ORIGINAL RESEARCH

Mathematical modeling of local balance in signed networks and its applications to global international analysis

Fernando Diaz-Diaz[1] · Paolo Bartesaghi[2] · Ernesto Estrada[1]

Abstract

Alliances and conflicts in social, political and economic relations can be represented by positive and negative edges in signed networks. A cycle is said to be positive if the product of its edge signs is positive, otherwise it is negative. Then, a signed network is balanced if and only if all its cycles are positive. An index characterizing how much a signed network deviates from being balanced is known as a global balance index. Here we give a step forward in the characterization of signed networks by defining a local balance index, which characterizes how much a given vertex of a signed network contributes to its global balance. We analyze the mathematical foundations and unique structural properties of this index. Then, we apply this index to the study of the evolution of international relations in the globe for the period 1816–2014. In this way we detect and categorize major historic events based on balance fluctuations, helping our understanding towards new mixed approaches to history based on network theory.

Keywords Signed networks · Structural balance · International relations · Geopolitical networks · Quantitative history

Mathematics Subject Classification 05C22 · 05C50 · 05C90 · 91D35 · 91F10

✉ Ernesto Estrada
  estrada@ifisc.uib-csic.es

  Fernando Diaz-Diaz
  fernandodiaz@ifisc.uib-csic.es

  Paolo Bartesaghi
  paolo.bartesaghi@unimib.it

1 Institute of Cross-Disciplinary Physics and Complex Systems, IFISC (UIB-CSIC), 07122 Palma de Mallorca, Spain

2 Department of Statistics and Quantitative Methods, University of Milano-Bicocca, Via Bicocca degli Arcimboldi 8, 20126 Milano, Italy

🙋 Springer

1 Introduction

Complex systems are, by definition, networked [1]. Therefore, their complexities demands the use of a combination of global and local graph-theoretic invariants to understand their structures and functions [2, 3]. A particular case of complex system is the representation of conflicting interactions by means of signed networks [4]. In this case, for instance, positive relations may refer to friendship, collaboration or alliances, and negative ties represent enmities, opposite view about a topic, or conflicts. When the set of vertices of a signed network can be split into two subsets such that every edge between nodes within each subset is positive, and negative edges only interconnecting vertices of the different sets, the system is called balanced [5, 6]. A balanced network does not contain negative cycles—that is, cycles in which the product of the edge signs is negative. When a signed network is not balanced, it is important to know how unbalanced it is, which will give rise to an index characterizing the degree of balance of the network as a whole [7–12]. Moreover, in an unbalanced network not every vertex contributes the same to the global level of balance of the graph. Thus, it is important to know the local degree of balance of the vertices. This concept was proposed by Harary [13] and further by Cartwright and Harary [6]. However, these authors did not systematically studied any particular index of the degree of local balance in signed graphs.

The theory of network balance emerged from the early days of social network analysis [14, 15], and nowadays has become an important paradigm for the analysis of sociometric relations ranging from individuals, to institutions and countries [7, 8, 10–12, 16]. In particular, balance theory plays an important role in the analysis of international relations (IR), where signed networks are used to represent countries and their alliances/conflicts [17–24]. This represents an alternative approach to the more classical, statistically-oriented, frameworks of "quantitative history" [25–27], and "cliodynamics" [28]. Coalition models for IR have also been proposed to account for local versus global alignment among countries, showing that, for example, the existence of two competing world coalitions leads to a more stable distribution of actors, while a single world leadership enables the emergence of unstable relationships [29]. These previous studies on networks in IR (NIR) have focused on detecting global patterns on the temporal evolution of the world as a whole. However, to connect NIR with historical narrative it is necessary to zoom in the individual contribution of every country to the balance/unbalance of the world in a given historic period. Consequently, here we define a local balance index based on algebraic graph theory and study several of its mathematical properties. In addition, we design an approach based on this index to identify the major events in the world history for the period between 1816 and 2014. In this way we connect a graph-based quantitative analysis of NIR with the qualitative, narrative account of history [27, 30]. The current work can be contextualized in a general framework of quantitative history based on the so-called "mixed method", which "combines quantitative and qualitative research techniques, methods, approaches, concepts or language into a single study" [31, 32].

The current paper is subdivided into two interrelated parts. In Part I we motivate, develop, analyze, interpret and compare a new local balance index. The goals of this part of the paper are: (1) to provide a motivation about the necessity for the

development of a local balance index; (2) to define mathematically a local balance index that accounts for the role of an individual vertex to the global balance in a signed network; (3) to analyze several of the mathematical characteristics of this local balance index remarking its structural interpretability; (4) to provide an interpretation of the local balance index in terms of a dynamical process taking place in signed (social) networks. In Part II we develop a general methodology for connecting significant changes in the local balance index of individual countries with historic events involving those countries at a given time. The goals of this second part of the paper are: (1) to apply the local balance index to the study of IR among countries in the world; (2) to design and test a general methodology allowing to match the changes (peaks and valleys) in the time series of countries' local balance with major historic events involving those countries at this time; (3) to perform a classification of major historic events on the basis of the nature of the change (peaks and valleys) occurring in the local balance; (4) to link the identification of peaks and valleys with a narrative of historic events; (5) to analyze the temporal series of local balance of individual countries in their respective networks of IR in the period 1816–2014.
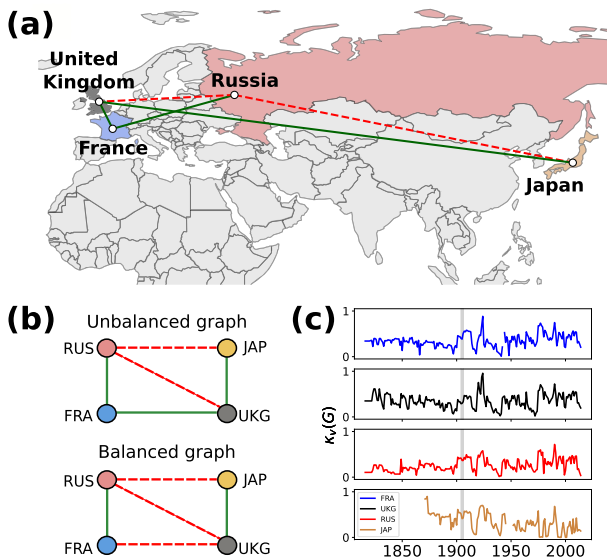
2 Preliminaries

A binary undirected signed network is represented by a triple $G = (V, E, \sigma)$, where V is the set of the n nodes, E is the set of the undirected edges $e = (v, w)$ with $v, w \in V$, and σ is a mapping $\sigma : E \rightarrow \{\pm 1\}$ which assigns to each edge a sign. The signed network can be completely characterized by the $n \times n$ signed adjacency matrix $A = [A_{vw}]$. Specifically, if there is no edge (v, w), $A_{vw} = 0$; otherwise, $A_{vw} = +1$ represents a positive edge, and $A_{vw} = -1$ represents a negative edge. We assume graphs without self-loops so that the matrix A has null diagonal entries. The underlying unsigned network, i.e. the graph obtained from G by ignoring the sign of the edges, is associated with the matrix $|A|$ of the absolute value entries of A. The absolute degree of a node v, denoted by k_v, is defined to be the number of edges incident to that node, irrespective of the edge signs, that is $k_v = \sum_w |A_{vw}|$. A non-empty path in which only the first and the last node coincide is called cycle. The sign of a cycle is the product of the signs of its edges. A signed graph G is *structurally balanced* if there are no negative cycles, or, equivalently, there are no cycles with an odd number of negative edges (see [6]). In a balanced network, the set V can be partitioned into two subsets such that every edge between them is negative, while every edge within them is positive (see [5]).

3 Part I. Local balance index

3.1 Motivation of the local balance index

Let us motivate the necessity of defining a local balance index with one example from the real-world. In the period 1904–1906 France and Russia were allies (Franco-Russian treaty), as were Japan and Britain (Anglo-Japanese alliance) and France and

Fig. 1 **a** Illustration of the relations between Great Britain, France, Russia and Japan during the Russo-Japanese war (1904–05). **b** The mathematical framework for the theory is the representation of IR as signed networks, in which positive edges (green solid lines) represent interstates alliances and negative ones (red dashed lines) are reserved for any kind of conflict/enmities between the parts. **c** Temporal evolution of the "local balance" of the countries in the examined time period. The gray area highlights the years in which the Russo-Japanese war took place. (Color figure online)

Britain (Entente Cordiale). However, Russia and Britain were enemies at the time when Russia went to war with Japan. Therefore, France, Great Britain and Russia formed a signed triangle [4] with two positive and one negative edge (see Fig. 1a). Then, if Britain made effective its support to Japan, Russia may declare war against its enemy Great Britain. This would oblige France to make effective its support to Russia as they were allies, but to enter into a contradiction with its new ally, the United Kingdom. This conflictive situation is a consequence of the fact that the triangle UK-Russia-France is unbalanced (Fig. 1b, top panel). The real consequence was that neither France supported Russia during the war, nor Great Britain made its alliance with Japan effective. These contradictions would not exist if the triangle between these countries were balanced (Fig. 1b, bottom panel), where two factions (Russia-France vs. Japan-UK) would be formed.

