, this similarity represents a possible micro-economically sounded bases for the economic complexity theory, towards which we address future work.

**Methods**

**Data.** Import–export data during the year 1995–2017 are extracted from the BACI-CEPII dataset<sup>36</sup>, which classifies goods according to the Harmonised System Codes 1992 (HS-1992) at the 6-digits level. To allow comparability with previously published results, we downscale the classification of traded goods to the 4-digits level. Our data include all the countries whose export share is worth at least  $10^{-5}$  of the total flux traded during the year (i.e., the total amount of dollars exported worldwide). This filters the noise arising by small export baskets. The Relative Comparative Advantage procedure is used to construct the incidence binary matrix  $\mathbf{M}$ , setting the threshold of RCA to 1 in line with the economic complexity framework<sup>12</sup>. RCA weights how much a product  $p$  counts within the export basket of the country  $c$ . This fraction is weighted by the ratio of the total monetary flux globally generated by the same product  $p$ , and the total monetary flux of all products traded worldwide during the reference year. In formulas,

$$RCA_{cp} = \frac{D_{cp}}{\sum_p D_{cp}} \bigg/ \frac{D_{cp}}{\sum_c D_{cp}} \tag{6}$$

where  $D_{cp}$  is the return in dollars of a country  $c$  exports through the product  $p$ . The input matrix  $\mathbf{M}$  is given by  $M_{cp} = 1$  if  $RCA_{cp} \geq 1$ , and 0 otherwise<sup>25</sup>. In this work we

also consider the direct use of  $RCA_{cp}$  as the input matrix for the computation of the metrics (this implies setting  $RCA_{cp} = M_{cp}$ ), whose results are shown in Supplementary Fig. 10.

**MR metrics.** The equations for the computation of the EC metrics according to the MR approach<sup>12,57</sup> are

$$\begin{cases} ECI_c = \frac{1}{k_c} \sum_p M_{cp} PCI_p, \\ PCI_p = \frac{1}{k_p} \sum_c M_{cp} ECI_c. \end{cases} \tag{7}$$

They can be mapped in our general framework by using

$$\begin{cases} X_c^A = ECI_c \sqrt{k_c}, \\ Y_p^A = PCI_p \sqrt{k_p}, \\ W_{cp}^A = M_{cp} / \sqrt{k_c k_p}. \end{cases} \tag{8}$$

The resulting matrix  $W_{cp}^A$  provides with the following symmetric proximity matrices

$$N_{cc}^A = \sum_p \frac{M_{cp} M_{cp}}{\sqrt{k_c} \sqrt{k_c k_p}}, \tag{9}$$

for countries, and

$$G_{pp}^A = \sum_c \frac{M_{cp} M_{cp}}{k_c \sqrt{k_p} \sqrt{k_p}}, \tag{10}$$

for products. We stress that the matrices  $N^A$  and  $G^A$  are symmetric ones thanks to the presence of the square roots of the degrees  $k_c$  and  $k_p$ , respectively, for which they differ from the corresponding asymmetric matrices that one would obtain directly from the original MR formulation.

Within our framework, the eigenvectors of the two matrices  $N^A$  and  $G^A$  associated to the largest eigenvalue  $\lambda_1 = 1$  are  $X_{c,1} = \sqrt{k_c}$  and  $Y_{p,1} = \sqrt{k_p}$ , from which the unitary eigenvectors of the MR framework are recovered through Eq. (8). The second eigenvectors  $X_{c,2}$  and  $Y_{p,2}$  of the matrices provide the  $ECI_c$  and  $PCI_p$  solutions using Eq. (8) instead.