As can be seen in Fig. 1b, top panel, UK–Japan–Russia formed a balanced triangle. Consequently, in total, Japan participates in one balanced triangle and in zero unbalanced ones, Russia and Great Britain are part of one balanced and one unbalanced triangle, and France participates in an unbalanced one. Thus, not all countries have the same balancing position in this historic event. To capture these differences in the contribution of individual nodes to the global balance of the network we will define a local balance index.

3.2 Definition and mathematical properties of the local balance index

A closer look at Fig. 1b, top panel, reveals that France is not only related to the rest of conflicting nations via a balanced triangle, but also by means of the unbalanced square RUS-JAP-UKG-FRA. The importance of considering cycles beyond triangles for characterizing the degree of balance of a network is well known since the pioneering works of Cartwright and Harary [6] (see also [11]). Instead of counting signed cycles, which is a computationally costly procedure, we will quantify the local balance by counting local closed walks (LCW). A LCW of length k starting at node v in a signed network G is given by $\left(A^k\right)_{vv}$, where A is the adjacency matrix of the network. The diagonal elements of the $k-$power of this matrix can also be written as: $(A^k)_{vv} =$ # of positive LCW of length $k - |$# of negative LCW of length $k|$.

Let us consider a LCW of length k with one negative edge $e = \{p, q\}$. Every node $v \neq p, q$ is at distance i from p and at distance $n - i - 1$ $(i = 1, \ldots, n - 2)$ from q. This means that the longer the LCW, the smaller the average influence of the negative edge on the rest of the nodes of the cycle. Among the several choices giving rise to specific matrix functions [33], we choose here to penalize every LCW of length k by $1/k!$, such that we have the following definition.

Definition 1 Let G be any signed graph with adjacency matrix A. Let $|A|$ be the entrywise absolute of the adjacency matrix A of the network. The local balance index of $v \in V$ is defined as:

$$\kappa_v(G) := \frac{\sum_{k=0}^{\infty}(A^k)_{vv}/k!}{\sum_{k=0}^{\infty}(|A|^k)_{vv}/k!} = \frac{(\exp A)_{vv}}{(\exp|A|)_{vv}}, \tag{1}$$

We now prove some general properties of the local balance index.

Proposition 1 *Let G be any signed connected graph and $v \in V$. Then, $\kappa_v(G)$ is bounded as follows:*

$$0 < \kappa_v(G) \leq 1. \tag{2}$$

Proof Let $\alpha_1 > \alpha_2 > \ldots > \alpha_n$ and $a_1, a_2, \ldots, a_n$ be the eigenvalues and associated eigenvectors of A, respectively. Similarly, let $\beta_1 > \beta_2 > \ldots > \beta_n$ and $b_1, b_2, \ldots, b_n$ be the eigenvalues and eigenvectors of $|A|$, respectively. Then, we can write Eq. (1) as:

$$\kappa_v(G) = \frac{\sum_{j=1}^{n}(a_j)_v^2 e^{\alpha_j}}{\sum_{l=1}^{n}(b_l)_v^2 e^{\beta_l}}, \tag{3}$$

where $(a_j)_v^2$ denotes the square of the vth entry of the vector a_j. From Eq. (3), it is clear that $\kappa_v(G) > 0$ for any v, since all terms within each sum are nonnegative and at least one of the $(a_j)_v$ is different from zero.

To prove that $\kappa_v(G) \leq 1$, we use the power-series expansion of the exponential function: $(e^A)_{vv} = \sum_k \frac{(A^k)_{vv}}{k!}$. We recall that $(A^k)_{vv}$ counts the difference between the number of positive and negative closed walks starting and ending at node v. Let us denote by $(W^+)^k_{vv}$ and $(W^-)^k_{vv}$ the number of positive and negative walks of length k starting and ending at the node v, respectively. Then, the local balance index can be rewritten as:

$$\kappa_v(G) = \frac{\sum_{k=0}^{\infty} \frac{1}{k!}[(W^+)^k_{vv} - (W^-)^k_{vv}]}{\sum_{k=0}^{\infty} \frac{1}{k!}[(W^+)^k_{vv} + (W^-)^k_{vv}]}. \tag{4}$$

Since $(W^+)^k_{vv} - (W^-)^k_{vv} \leq (W^+)^k_{vv} + (W^-)^k_{vv}$, then $\kappa_v(G) \leq 1$. $\qquad\square$

Let us now recall two basic concepts of balance theory on signed graphs.

Definition 2 Let G be any signed, connected graph and let C be any cycle in G. The sign of the cycle C is the product of the signs of all edges in the cycle. C is called positive if its sign is positive, otherwise it is negative.

Definition 3 Let G be any signed, connected graph. The signed graph G is balanced if and only if, all of its cycles are positive.

We then have the following property of the local balance index.

Proposition 2 *Let G be any signed connected graph. G is balanced if and only if $\kappa_v(G) = 1$ for any node v.*

Proof Recall that the global balance index of G is defined as [10]

$$\kappa(G) := \frac{tr\left(e^A\right)}{tr\left(e^{|A|}\right)},$$

where tr is the trace of the corresponding matrix. Then, we can write:

$$\kappa(G) = \frac{\sum_{v=1}^{n}\left(\kappa_v(G)\, SC_v\right)}{\sum_{v=1}^{n} SC_v}, \tag{5}$$

where $SC_v = (e^{|A|})_{vv}$ is the subgraph centrality of a node v [34]. It has been proved in previous works that $\kappa(G) = 1$ if and only if G is balanced (see [10]). By Eq. (5), this is true if and only if and only if $\kappa_v(G) = 1$ for all $v \in V$. $\qquad\square$

The lower bound of $\kappa_v(G)$ can be improved by virtue of the following proposition.

Proposition 3 *Let G be any signed connected graph and let $k_{max} = \max_{v \in V} k_v$. Then,*

$$\kappa_v(G) \geq e^{-2k_{max}}. \tag{6}$$

Proof Let us denote the smallest eigenvalue of A by α_{min}. Then, using the spectral decomposition of A, we obtain the following inequality:

$$(e^A)_{vv} = \sum_{j=1}^{n}(a_j)_v^2 e^{\alpha_j} \geq e^{\alpha_{min}} \sum_{j=1}^{n}(a_j)_v^2 = e^{\alpha_{min}}$$

In the last step, we have used the fact that $\sum_j (a_j)_v^2 = 1$, which follows from the orthonormality of the eigenbasis $\{a_j\}$. To find a lower bound for α_{min}, we use Gershgorin's theorem. According to it, all eigenvalues of A lie on the union of the disks D_v centered at A_{vv} and with radius $R_v = \sum_{w \neq v}|A_{vw}|$. Because A is the adjacency matrix of a graph without self-loops, $A_{vv} = 0$ for all v and $R_v = \sum_{w \neq v}|A_{vw}| = k_v$, i.e., the radius of each Gershgorin disk is the absolute degree of the corresponding node. Consequently, all eigenvalues must lie within a disk centred at the origin and with radius k_{max}. As a consequence, $\alpha_{min} \geq -k_{max}$ and $(e^A)_{vv} \geq e^{-k_{max}}$. Repeating the same calculation for $e^{|A|}$, we find that $(e^{|A|})_{vv} \leq e^{\beta_{max}} \leq e^{k_{max}}$. Finally, combining these two inequalities results in $\kappa_v(G) = \frac{(e^A)_{vv}}{(e^{|A|})_{vv}} \geq e^{-2k_{max}}$. □

Corollary 1 *Let G be any signed connected graph. Then,* $\kappa_v(G) \geq e^{2-2n}$.

We now find which graph has the lowest possible $\kappa_v(G)$.

Proposition 4 *Let K_n^- be the complete graph with all-negative edges. Then,*

$$\lim_{n \to \infty} \kappa_v(K_n^-) = 0, \ \forall v \in V. \tag{7}$$

Proof The spectrum of the complete graph is $Sp(K_n) = \{(-1)^{[n-1]}, (n-1)^{[1]}\}$, while the spectrum of the complete all-negative graph is $Sp(K_n^-) = \{(1-n)^{[1]}, 1^{[n-1]}\}$ (the superindices indicate the multiplicity of each eigenvalue). Since K_n^- is a complete all-negative graph, all nodes are equivalent and therefore $(e^A)_{vv}$ is independent of v. Consequently, $(e^A)_{vv} = \frac{tr(e^A)}{n} = \frac{e^{1-n}+(n-1)e}{n}$. Similarly, $(e^{|A|})_{vv} = \frac{e^{n-1}+(n-1)e^{-1}}{n}$. Substituting in the expression of the local balance index defined in Eq. (1), we finally get:

$$\kappa_v(K_n^-) = \frac{e^{1-n} + (n-1)e}{e^{n-1} + (n-1)e^{-1}},$$

which becomes zero in the limit $n \to \infty$. □

The following result regarding the local balanced of cycle graphs [35] is proved in the Sect. 1 of the Supplementary Material (SM).

Proposition 5 *Let C_n^{k-} be a signed cycle of length n and k negative edges. Then, the local balance index of any node in C_n^{k-} is*

$$
\kappa_v\left(C_{2l}^{k-}\right) =
\begin{cases}
\dfrac{\sum_{j=0}^{2l-1} \exp\left(2\cos\left(\dfrac{(2j+1)\pi}{2l}\right)\right)}{\sum_{j=0}^{2l-1} \exp\left(2\cos\left(\dfrac{j\pi}{l}\right)\right)} & k\ \text{odd}, \\[20pt]
1 & k\ \text{even},
\end{cases}
\tag{8}
$$

for even cycles and

$$
\kappa_v\left(C_{2l+1}^{k-}\right) =
\begin{cases}
\dfrac{\sum_{j=0}^{2l} \exp\left(2\cos\left(\dfrac{2j\pi}{2l+1}\right)\right)}{\sum_{j=0}^{2l} \exp\left(2\cos\left(\dfrac{(2j+1)\pi}{2l+1}\right)\right)} & k\ \text{odd}, \\[20pt]
1 & k\ \text{even},
\end{cases}
\tag{9}
$$

for odd ones.

Remark 1 The previous result illuminates several important structural aspects of the local balance in cycles. First, that every node in a signed cycle will have the same local balance independently of how close it is from the negative edges. Second, that the parity of the number of negative edges but not their number is what determines balance in signed cycles. Third, that for sufficiently large *n* the presence of the negative edge becomes negligible, fulfilling our intuition that the umbalance is "diluted" in very large cycles. This statement is expressed in the following corollary.

Corollary 2 *Let C_n^{k-} be a signed cycle of length n and k negative edges. Then $\kappa_v\left(C_{n\to\infty}^{k-}\right) \to 1$.*

Finally, we give a result, also proved in Sect. 1 of the SM, concerning a type of graph whose structure is found in several practical applications, such as NIR, as we will soon show.

Definition 4 Let K_n be a complete graph with *n* nodes. Let $l \leq n-1$. Then, pick *l* nodes in K_n to form the set *S* and change the sign of their edge (v_i, v_j) to negative if and only if $\{v_i, v_j\} \subset S$. Then, the set *S* forms a negative clique of *l* nodes in K_n and we will denote the resulting graph as $K_n(K_l^-)$.

Proposition 6 *Let $K_n(K_l^-)$ be a graph defined as before. $\kappa_v\left(K_n(K_l^-)\right)$ tends to zero when $l \to \infty$ and $m := n-l$ remains finite.*

3.3 Local balance and diffusion of 'information'

We now show how the local balance index can be obtained in an alternative way as a result of a non-conservative diffusion process. It has been claimed that diffusion

[36] of information, generally speaking, in social systems can be a non-conservative process [37–39]. By non-conservative it is meant that the amount of information at a given time can be smaller/bigger than that at the initial time. Lerman and Ghosh [40] defined the non-conservative Laplacian of a network as a way to model such kind of diffusive processes. In the case of a signed network with adjacency matrix A the Lerman-Ghosh Laplacian could be read as: $\mathcal{L} := \chi I - A$ and for the unsigned version of the same network, we should define $\tilde{\mathcal{L}} := \chi I - |A|$, where χ is a scalar parameter and I is the identity matrix.

Then, the non-conservative diffusion on the signed graph is described by

$$\dot{u}(t) = -\mathcal{L}u(t), u(0) = u_0, \tag{10}$$

where $u(t)$ is the state vector at time t and u_0 is the initial state. The solution of Eq. (10) is then

$$u(t) = e^{-\mathcal{L}t}u_0 = e^{-\chi t}e^{tA}u_0.$$

This solution can be split into two terms:

$$e^{-\chi t}\left\{\begin{bmatrix} \left(e^{tA}\right)_{11}(u_0)_1 \\ \vdots \\ \left(e^{tA}\right)_{nn}(u_0)_n \end{bmatrix} + \begin{bmatrix} \left(e^{tA}\right)_{12}(u_0)_2 + \cdots + \left(e^{tA}\right)_{1n}(u_0)_n \\ \vdots \\ \left(e^{tA}\right)_{n1}(u_0)_1 + \cdots + \left(e^{tA}\right)_{nn-1}(u_0)_{n-1} \end{bmatrix}\right\},$$

where the first term accounts for the amount of diffusive information that departs from a node and returns to (or is retained at) it.

Obviously, we can write down the same equation for the unsigned version of the network, such that:

$$\dot{\tilde{u}}(t) = -\tilde{\mathcal{L}}\tilde{u}(t), \tilde{u}(0) = u_0,$$

with the same initial condition as before. Thus, $\tilde{u}(t) = e^{-\chi t}e^{t|A|}u_0$.

If we want to know how much information is "lost" in a non-conservative diffusion on a given network as a consequence of the existence of signed edges in it, we can calculate the ratio $u_v(t)/\tilde{u}_v(t)$ with initial condition $(u_0)_i = \delta_{iv}$ (where δ_{iv} denotes the Kronecker delta), such that

$$\frac{u_v(t)}{\tilde{u}_v(t)} = \frac{e^{-\chi t}\left(e^{tA}\right)_{vv}}{e^{-\chi t}\left(e^{t|A|}\right)_{vv}} = \frac{\left(e^{tA}\right)_{vv}}{\left(e^{t|A|}\right)_{vv}}.$$

Therefore, the local balance index can be interpreted as the ratio of the information "lost" in a non-conservative diffusion on a signed network when $t = 1$.

Finally, we remark that the parameter $t \in \mathbb{R}^+$ can also be interpreted as an overall weight that is applied to every edge of the network. Consequently, we can define a

weighted version of the local balance index:

$$\kappa_v (G, t) := \frac{\left(e^{tA}\right)_{vv}}{\left(e^{t|A|}\right)_{vv}}.$$

We have the following result:

Proposition 7 *Let $\kappa_v (G, t)$ be the local balance index of a node v in an unbalanced graph G, where every edge is weighted by a factor $t \in \mathbb{R}^+$. Then,*

$$\lim_{t \to \infty} \kappa_v (G, t) = 0, \tag{11}$$

$$\lim_{t \to 0} \kappa_v (G, t) = 1. \tag{12}$$

Proof On the one hand,

$$\lim_{t \to \infty} \kappa_v (G, t) = \lim_{t \to \infty} \frac{(a_1)^2_v e^{t\alpha_1}}{(b_1)^2_v e^{t\beta_1}} = \lim_{t \to \infty} \frac{(a_1)^2_v}{(b_1)^2_v} e^{t(\alpha_1 - \beta_1)} = 0,$$
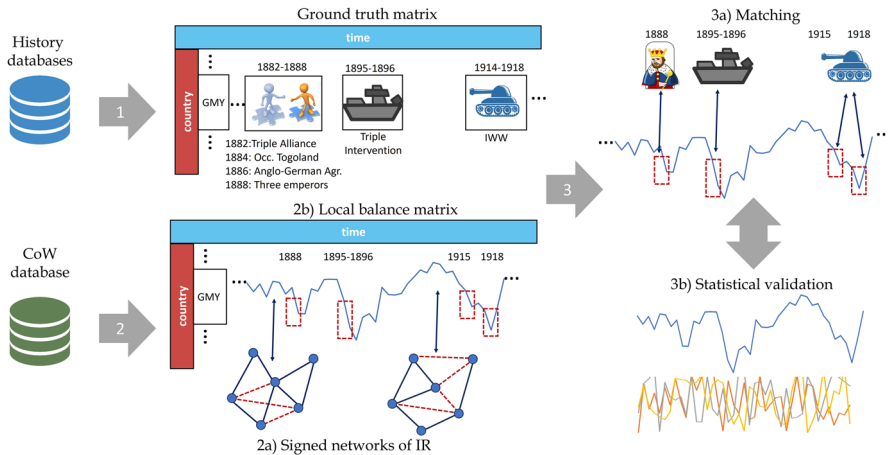
because $\beta_1 > \alpha_1$ (theorem 3.9 in [16]). On the other hand,

$$\lim_{t \to 0} \kappa_v (G, t) = \frac{\sum_{j=1}^{n}(a_j)^2_v}{\sum_{j=1}^{n}(b_j)^2_v} = 1.$$

$\square$

4 Part II. Application of local balance to global international relations

The main goal of this second part of the paper is to apply the local balance index to an in-depth analysis of the global international relations (IR) for the period between 1816 and 2014. While previous analysis of the IR for similar periods have focused on the study of the global balance of signed networks [7, 11], here we concentrate our analysis on the countries' role on this global balance. Therefore, by an in-depth analysis we refer to the fact that we will study the connection between the countries' local balance index–a purely topological index derived from signed networks of IR– with major historical events involving these countries in specific periods of time. For conducting this analysis we base the current study in the general methodological scheme illustrated in Fig. 2. We detail the methodological aspects of each step of this scheme in the following separate subsections.