**FC metrics.** The non-linear FC algorithm defines the values of complexity, Fitness,  $F_c$ , for countries and Quality,  $Q_p$ , for products as<sup>15</sup>

$$\begin{cases} \tilde{F}_c^{(n+1)} = \sum_p M_{cp} Q_p^{(n)}, & F_c^{(n+1)} = \frac{\tilde{F}_c^{(n+1)}}{\left(\sum_c \tilde{F}_c^{(n+1)}\right)/C}; \\ \tilde{Q}_p^{(n+1)} = \frac{1}{\sum_c M_{cp} \tilde{F}_c^{(n)}}, & Q_p^{(n+1)} = \frac{\tilde{Q}_p^{(n+1)}}{\left(\sum_p \tilde{Q}_p^{(n+1)}\right)/P}; \end{cases} \tag{11}$$

where  $C$  and  $P$  are the number of exporting countries and exported products, respectively. In Eq. (11),  $\tilde{F}_c^{(n+1)}$  and  $\tilde{Q}_p^{(n+1)}$  are the intermediate values of  $F_c^{(n+1)}$  and  $Q_p^{(n+1)}$  obtained at each iteration  $(n + 1)$ <sup>15</sup>. At each step, the intermediate values are normalised by their algebraic means, in this way providing the final values  $F_c^{(n+1)}$  and  $Q_p^{(n+1)}$ . The normalisation is required for the stabilisation of the non-linear map in Eq. (11)<sup>35</sup>.

The system in Eq. (11) can be written in closed and non-iterative form as

$$\begin{cases} F_c = c_F \sum_p M_{cp} Q_p, \\ Q_p = c_Q \frac{1}{\sum_c M_{cp} F_c}, \end{cases} \tag{12}$$

in which we have embedded the normalisation procedure by introducing the parameters  $c_F$  and  $c_Q$ , namely  $c_F = \frac{C}{\sum_p Q_p k_p}$  and  $c_Q = \frac{\sum_p Q_p k_p}{P}$ . Equation (11) can be seen as the simplest numerical solution of Eq. (12).

Equation (12) represents a functional relationship between the vectors of values  $F_c$  and  $Q_p$  and, in particular, the Quality values can be formally expressed as  $Q_p = h(F_1, F_2, \dots, F_C)$ ,  $c = [1, \dots, C]$ , where  $h(F_1, F_2, \dots, F_C)$  is a non-linear function of the  $C$  – Fitness values. In order to map the FC algorithm onto the linear  $X_c$ – $Y_p$  framework, we linearise the function  $h(F_c)$  using the Taylor’s series and expanding the function around the value  $F_c = k_c$ , which is known to dominate the information contained in  $F_c$ <sup>19</sup>. Moreover,  $k_c$  is the first result of the map at iteration  $n = 1$ . The Taylor’s expansion provides a linear expression to evaluate the Quality of the products, namely

$$\begin{cases} F_c \simeq c_F \sum_p M_{cp} Q_p, \\ Q_p \simeq \frac{c_Q}{(k'_p)^2} \sum_c \frac{M_{cp} F_c}{k_c^2}, \end{cases} \tag{13}$$

where  $k'_p = \sum_c M_{cp}/k_c$ . Notice that the system in Eq. (13) is an eigen-problem, thus it can be solved without the use of iterative algorithms. This avoids the convergence problem which is known to affect the system in Eq. (11)<sup>29</sup>, due to the hyperbolic



nature of the second equation<sup>35</sup>. As stated in the main text, the linearisation of the original definition of  $F_c$  and  $Q_p$  only biases the results for <0.5% (on average in time according to Pearson’s correlation coefficient, Supplementary Figs. 2 and 7).

Taking the linearised equations in Eq. (13) as the starting point, the mapping of FC metrics within our framework is given as

$$\begin{cases} X_c^B = F_c/k_c, \\ Y_p^B = Q_p \cdot k'_p, \\ W_{cp}^B = M_{cp}/(k_c k'_p), \end{cases} \quad (14)$$

where we neglect the rescaling factors  $c_F$  and  $c_Q$ , since their roles of stabilising the numerical values is not anymore required due to linearity, thus reducing the number of unknowns in the system.

The resulting matrix  $W_{cp}^B = M_{cp}/k_c k'_p$  provides with the following symmetric proximity matrices

$$N_{cc'}^B = \sum_p \frac{M_{cp} M_{c'p}}{k_c k_{c'} (k'_p)^2}, \quad (15)$$

for countries, and

$$G_{pp'}^B = \sum_c \frac{M_{cp} M_{c'p'}}{k_c^2 k'_p k'_{p'}}, \quad (16)$$

for products. The linearised values for Fitness and Quality are recovered from the eigenvectors of the proximity matrices associated to the largest eigenvalue  $\lambda_1$ , from which holds  $X_{c,1} = F_c/k_c$  and  $Y_{p,1} = Q_p k'_p$ .