Fig. 2 Illustration of the general scheme for relating the local balance index of signed networks with major historic events which have occurred in the World in the period 1816–2014. First, a ground truth matrix of events is created from databases of historic events for each of the 217 countries analyzed (1). Independently, the networks of international relations (2a) are created from the information provided at the CoW database and the time series of country's local balance are built (2b). Meaningful peaks and valleys in the time series of local balance are detected (2b). Then, a matching process (3a) allows to pair the peaks/valleys detected in the time series with the events recorded in the ground truth. The statistical significance (3b) of these matchings is then analyzed and reported

4.1 Methodology and statistical assessment

4.1.1 Construction of signed networks of IR

For every year in the period 1816–2014 we construct a signed network of IR, which corresponds to the step 2a in Fig. 2. The nodes of each network correspond to a subset of the 217 countries studied (see Table 5 in the SM for a list). The connections between the countries correspond to signed relations reported in databases different from the sources used for constructing the ground truth matrix. These databases were collected from [41, 42] for alliances, [43] for enmities, and [44] for strategic rivalries. The coding of inter-country relations is as follows. Two states are assigned a score of $+1$ at year t if they had any type of alliance (defense pact, offense pact, nonaggression pact, neutrality pact) during that year and zero otherwise. Consultation pacts (see [42]) are excluded. On the other hand, a dyad is assigned a score of -1 at time t if members had militarized interstate disputes at time t or were considered to be strategic rivals at this time according to the Colaresi et al. [44] data. Note that alliances may be asymmetric (see [41] for further explanation), while all enmities are symmetric. Since the number of asymmetric edges was very small, we converted data for each year into an undirected, unweighted, signed network where nodes represent countries and edges their relations in a given year. If two countries happen to be part of both an alliance and a conflict in the same year, we consider the conflict to be dominant, and therefore the edge to be negative.

Data are thus arranged in a temporal network spanning 199 years with a total of 217 nodes. We emphasize that the number of nodes is far from constant over time, as new states are created and old empires disappear. In fact, the number of nodes typically increases, from a minimum of 23 nodes in 1816 to a maximum of 195 nodes in 2014. The average number of nodes per year is 66. The edge statistics follows a similar trend: it increases from 88 links in 1816 to 8124 in 2014, with an average number of 1213 edges per year. On average, 41.22% of the edges are negative, although this number also fluctuates over time: a minimum of 1.49% negative interactions is reached in 2007, while the maximum of 97.30% is observed in 1880. It should also be noted that annual networks are typically not connected and often have many isolated nodes. The giant component represents on average 69.42% of the entire network, with a rather exceptional minimum of 33.33% in 1871 and a maximum of 95.38% in 2011 and 2013. Over the last 50 years, the giant component always exceeds 90% of the whole network. Since it selects the main actors on the geopolitical scene, the individual local balance values have been computed for such a component only.

4.1.2 Construction of ground truth matrix

First we create an empty matrix Γ with $N_c = 217$ rows, each one corresponding to one of the 217 countries analyzed in this work. Every one of the $N_t = 199$ columns corresponds to one of the years in the period 1816–2014 which are covered by this study. Then, we populate this matrix by using information from different sources about historic events involving a given country for specific years (see step 1 in Fig. 2). Notice that these sources of historic data are independent of those used before to create the NIR. Historical events are classified into two groups: 'peaceful' events and 'conflicting' events. Peaceful events are assigned a (+1) sign while conflicting events are assigned a (−1) sign. Assigning (+1) to peaceful events and (−1) to conflicting events is a conventional agreement. Reverting such assignment would not affect the obtained results. Peaceful events include:

1. The establishment of treaties, accords or agreements after wars;
2. The implementation of democratic procedures, like elections, change of regime or adoption of constitutions;
3. Countries becoming independent from their colonial domination.

On the other hand, conflicting events include:

1. The outbreak of an armed or trade conflict between two or more countries, or the entry of a country into an existing conflict, and in some cases clearly identified battles, invasions, or expeditions;
2. Revolutions, insurrections, or civil wars within a single nation, but with outside nations involved in support of certain warring factions. In this context, wars of independence from foreign domination, such as those in 19th-century Europe or 20th-century Africa, deserve special mention;
3. The effects of coups d'etat, which alter geopolitical arrangements and relations with foreign nations;
4. Crises in the broad sense, that is, the institutional, military, or economic collapse of a given nation brought about by international political conditions.

For example, given that Germany (GMY) is one of the participants in the Triple Alliance of 1882, we assign a (+1) sign to the entry corresponding to the pair (GMY, 1882) in the ground truth matrix. Similarly, due to the French involvement in the Sino-French War (1884–1885), we assign a (−1) sign to the pair (FRN, 1884).

To be consistent with the IR database (see next section), we only include events if the involved country appears in the IR database in the year of the event and in the previous one. For example, the Indian Rebellion of 1857 does not translate to a (−1) in (IND, 1857), because India is not considered an autonomous entity in the IR dataset until 1947 [notice, however, that this event does translate into a (−1) in the entry (UKG, 1857)].

In total, we have considered 1533 major historic events. The countries participating in the largest number of historic events are: UK (76), Russia (61), USA (61), France (52), Italy (45), Germany (44), Turkey (43) and China (42). They coincide with the major actors of history during the centuries XIX-XXI. There are 47 countries, mainly small ones, for which there are no recorded major historical event. The years with the largest number of events are: 1999 (44), 1992 (34), 1979 (34), 1989 (33), 1975 (28), 1991 (27), 1990 (26), 2011 (24), and 1967 (23). There are only 11 years for which there is no major historic event recorded, all of them in the XIX century.

4.1.3 Detection of peaks and valleys

Here we construct a local balance matrix T of dimensions $N_c \times N_t$ based on the detection of peaks and valleys, that is positive and negative changes, in the temporal series of the local balance index of every country in its signed network (see step 2b in Fig. 2). That is, for each year t, we extracted the signed adjacency matrix of the network and calculated the local balance $\kappa_v(t)$ for each country v. Afterwards, we aggregated the data from every year, obtaining a time series of the local balance values for each country (see, for instance, Fig. 1c). Then, we computed the difference in local balance for consecutive years, $\Delta\kappa_v(t) = \kappa_v(t) - \kappa_v(t-1)$, and identified the largest decreases or increases. We assign a (−1) sign to a country-year pair if $\Delta\kappa_v(t) \leq -0.1$ and $\kappa_v(t) \leq +0.5$, and a (+1) sign if $\Delta\kappa_v(t) \geq 0.1$ and $\kappa_v(t) \geq +0.5$. The conditions included for $\kappa_v(t)$, e.g. $\kappa_v(t) \geq +0.5$ and $\kappa_v(t) \leq +0.5$ respectively, guarantee that the system at the analyzed time has a significant degree of (un)balance. Since the method relies on differences in κ_v, events occurring in the year in which a country enters the dataset cannot be detected, as the balance for the year $t-1$ cannot be calculated.

We identified 1829 valleys and 834 peaks. However, the events are not homogeneously distributed over the time range. For instance, the countries with the largest number of valleys across the period studied are: Iran (34), Turkey (34), UK (32), USA (30), France (29), Yugoslavia (28), Italy (27), Greece (27), Japan (27), Spain (26), Romania (24) and Germany (24). There are 17 countries without any valley. On the other hand, the countries with the largest number of peaks are: USA (27), Italy (23), Romania (21), Greece (20), UK (19), Portugal (19) and Spain (18). There are 52 countries for which no peaks were detected based on the criterion used here.

4.1.4 Analysis of the matching of peaks/valleys with historic events

We are now in the position to test the similarities between the ground truth matrix Γ and the local balance matrix T. We employ two similarity criteria: the Pearson correlation coefficient between matrices Γ and T:

$$s_P(\Gamma, T) = \frac{Cov(\tilde{\Gamma}, \tilde{T})}{\sqrt{Var(\tilde{\Gamma})Var(\tilde{T})}} \tag{13}$$

(where $\tilde{\Gamma}$ and $\tilde{T}$ are one-dimensional vectors obtained by concatenating the rows of Γ and T in a standard matrix flattening procedure), and the Frobenius similarity of the pairs of matrices:

$$s_F(\Gamma, T) = \frac{tr(\Gamma^T T)}{\sqrt{tr(\Gamma^T \Gamma)tr(T^T T)}}. \tag{14}$$

The Frobenius similarity is the equivalent for matrices to the cosine similarity index between two vectors. In Sect. 2 of the SM, we give a more detailed description of these two similarity indices.

We need now to check whether the similarities found between the ground truth and local balance matrices are statistically significant. For this test we create the null model matrix T' as follows. For each country we randomize the corresponding time series 500 times, and then we construct the matrix T' by applying the methodology that was previously described for the matrix T. Given that through the time range analyzed, many countries have been created or have disappeared, the balance time series for many countries contains missing data, corresponding to the years in which those countries did not exist. Consequently, we exclude the missing values from the shuffling; that is, we keep the periods of missing data fixed and only randomize the time periods where the given country exists.

Using s_P and s_F as similarity indices, we find that the matrix of peaks and valleys and the ground truth matrix have a statistically significant similarity, ten times bigger than that of the randomized truth model (0.306 vs. 0.039 for Pearson and 0.313 vs. 0.052 for Frobenius). Both quantitative similarity measures indicate that the matrix of peaks and valleys based on the real time series of local balance is significantly more similar to the ground truth of historic events than its randomization. To test whether these similarity indices are significantly different or not we use the statistical tests included in [45]. All of them confirmed that the similarity with the real matrix is significantly larger than that with the randomized one. We have also tested the similarity between the absolute values of the matrices; i.e., indicating only the existence of a major event and not its nature, and we also obtain that the correlations with the real local balance matrix are statistically more significant than the correlations with the randomized matrix.
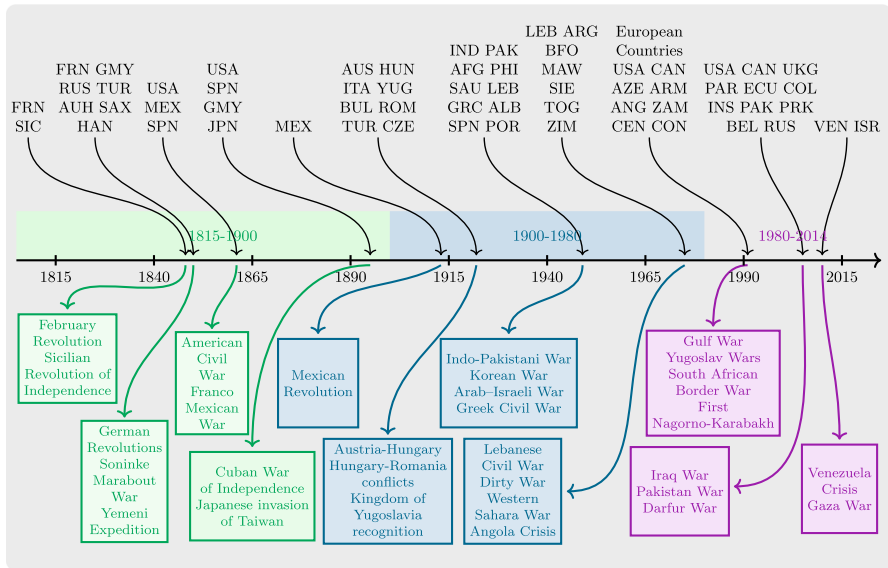
4.1.5 Correlations with GeoPolitical risk indices

As a further test of the statistical significance of the method developed here, we check for several individual countries the connection between the local balance and the perceived geopolitical risk. This analysis is based on the GeoPolitical Risk (GPR) index [46], which measures the frequency of newspaper articles that discuss adverse geopolitical events. We employ the historic variant, that aggregates the data of three different U.S. newspapers to provide monthly measures of the risk of 44 different countries between the years 1900 and 2023. Then, we calculate the Spearman's correlation coefficient [47] between GPR and local balance time series for the nine countries whose time overlap includes 100 or more years: China, France, United Kingdom, Italy, Japan, Portugal, Russia/USSR, Turkey and the USA. We also compare the correlation between the GPR time series and 500 randomizations of each local balance time series, following the procedure outlined in Sect. 3.1.4. The resulting rank correlation coefficients are provided in Table 2 in the SM, where we observe significant correlations for most of the countries. For instance, Spearman rank correlation for the USA, France and Portugal are -0.548, -0.414 and -0.367, respectively, while for the same countries in the analogous random networks the rank correlations are 0.000486, -0.001147 and 0.006732, respectively. Interestingly, the weakest correlations are found in non-European countries, a fact that suggests a possible Western bias in the U.S. newspapers used to compute the GPR index.
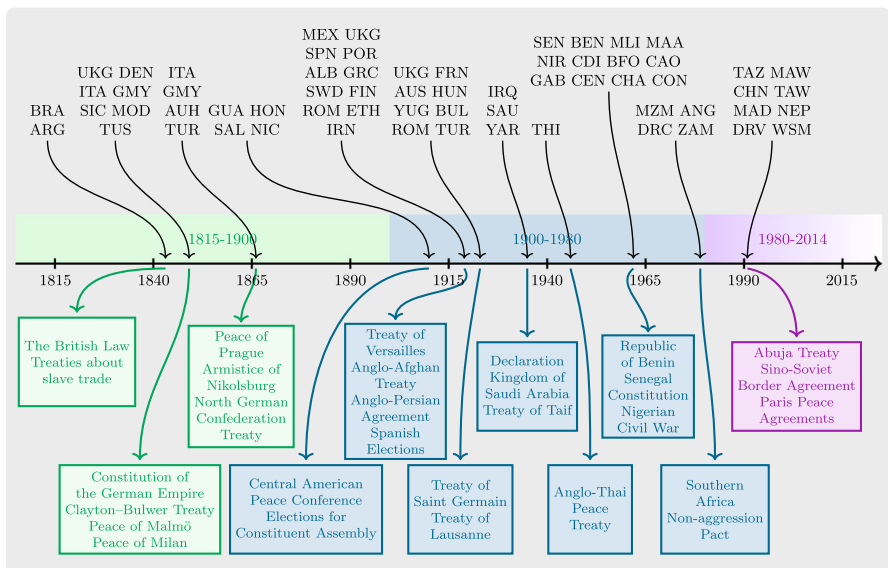
Based on the similarity criteria between the ground truth matrix and the matrix of peaks and valleys, and the confidence provided by the correlation of the local balance index with GPR we proceed to match the peaks and valleys with the corresponding historic events that match them in terms of countries and exact year of occurrence.

4.2 Linking local balance with historical narrative

Now that the statistical significance of the current method is confirmed, we can perform an *a posteriori* analysis of the matching between peaks and valleys in terms of the nature of the historic events. In Table 3 in the SM, we provide 12 examples of sharp drops in local balance, which include for instance a drop of 0.138 in 1848 for France due to the European revolutions taking place this year or a big drop of 0.422 in 1916 for Mexico due to the Mexican revolution. It is clear that the drop in balance happens exactly at the time when a conflicting historical event is taking place in the analyzed country. Similarly, in Table 4 in the SM, we give 12 instances of local balance increases, which include for instance a big peak of 0.417 for Nigeria in 1963 due to the creation of the first Nigerian republic or a peak of 0.407 for Austria–Hungary in 1866 due to the Armistice of Nikolsburg and the Peace of Prague reached this year. Again, we observe a match between the sharp increase in balance and the involvement of the country in treaties, protocols or independence events. In Fig. 3 we illustrate a timeline containing the main events corresponding to valleys detected by the current methodology in the XIX, XX and XXI century, while in Fig. 4 we gather the main events that correspond to peaks. In Sect. 3 of the SM, we include more detailed figures illustrating timelines for each of the centuries analyzed here (Figs 4, 5 and 6 in the SM.).