Notice that the computation of the GENEPEY index entails interpreting the matrices  $\mathbf{N}$  and  $\mathbf{G}$  as proximity matrices, thus setting their diagonal elements to same constant values: we here set  $N_{cc}^B = G_{pp}^B = 0$ . Even when the matrices  $\mathbf{N}$  and  $\mathbf{G}$  are interpreted as proximity matrices (i.e., their diagonal is set to zero), very good correlations are obtained between linearly and non-linearly computed values (Supplementary Figs. 2 and 7).

**Relation to EXPY metric.** It is easy to verify the similarity of the relation in Eq. (3) (with the elements  $N_{cc}$  as given in Eq. (15)) to compute the  $X_c$  values with the expression to compute the productivity of a country according to the EXPY<sup>2</sup>. In fact, by recalling the weighted incidence matrix of the export volumes in dollars,  $D_{cp}$ , and the strengths of countries and products such that:

$$k_c = \sum_p D_{cp}, \quad k_p = \sum_c D_{cp}, \quad k'_p = \sum_c \frac{D_{cp}}{k_c};$$

one has that the productivity level of a product, named PRODY, is given as

$$PRODY_p = \sum_c \frac{D_{cp}}{k_c k'_p} R_c,$$

being  $R_c$  the GDP per capita of the country  $c$ . EXPY, as a function of the PRODY, is computed as

$$\begin{aligned} EXPY_c &= \sum_p \frac{D_{cp}}{k_c} PRODY_p = \sum_p \frac{D_{cp}}{k_c} \sum_{c'} \frac{D_{c'p}}{k_{c'} k'_p} R_{c'} \\ &= \sum_{c'} \sum_p \frac{D_{cp} D_{c'p}}{k_c k_{c'} k'_p} R_{c'}. \end{aligned} \quad (17)$$

EXPY mainly differs from the GENEPEY approach, and thus the linearised FC one, – apart from a rescaling factor  $k'_p$  (see Eq. (15)) – for the embedding of the exogenous information on the GDP per capita which replaces the country–country relation in Eq. (3).

**The GENEPEY index.** The description on how the GENEPEY index is derived from the eigenvectors of the proximity matrices is here exemplified for countries, and the same procedure applies for the index as referred to products. In fact, to obtain the GENEPEY index for products, it is sufficient to replace in the following the terms  $X_{c,1}$ ,  $X_{c,2}$  and  $\mathbf{N}$  with  $Y_{p,1}$ ,  $Y_{p,2}$  and  $\mathbf{G}$ , respectively.

The GENEPEY index for countries combines the eigenvectors corresponding to the two largest eigenvalues of the symmetric proximity matrix  $\mathbf{N}$  (see Eq. (15)). The manner how the information obtained from the two eigenvectors is squeezed into the unique measure in Eq. (5) finds its roots in the recast of the network centrality problem into an estimation exercise. The main steps of this procedure follow, and we refer the readers to the original work<sup>28</sup> for a more detailed explanation.

The matrix  $\mathbf{N}$  describes the weighted adjacency matrix of the undirected network whose nodes are the countries and edges the similarities among them. The eigenvectors of this matrix represent centrality measures of the nodes. Our aim is to use the eigenvectors to least-square estimate the matrix  $\mathbf{N}$ . Firstly, we introduce a centrality-dependent estimator function  $\zeta$ . In the case of the eigenvector centrality, such function linearly depends on the eigenvectors  $X_{c,1}$  and  $X_{c,2}$ , corresponding to the two largest eigenvalues  $\lambda_1$  and  $\lambda_2$  of the matrix  $\mathbf{N}$ <sup>27,28,58</sup>. In formulas

$$\zeta(\lambda_i, \mathbf{X}_{c,i}, \mathbf{X}_{s,i}) = \sum_{i=1}^2 \lambda_i X_{c,i} X_{s,i}, \quad (18)$$

where  $i = [1, 2]$  and  $c$  and  $s$  run in the range  $[1, C]$ , being  $C$  the number of countries in the matrix. The function  $\zeta$  minimises the squared errors between the matrix elements and the corresponding estimates; namely

$$SE = \sum_c \sum_s \left( N_{cs} - \zeta(\lambda_i, X_{c,i}, X_{s,i}) \right)^2. \quad (19)$$