Fig. 3 Timeline of the major events corresponding to valleys detected in the XIX, XX and XXI centuries. A table relating each country to its corresponding code is provided in the SM

Fig. 4 Timeline of the major events corresponding to peaks detected in the XIX-XXI centuries. A table relating each country to its corresponding code is provided in the SM
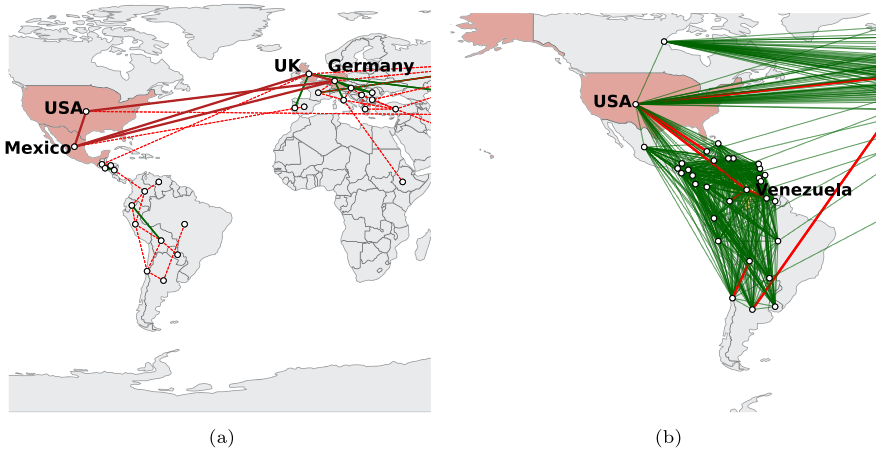
However, we still need to connect the quantitative analysis with a qualitative narrative of the historic events happening in those countries. To do so, we use here a "mixed method" based on the principles described in [31, 32]. Specifically, we retrospectively analyze the historical situation in which the individual countries were involved that year to create a narrative. This analysis reveals some well-known major historic events and their main players, but also some more subtle historic situations.

4.2.1 Historic events associated with valleys

As a first example that allows to bridge these quantitative findings with a historic narrative, we will focus on the period 1848–1850, which witnessed the largest and most widespread revolutionary wave in the history of Europe. Historians agree in attributing the February 1848 Revolution in France, which led to the collapse of Monarchy, as the spark that triggered the fire across Europe [48]. The node of France in the IR network is characterized by low values of κ_v in its whole recent history, but in the years preceding the revolution of 1848, the value of κ_v was around 0.40. In 1848 it drops to 0.25 and in 1850 to 0.13. In the meanwhile, the uprisings spread throughout Europe, in particular in the Italian and German states. In the early months of 1848, the Sicilian Revolution took place and the value of κ_v associated with the no longer existing Kingdom of the Two Sicilies plunges from a value of 0.96 in 1847 to a value of 0.22 in 1848.

In 1849 there is a general increase in the κ_v values for all states then in existence, but again in 1850 we see a new crash. In fact, the German revolutions of 1848–1849 attempted to transform the Confederation into a unified German Federal State; but in 1850 the Diet was re-established after the revolution was crushed by Austria and Prussia [49]. This is observed by the sharp drops in the values κ_v of many German States: the Kingdoms of Saxony and of Hanover both drop from 0.97 in 1849 to 0.41 in 1850, the Kingdoms of Bavaria and of Wurttemberg from 0.70 to 0.38 and the Grand Duchy of Baden from 0.70 to 0.38. Surprisingly, the links of Hanover, Saxony and the Grand Duchy of Baden in 1850, are all positive! And other German countries involved in such revolutions have at most one negative edge only. Therefore, the local balance index manages to capture the situation of local instability much better than other local network indicators, such as the number of negative edges incident to every country.

We now pick an example connecting our approach with historic narrative from the XX century. A dramatic drop in the value of the local balance index is witnessed by Mexico in 1913 and 1917. The κ_v values for Mexico in the early years of the XX century averaged around 0.96. However, in 1913 this value suddenly collapsed to 0.44 and to 0.37 in 1917. In 1912 Mexico's only link to the global network is through the U.S., which is a negative edge. In 1913 the IR of Mexico are described by the subgraph illustrated in Fig. 5a. The majority of the edges in Fig. 5a are negative. The imbalance observed here is predominantly produced by the two negative triangles Mexico-USA-Germany and Mexico-UK-Germany. Therefore, there is a key role played by Germany in unbalancing the IR concerning Mexico in 1913. If the Germany-Mexico link were positive, the balance of this subgraph would significantly increase. The connection between Germany and Mexico during this period has been "relatively unnoticed by historians until recently" [50]. The narrative of events is detailed in Sect. 3 of the SM

Fig. 5 Illustration of the IR involving (**a** Mexico in 1913 and **b** Venezuela in 2010. In this map and all the following ones, green edges represent alliances and red edges depict conflicts or tensions

and mainly refers to the efforts of Germany to draw Mexico into war with the United States with main goal of distract the last from an eventual European war.

Finally, we consider a historic scenario taking place in the XXI century. It refers to the Venezuelan economic crisis and how this destabilized its IR by dropping dramatically its local balance. Venezuela exhibits a very low local balance for the entire period, from the end of World War II to the present, with an average of 0.25. Nevertheless, in 2009 its value was 0.40, followed by a dramatic drop the year after to a value of 0.008, one of the lowest values ever recorded for all countries and for all years analyzed. In 2009, Venezuela was well established in the South American cluster with only two negative edges, one with Colombia and another with Guyana, and, very importantly, with a positive edge with the U.S. In 2009 the link with the U.S. represented for Venezuela a bottleneck through which it connected to the rest of the world, outside South America. Now, in 2010 it happens that this link suddenly becomes negative as can be seen in Fig. 5b, and this is the only difference with the previous years. All the edges within the South American cluster preserve their sign from 2009 to 2010. The events that sparked this situation are described in Sect. 3 of the SM and basically account for the economic crisis occurring in Venezuela in 2009–2010 and the subsequent deterioration of internal social situation which triggered sanctions from the European Union and United States.
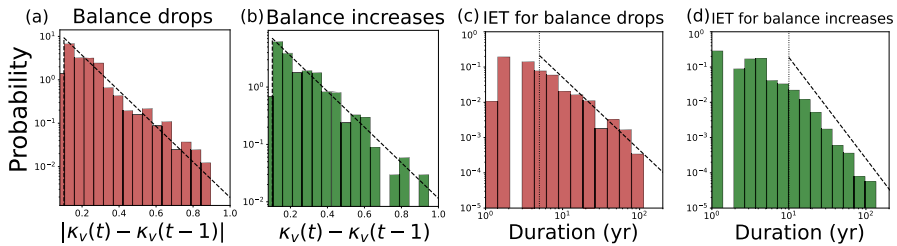
4.2.2 Historic events associated with peaks

Let us now perform an analysis for peaks similar to the one done before for valleys. We observe that events corresponding to peaks are more frequent in the XIX and XX centuries, with practically no one in the XXI century. Some of the detected events corresponding to peaks are illustrated in Fig. 4.

One of the historic periods showing significant incremental peaks in the local balance is the XIX century in South America. This was a very convulse period, which started with the independence of several of these regions from the Spanish and Portuguese empires, forming new states. It continues with the fights to balance power in the region, which also involves some global powers like Great Britain, France and Portugal. The local balance of Brazil increased from 0.58 in the previous year to 0.96, due to the removal of the negative edge of Brazil with Great Britain (maps in Fig. 1 in the SM). Prior to 1843, Brazil had a clear confrontation towards Great Britain due to the British efforts to suppress the slave traffic [51, 52]. In 1843 Argentina and Chile enter in conflict over Patagonia and the Strait of Magellan [53] adding volatility to the region. Argentina drops again its local balance in 1845-46, possibly due to its conflict with France and Great Britain. Only after 1847 the local balance of Argentina is recovered to $\kappa_i \approx 1$ due to negotiations to end the hostilities. Both Argentina and Brazil had territorial ambitions over Paraguay and Uruguay. Thus, a strategic alliance of Paraguay with Brazil against Argentina helped to increase the former's balance index in the South American network in the transition from 1850 to 1851. A similar strategic alliance with Peru allowed Venezuela to increase its balance from 0.42 in 1858 to 0.88 a year later [54].

Another dramatic increase of κ_v observed in this work occurred for Mozambique in 1977 and Angola in 1979. Mozambique had $\kappa_v \approx 0.008$ in 1976 and Angola had $\kappa_v \approx 0.005$ in 1978. They become 0.602 and 0.616, respectively, in 1977 and 1979, respectively. The events most likely associated with this increase in the local balance are the Angolan and the Mozambican civil wars. In Sect. 3 of the SM, we describe the narrative of these wars which were developed in a Cold War scenario in which a perfectly balanced situation of two-blocs occurs in southern Africa with the participation of many countries around the world.