Secondly, at a fixed  $i^*$ , each eigenvector  $\mathbf{X}_{i^*}$  solves the minimisation problem<sup>28</sup>

$$\frac{\partial SE}{\partial \mathbf{X}_{i^*}} = 0.$$

In this multidimensional setting on eigenvector centrality, the ranking of the network’s nodes (i.e., the countries) is given by the adoption, from the commonality analysis, of the concept of the unique contribution of the  $X_{c,i}$  variables. The unique contribution is defined as the drop in the coefficient of determination  $R^2$  induced by excluding the variables  $X_{c,i}$  ( $i = [1, 2]$ ) considered in the estimator function  $\zeta$ , in Eq. (18), from the estimation procedure<sup>28</sup>. The core concept of the unique contribution is that, the larger the drop, the larger is the contribution of the  $c$ -th values in the reconstruction of the matrix  $\mathbf{N}$  and, in this application, the more central the  $c$ -th node is in the network topology under analysis. Hence, according to this approach, we define the GENERALISED Economic complexity index (GENEPEY) for the country  $c$  as the unique contribution of its complexity values  $X_{c,i}$  as computed by the formula given in Eq. (5).

**Reporting summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

**Data availability**

The trade data supporting the findings of this study are available upon request from the BACI-CEPII database (Gaulier and Zignago<sup>36</sup>). Downloads may require paid subscription.

The GDP PPP and population data used in this work are provided by the World Bank and publicly and freely available at [<https://data.worldbank.org/>].

The data on the coordinates of countries are provided by the Portland State University and publicly and freely available at [<https://www.pdx.edu/econ/country-geography-data/>].

The pollinators-plants networks are freely available at [[www.web-of-life.es](http://www.web-of-life.es)].

The results of the GENEPEY index for countries during the period of analysis are publicly and freely available at [<https://zenodo.org/record/3876721>]. Other results are available from the authors upon request.

**Code availability**

The code for the computation of the GENEPEY index is publicly and freely available at [<https://zenodo.org/record/3876721>].

Received: 25 October 2019; Accepted: 8 June 2020;

Published online: 03 July 2020

**References**

1. Brillat-Savarin, J. A. *Physiologie du goût* (Charpentier, 1841).
2. Hausmann, R., Hwang, J. & Rodrik, D. What you export matters. *J. Economic Growth* **12**, 1–25 (2007).
3. Schumpeter, J. *The theory of economic development* (Transaction Publishers, 1934).
4. Aghion, P., & Howitt, P. A model of growth through creative destruction. Tech. Rep. No. w3223 (National Bureau of Economic Research, 1990).
5. Romer, P. M. Endogenous technological change. *J. Political Econ.* **98**, S71–S102 (1990).
6. Rodrik, D. Goodbye Washington consensus, hello Washington confusion? A review of the World Bank’s economic growth in the 1990s: learning from a decade of reform. *J. Econ. Lit.* **44**, 973–987 (2006).
7. Rodrik, D. *One economics, many recipes: Globalization, institutions, and economic growth* (Princeton University Press, 2008).
8. Nelson, R. R. *An evolutionary theory of economic change* (Harvard University Press, 2009).
9. Easterly, W. & Levine, R. What have we learned from a decade of empirical research on growth? It’s not factor accumulation: stylized facts and growth models. *World Bank Econ. Rev.* **15**, 177–219 (2001).
10. Hulten, C. R. in *New developments in productivity analysis*. 1–54 (University of Chicago Press, 2001).
11. Hausmann, R. et al. *The Atlas of Economic Complexity: Mapping paths to prosperity* (MIT Press, 2014).
12. Hidalgo, C. A. & Hausmann, R. The building blocks of economic complexity. *Proc. Natl Acad. Sci. USA* **106**, 10570–10575 (2009).
13. Inoua, S. A simple measure of economic complexity. *arXiv*. preprint arXiv:1601.05012 (2016).

14. Mariani, M. S., Vidmer, A., Medo, M. & Zhang, Y.-C. Measuring economic complexity of countries and products: which metric to use? *Eur. Phys. J. B* **88**, 293 (2015).
15. Tacchella, A., Cristelli, M., Caldarelli, G., Gabrielli, A. & Pietronero, L. A new metrics for countries' fitness and products' complexity. *Sci. Rep.* **2**, 723 (2012).
16. Caldarelli, G. et al. A network analysis of countries' export flows: firm grounds for the building blocks of the economy. *PLoS ONE* **7**, 47278 (2012).
17. Cristelli, M., Gabrielli, A., Tacchella, A., Caldarelli, G. & Pietronero, L. Measuring the intangibles: a metrics for the economic complexity of countries and products. *PLoS ONE* **8**, e70726 (2013).
18. Kemp-Benedict, E. An interpretation and critique of the Method of Reflections. MPRA Paper No. 60705 (MPRA, 2014).
19. Morrison, G. et al. On economic complexity and the fitness of nations. *Sci. Rep.* **7**, 15332 (2017).
20. Felipe, J., Kumar, U., Abdon, A. & Bacate, M. Product complexity and economic development. *Struct. Change Econ. Dyn.* **23**, 36–68 (2012).
21. Poncet, S. & de Waldemar, F. S. Export upgrading and growth: the prerequisite of domestic embeddedness. *World Dev.* **51**, 104–118 (2013).
22. Alsén, A. et al. National strategy for Sweden: From wealth to well-being. Tech. Rep. (The Boston Consulting Group, 2013).
23. Cristelli, M., Tacchella, A., Cader, M., Roster, K. & Pietronero, L. On the predictability of growth. Policy Research Working Paper No. WPS8117 (The World Bank, 2017).
24. Brito, S., Magud, N. E. & Sosa, S. Real exchange rates, economic complexity, and investment. Working Paper No. 18/107 (International Monetary Fund, 2018).
25. Balassa, B. Trade liberalisation and “Revealed” Comparative Advantage. *Manch. Sch.* **33**, 99–123 (1965).
26. Newman, M. E. *Network - An Introduction* (Oxford University Press, 2010).
27. Golub, G. H. & Van Loan, C. F. *Matrix Computations*. Vol. 3 (JHU Press, 2012).
28. Sciarra, C., Chiarotti, G., Laio, F. & Ridolfi, L. A change of perspective in network centrality. *Sci. Rep.* **8**, 15269 (2018).
29. Pugliese, E., Zaccaria, A. & Pietronero, L. On the convergence of the Fitness-Complexity Algorithm. *Eur. Phys. J. Spec. Top.* **225**, 1893–1911 (2016).
30. Baudena, M. et al. Revealing patterns of local species richness along environmental gradients with a novel network tool. *Sci. Rep.* **5**, 11561 (2015).
31. Domínguez-García, V. & Muñoz, M. A. Ranking species in mutualistic networks. *Sci. Rep.* **5**, 8182 (2015).
32. Tu, C., Carr, J. & Suweis, S. A data driven network approach to rank countries production diversity and food specialization. *PLoS ONE* **11**, e0165941 (2016).
33. Lin, J.-H., Tessone, C. & Mariani, M. Nestedness maximization in complex networks through the fitness-complexity algorithm. *Entropy* **20**, 768 (2018).
34. Albeaik, S., Kaltenberg, M., Mansour, A. & Hidalgo, C. Improving the economic complexity index. *arXiv preprint arXiv:1707.05826* (2017).
35. Gabrielli, A. et al. Why we like the ECI+ algorithm. *arXiv preprint arXiv:1708.01161* (2017).
36. Gaulier, G. & Zignago, S. *BACI: International trade database at the product-level*. Working Paper No. 2010 – 23 (CEPII, 2010).
37. Soete, L., Schneegans, S., Eröcal, D., Angathevar, B. & Rasiah, R. *UNESCO Science Report: Towards 2030* (UNESCO, 2015).
38. Randall, J. E. & Ironside, R. G. Communities on the edge: an economic geography of resource-dependent communities in Canada. *Can. Geographer/Le Géographe Canadien* **40**, 17–35 (1996).
39. Battersby, B. et al. International trade performance: The gravity of Australia's remoteness. Treasury Working Paper No 2005 – 03 (The Treasury Australian Government, 2005).