4.3 Analysis of local balance time series

This section is dedicated to extracting empirical probability distributions from the time series of local balances for all countries. Specifically, we analyze two magnitudes that encapsulate key properties of the balance time series: the size of the balance drops and increases (i.e., the probability that the local balance experiences a drop or increase of size x, represented as $P(x) := P[\kappa_v(t) - \kappa_v(t-1) = x]$); and the inter-event time (IET) distribution for drops and increases (i.e., the probability that two consecutive drops/increases are separated by a time interval t). The analysis is based on the extraction of histograms that describe each distribution, followed by the identification of the most plausible probability distribution and an estimation of its parameters. To estimate the exponent of power-law distributions, we have used the maximum likelihood estimation algorithm implemented in the powerlaw python package [55]. We have limited the range of the power law distribution to the heavy-tail region; in other words, we have imposed that $x > x_{min}$, where x_{min} is another estimated parameter. To compare the goodness of fit of the exponential and power-law distributions, we have employed the Kolmogorov-Smirnov (KS) statistical test. Since we are interested in capturing the statistics of historically relevant events, we have

Fig. 6 Probability distribution of the size of the local balance peaks (**a**) and valleys (**b**) and of the inter-event times (IET) between two consecutive peaks (**c**) and between two consecutive valleys (**d**). We also plot the best fit to each empirical distribution (dashed lines). The dotted vertical line indicates the x_{min} after which we fit to a power-law distribution. (Color figure online)

filtered out of the size distribution any event for which $|\kappa_v(t) - \kappa_v(t-1)| < 0.1$. The results of this analysis are presented in Fig. 6.

In Fig. 6a, b, we illustrate the frequency of balance drops and increases as a function of their size. Both histograms represent an exponential distribution: the distribution of balance increases follows an exponential decay of the form $P(x) = \exp[-7.16(x - 0.1)]$ (p-value: $p < 10^{-6}$); while the distribution of balance drops decays as $P(x) = \exp[-9.43(x - 0.1)]$ ($p < 10^{-4}$). One possible mechanism that could generate such distributions is a scaled Poisson random process in which random events happen at a constant rate α. Indeed, it is well-known that the inter-event time distribution of any scaled Poisson process is given by $P(x) = \alpha e^{-\alpha x}$. This suggests that the mechanisms that determine the size of a given historical drop or increase event are not very different from a scaled Poisson random process.

In contrast, the distribution of inter-event times, shown in Fig. 6c, d, does not follow an exponential distribution. Instead, we find that power-law distributions fit the data more precisely: for the IET of balance increases, the data is best fit to $P(t) \sim t^{-2.9}$ (p-value: $p = 0.004$, $x_{min} = 10$); while for the balance drops IET we fit the histogram to $P(t) \sim t^{-2.1}$ ($p < 10^{-4}$, $x_{min} = 5$), where t is expressed in years. The estimated exponents for both fits are smaller than three, meaning that the estimated probability distributions have infinite variance. This suggests the existence of a "scale-free" IET distribution, similar to those found in several other socio-technical systems [56, 57]. Significant events are not observable in the XXI century, possibly because the "noise" level of the local balance (continuous increments and drops) makes them statistically not relevant.

4.3.1 Correlations in local balance time series

We now examine the correlations between the behavior of different countries in terms of local balance over time. The results are graphically illustrated in Fig. 7 in Sect. 3 of the SM. In panel (a), we show the Pearson correlation plot for the 'spatial distribution' of the $\kappa_v(t)$, i.e. for pairs of different years, computed on the subset of countries actually existing in both the years examined. In panel (b) we show the Pearson correlation plot for the time series $\kappa_v(t)$ for pairs of countries, computed on the intersection of their corresponding existence time intervals.

The correlogram in panel (a) reveals a clear historical caesura corresponding to the year 1949, after the Second World War: the years before 1949 are predominantly positively correlated with each other, as are the years after 1949; but the years in the two blocs are typically negatively correlated. This implies a noticeably different distribution of the local balance among countries before and after that year.

On the other hand, the correlogram in panel (b) shows a positive correlation in the time evolution of the individual balance of most countries. In fact, throughout the whole set of nations only 14.38% of the pairs of nations are negative correlated. Among the nations with the highest number of negative correlations we find many African countries. For example, Kenya is anticorrelated with 83 countries, Mauritania with 82, and Mali and Nigeria with 76. On the other hand, among the nations with the highest number of positive correlations we find Ukraine, Belarus, Armenia, Azerbaijan and Kazakhstan with only one anticorrelated country.

If we instead ask ourselves which pair of countries shows the largest correlation throughout their time series, we find that this pair is comprised of Spain and Portugal, with a value of the correlation coefficient $\rho = 0.899$. Another highly correlated pair of countries is France and the United Kingdom ($\rho = 0.765$). It is worth noting that the last major conflict between United Kingdom and France was that of the Napoleonic Wars (1793–1815), which finished one year before the starting year of the data set. Other pairs of maximally correlated countries are Russia with France ($\rho = 0.746$), Germany with Italy ($\rho = 0.759$), Italy with Turkey ($\rho = 0.799$), and Yugoslavia with Romania ($\rho = 0.806$).

One important global trend that can be acknowledged is the growth of correlation within specific clusters of nations in the second half of the XX century, after World War II. For instance, let us consider the founding member countries of NATO, on the one hand, and the Warsaw Pact, on the other. The mean internal correlation between the values of the local $\kappa_v(t)$ in the period 1945–2014 is $\rho = 0.890$ for the NATO cluster and $\rho = 0.899$ for the Warsaw Pact cluster. If we consider the whole interval 1816–2014, the mean correlations within the two clusters are $\rho = 0.635$ and $\rho = 0.835$, respectively. Thus, we conclude that NATO has contributed much more to changing the *status quo* than the Warsaw Pact, which just confirmed a pre-existing system of consistent behaviors. Let us remark that the role of supranational coalitions, such as NATO and the former Warsaw Pact, in stabilizing cooperation, reducing instability, and creating homogeneous behavior among coalition members has recently been highlighted, for example, in [58].

5 Conclusion

With this work we have been aimed at demonstrating how mathematics can help to reveal that "not everything in history is contingent and particular" [28]. We do so by using algebraic graph theory techniques, which combined with appropriate data analysis tools reveal patterns in history from the individual participation of countries in the global web of IR. For instance, the current work informs that the probability that two events are separated by a time t is best fit to the power law $p(t) \sim t^{-2.9}$. This means that while the probability that such events are separated by only 1–2

years is about 70%, it drops to nearly 10% for intervals of 2–3 years and it is almost negligible for separations longer than 5 years. We have found that the main reason why such events occur so frequently is because they are populated by different countries. Indeed, if $p(N)$ is the probability that a given country is involved in N major events dropping the balance in the whole period of 199 years, then $p(N) \sim N^{-1.405}$. This means that the probability that a given country is involved in only one of such events is about 42%, but it drops to 16% for being involved in two events in the whole period. The probability that one single country is involved in 10 events in 199 years is only 2%. But history makes exceptions! Chile and Argentina were involved in 15 events, Peru in 16, Brazil in 19 and Spain in 20, mainly due to the convulsive situation existing in South America in the XIX century analyzed before. To sum up, the local balance index captures information about how an individual country reacts to the change of balance of the rest of states in the world at a given time. Therefore, the local balance index can be thought of as a quantitative candidate for translating mathematically what historians and political scientists call the "balance of power".

Declarations

Conflict of interest The authors declare no conflict of interest.