40. Guttmann, S. & Richards, A. Trade openness: an Australian perspective. *Aust. Econ. Pap.* **45**, 188–203 (2006).
41. Mealy, P., Farmer, J. D. & Teytelboym, A. Interpreting economic complexity. *Sci. Adv.* **5**, 1705 (2019).
42. Tacchella, A., Mazzilli, D. & Pietronero, L. A dynamical systems approach to gross domestic product forecasting. *Nat. Phys.* **14**, 861 (2018).
43. Allen, F., Qian, J. & Qian, M. Law, finance, and economic growth in China. *J. Financial Econ.* **77**, 57–116 (2005).
44. Cetorelli, N. & Goldberg, L. S. Global banks and international shock transmission: evidence from the crisis. *IMF Econ. Rev.* **59**, 41–76 (2011).
45. Bordini, R. *Exceptional measures: The Shanghai stock market crash and the future of the Chinese economy*. In-depth Analysis September 2015-PE 549.067 (European Union, 2015).
46. Pritchett, L. & Summers, L. H. *Asiaphoria meets regression to the mean*. Tech. Rep. No. w20573 (National Bureau of Economic Research, 2014).
47. Dobbs, R. et al. *Urban world: Cities and the rise of the consuming class* (McKinsey & Company, 2012).
48. UNDESA. *World Population Prospects: the 2012 revision, volume II, Demographic Profiles* (st/esa/ser.a/345) (United Nations, 2013).
49. Brin, S. & Page, L. The anatomy of a large-scale hypertextual Web search engine. *Comput. Netw.* **30**, 101–117 (1998).
50. Everett, M. & Borgatti, S. The dual-projection approach for two-mode networks. *Soc. Netw.* **35**, 204–210 (2013).
51. Pugliese, E., Chiarotti, G. L., Zaccaria, A. & Pietronero, L. Complex economies have a lateral escape from the poverty trap. *PLoS ONE* **12**, e0168540 (2017).
52. Liao, H. & Vidmer, A. A comparative analysis of the predictive abilities of economic complexity metrics using international trade network. *Complexity* **2018**, 1–12 (2018).
53. Brummitt, C. D., Gomez-Lievano, A., Hausmann, R. & Bonds, M. H. Machine-learned patterns suggest that diversification drives economic development. *arXiv preprint arXiv:1812.03534* (2018).
54. Tacchella, A., Cristelli, M., Caldarelli, G., Gabrielli, A. & Pietronero, L. Economic complexity: conceptual grounding of a new metrics for global competitiveness. *J. Econ. Dyn. Control* **37**, 1683–1691 (2013).
55. Smith, R. L., Smith, T. M., Hickman, G. C. & Hickman, S. M. *Elements of Ecology* (Benjamin Cummings Menlo Parie, CA, 1998).
56. Frankel, J. A. & Romer, D. H. Does trade cause growth? *Am. Econ. Rev.* **89**, 379–399 (1999).
57. Albeaik, S., Kaltenberg, M., Alsaleh, M. & Hidalgo, C. A. 729 new measures of economic complexity (addendum to Improving the Economic Complexity Index). *arXiv preprint arXiv:1708.04107* (2017).
58. Bonacich, P. Factoring and weighting approaches to status scores and clique identification. *J. Math. Sociol.* **2**, 113–120 (1972).

### Acknowledgements

We acknowledge ERC funding from the CWASI project (ERC-2014-CoG, project 647473).

### Author contributions

C.S., G.C., L.R., and F.L. conceived and designed the study. C.S. prepared the data and conducted the experiments. C.S., G.C., L.R., and F.L. analysed the data. CS produced the figures and wrote the manuscript. C.S., G.C., L.R., and F.L. edited the manuscript. All authors reviewed the manuscript.

### Competing interests

The authors declare no competing interests.

### Additional information


**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41467-020-16992-1>.

**Correspondence** and requests for materials should be addressed to C.S.

**Peer review information** *Nature Communications* thanks Matthieu Cristelli and Andrea Tacchella for their contribution to the peer review of this work. Peer reviewer reports are available.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020