References

1. Estrada, E.: What is a complex system, after all? Found. Sci. (2023). https://doi.org/10.1007/s10699-023-09917-w
2. Wasserman, S., Faust, K.: Social network analysis: methods and applications. Cambridge University Press, Cambridge (1994)
3. Estrada, E.: The structure of complex networks: theory and applications. Oxford University Press, Oxford (2012)
4. Zaslavsky, T.: Signed graphs. Discrete Appl. Math. **4**, 47–74 (1982)

5. Harary, F.: On the notion of balance of a signed graph. Mich. Math. J. **2**, 143–146 (1953)
6. Cartwright, D., Harary, F.: Structural balance: a generalization of Heider's theory. Psychol. Rev. **63**(5), 277–293 (1956)
7. Kirkley, A., Cantwell, G.T., Newman, M.E.: Balance in signed networks. Phys. Rev. E **99**(1), 012320 (2019)
8. Kunegis, J., Schmidt, S., Lommatzsch, A., Lerner, J., De Luca, E.W., Albayrak, S.: Spectral analysis of signed graphs for clustering, prediction and visualization. In: Proceedings of the 2010 SIAM International Conference on Data Mining (SDM), pp. 559–570. Society for Industrial and Applied Mathematics (2010)
9. Aref, S., Dinh, L., Rezapour, R., Diesner, J.: Multilevel structural evaluation of signed directed social networks based on balance theory. Sci. Rep. **10**(1), 15228 (2020)
10. Estrada, E., Benzi, M.: Walk-based measure of balance in signed networks: detecting lack of balance in social networks. Phys. Rev. E **90**, 042802 (2014). https://doi.org/10.1103/PhysRevE.90.042802
11. Estrada, E.: Rethinking structural balance in signed social networks. Discrete Appl. Math. **268**, 70–90 (2019). https://doi.org/10.1016/j.dam.2019.04.019
12. Talaga, S., Stella, M., Swanson, T.J., et al.: Polarization and multiscale structural balance in signed networks. Commun. Phys. **6**, 349 (2023). https://doi.org/10.1038/s42005-023-01467-8
13. Harary, F.: On local balance and *n*-balance in signed graphs. Mich. Math. J. **3**(1), 37–41 (1955)
14. Heider, F.: Attitudes and cognitive organization. J. Psychol. **21**, 107–12 (1946)
15. Heider, F.: The psychology of interpersonal relations. Psychology Press, USA (2013)
16. Tian, Y., Lambiotte, R.: Spreading and structural balance on signed networks. SIAM J. Appl. Dyn. Syst. **23**(1), 50–80 (2024). https://doi.org/10.1137/22M1542325
17. Maoz, Z., Terris, L.G., Kuperman, R.D., Talmud, I.: International relations: a network approach. In: Mintz, A., Russett, B. (eds.) New directions for international relations: confronting the method-of-analysis problem, pp. 35–64. Lexington Books, Lanham, Maryland (2005)
18. Hafner-Burton, E.M., Kahler, M., Montgomery, A.H.: Network analysis for international relations. Int. Organ. **63**(3), 559–592 (2009)
19. Doreian, P., Mrvar, A.: Structural balance and signed international relations. J. Soc. Struct. **16**, 1–49 (2015)
20. Maoz, Z.: Network science and international relations. In: Thompson, W.R. (ed.) Oxford research encyclopedia of politics. Oxford University Press, Oxford (2017). https://doi.org/10.1093/acrefore/9780190228637.013.517
21. Kacziba, P.: The network analysis of international relations: overview of an emergent methodology. J. Int. Stud. **14**, 155–171 (2021)
22. Harary, F.: A structural analysis of the situation in the Middle East in 1956. J. Confl. Resolut. **5**(2), 167–178 (1961)
23. Vinogradova, G., Galam, S.: Dissolution of a global alliance. Policy Complex Syst. **1**(1), 93–107 (2014). https://doi.org/10.18278/jpcs.1.1.5
24. Vinogradova, G., Galam, S.: Global alliances effect in coalition forming. Eur. Phys. J. B (2014). https://doi.org/10.1140/epjb/e2014-50264-4
25. Gauthier, L.: Putting clio back in cliometrics. Hist. Theory **61**(2), 289–311 (2022)
26. Maxwell, J.A., Miller, B.A.: Categorizing and connecting strategies in qualitative data analysis. In: Handbook of Emergent Methods, pp. 461–476. Guilford Press, New York (2008)
27. Anderson, M.: Quantitative history. In: The Sage Handbook of Social Science Methodology, pp. 246–263. Sage Publications, London (2007)
28. Turchin, P.: Arise 'cliodynamics'. Nature **454**, 34–35 (2008)
29. Galam, S.: Fragmentation versus stability in bimodal coalitions. Phys. A Stat. Mech. Appl. **230**(1), 174–188 (1996). https://doi.org/10.1016/0378-4371(96)00034-9
30. Tamura, E.H.: Narrative history and theory. Hist. Edu. Q. **51**(2), 150–157 (2011)
31. Johnson, R.B., Onwuegbuzie, A.J.: Mixed methods research: a research paradigm whose time has come. Edu. Res. **33**(7), 14–26 (2004). https://doi.org/10.3102/0013189X033007014
32. Onwuegbuzie, A.J., Johnson, R.B., Collins, K.M.: Call for mixed analysis: a philosophical framework for combining qualitative and quantitative approaches. Int. J. Mult. Res. Approaches **3**(2), 114–139 (2009)
33. Higham, N.J.: Functions of matrices: theory and computation. SIAM, USA (2008)
34. Estrada, E., Rodríguez-Velázquez, J.A.: Subgraph centrality in complex networks. Phys. Rev. E **71**(5), 056103 (2005)

35. Akbari, S., Belardo, F., Dodongeh, E., Nematollahi, M.A.: Spectral characterizations of signed cycles. Linear Algebra Appl. **553**, 307–327 (2018)
36. Masuda, N., Porter, M.A., Lambiotte, R.: Random walks and diffusion on networks. Phys. Rep. **716**, 1–58 (2017)
37. Zeng, A., Yeung, C.H.: Predicting the future trend of popularity by network diffusion. Chaos Interdiscip. J. Nonlinear Sci. (2016). https://doi.org/10.1063/1.4953013
38. Liu, L., Tang, J., Han, J., Yang, S.: Learning influence from heterogeneous social networks. Data Min. Knowl. Discov. **25**, 511–544 (2012)
39. Ghosh, R., Lerman, K., Surachawala, T., Voevodski, K., Teng, S.H.: Non-conservative diffusion and its application to social network analysis. Soc. Inform. Netw. (2011). https://doi.org/10.48550/arXiv.1102.4639
40. Lerman, K., Ghosh, R.: Network structure, topology, and dynamics in generalized models of synchronization. Phys. Rev. E **86**(2), 026108 (2012)
41. Leeds, B.A., Ritter, J.M., Mitchell, S.M., Long, A.G.: Alliance treaty obligations and provisions, 1815–1944. Int. Interact. **28**(3), 237–260 (2002)
42. Leeds, B.A.: Alliance treaty obligations and provisions (ATOP) codebook. Rice University, http://atop.rice.edu/home (2005)
43. Maoz, Z., Johnson, P.K., Kaplan, J., Ogunkoya, F., Shreve, A.P.: The dyadic militarized interstate disputes (MIDs) dataset version 3.0: logic, characteristics, and comparisons to alternative datasets. J. Confl. Resolut. **63**(3), 811–835 (2019)
44. Colaresi, M.P., Rasler, K., Thompson, W.R.: Strategic rivalries in world politics: position space and conflict escalation. Cambridge University Press, Cambridge (2008)
45. Diedenhofen, B., Musch, J.: cocor: a comprehensive solution for the statistical comparison of correlations. PloS one **10**(4), 0121945 (2015)
46. Caldara, D., Iacoviello, M.: Measuring geopolitical risk. Am. Econ. Rev. **112**(4), 1194–1225 (2022)
47. Spearman, C.: The proof and measurement of association between two things. Am. J. Psychol. **15**(1), 72–101 (1904)
48. Strandmann, H.P., Evans, R.J.W.: The revolutions in Europe, 1848–1849: from reform to reaction. Oxford University Press, Oxford (2000)
49. Hahn, H.J.: The 1848 revolutions in German-speaking Europe. Longman, USA (2001)
50. Leffler, J.J.: Germany, Mexico, and the United States, 1911–1917. In: PhD Dissertation, University of Rhode Island (1982). https://digitalcommons.uri.edu/dissertations/AAI8301334/
51. Adams, J.E.: The abolition of the Brazilian slave trade. J. Negro Hist. **10**(4), 607–637 (1925)
52. Campbell, C.J.: Making abolition Brazilian: British law and Brazilian abolitionists in nineteenth-century Minas Gerais and Pernambuco. Slavery Abolit. **36**(3), 521–543 (2015)
53. Burr, R.N.: The balance of power in nineteenth-century South America: an exploratory essay. Hisp. Am. Hist. Rev. **35**(1), 37–60 (1955)
54. Kroupa, M.A.: International trade relations of Venezuela. In: Master's Thesis, Loyola University Chicago (1942). https://ecommons.luc.edu/luc_theses/251
55. Alstott, J., Bullmore, E., Plenz, D.: powerlaw: a python package for analysis of heavy-tailed distributions. PLOS ONE **9**, 85777 (2014)
56. Barabási, A.-L.: The origin of bursts and heavy tails in human dynamics. Nature **435**(7039), 207–211 (2005)
57. Karsai, M., Kivelä, M., Pan, R.K., Kaski, K., Kertész, J., Barabási, A.-L., Saramäki, J.: Small but slow world: How network topology and burstiness slow down spreading. Phys. Rev. E **83**(2), 025102 (2011)
58. Galam, S.: The dynamics of alliances: the case of ukraine and russia. J. Comput. Sci. **71**, 102058 (2023). https://doi.org/10.1016/j.jocs.2023.102058