Modeling Small Systems Through the Relative Entropy Lattice

YAIR NEUMAN[1] AND DAN VILENCHIK [ID] [2]
[1]The Department of Brain and Cognitive Sciences and Zlotowski Center for Neuroscience, Ben Gurion University of the Negev, Beersheba 8410501, Israel
[2]The Department of Communication Systems Engineering, Ben Gurion University of the Negev, Beersheba, Israel

Corresponding author: Yair Neuman (yneuman@bgu.ac.il)

ABSTRACT There are certain contexts, where we would like to analyze the behavior of small interacting systems, such as sports teams. While large interacting systems have drawn much attention in the past years, let it be physical systems of interacting particles or social networks, small systems are short of appropriate quantitative modeling and measurement tools. We propose a simple procedure for analyzing a small system through the degree in which its behavior at different granularity levels (e.g., dyads) non-linearly diverges from the simple additive behavior of its sub-units. For example, we may model the behavior of a soccer team by measuring the extent to which the behavior changes when we move from individual players to dyads, triads, and so on. In this paper, we address the challenge of modeling small systems in terms of measuring divergence from additivity at different granularity levels of the system. We present and develop a measure for quantifying divergence from additivity through what we term a *Relative Entropy Lattice*, and illustrate its benefits in modeling the behavior of a specific small system, a soccer team, using data from the English Premier League. Our method has practical implications too, such as allowing the coach to identify "hidden" weak spots in the team's behavior.

INDEX TERMS Modeling, data analysis, small systems, measurement, non-additivity, relative entropy, Kullback-Leibler divergence, sport analytics.

I. INTRODUCTION

The "emerging" [1] or "synergetic" [2] behavior of systems has been intensively studied in various disciplines but there is a lack of appropriate tools for quantitatively analyzing small interactive systems. For instance, a common approach for studying emergent behavior is to consider it in terms of "phase transition". Haken [2] epitomizes this approach by proposing to model "synergetic" behavior in terms of phase transition. However, there are two difficulties with this approach for studying what may be described as the divergence from additivity of some small systems. First, small systems, from sports teams to families cannot be studied using the conceptual and mathematical tools of statistical mechanics as by definition the scope of statistical mechanics is large populations. Second, the idea of measuring phase transition involves the study of an observed qualitative change in the system. In other words, there are contexts in which

we focus our efforts on positively identifying a qualitative change. There are some other contexts though, in which we would simply like to measure the extent to which the system's behavior continuously scales, with no specific focus on a qualitative change. For instance, it has been recently shown that the entropy of a soccer team, as measured by its ball passing behavior, is correlated with its rank at the end of the season in a way that a more ordered pattern of ball passes is positively correlated with a better rank [3]. In this context, we may want to measure the extent to which the ordered behavior of players scales up when moving from individual players to interacting dyads, triads and so on. The shortcoming in methodological tools for measuring the behavior of small systems, might have detrimental consequences for our ability to understand and control them. For example, in trying to form a winning basketball team, a coach must be able to model the scaling behavior of different combinatorial "blocks" of players that he may recruit to the team. This process must go beyond the easily measured performance of individual players and has heavy professional and

financial consequences. To the best of our knowledge, such a process is usually guided by intuition and rules of thumb, without scientific and interpretable tools that may support a rational process of decision making.

From the perspective of statistical entropy, the emerging behavior, previously discussed, may be conceptualized in term of divergence from additivity, where additivity means that the entropy of a system, composed of two statistically independent subsystems, is the sum of the entropy of the subsystems [4], [5]. The additivity of entropy is evident only in cases of non-interacting systems. Therefore the emerging behavior of a system composed out of interacting sub-systems may be better conceptualized in terms of non-additivity (or divergence from additivity), that may be generally defined as: "the nonlinear function associating the system's granularity level and its components' degrees of freedom" [6]. As entropy is the amount of information encoded in the system, we may measure the extent to which the information encoded in the system diverges from the additivity of the information encoded in its sub-systems. For example, we may model the behavior of a married couple by measuring the entropy of their free time activities. This behavior diverges from additivity in the sense that its entropy is quite different from the one of its sub-units: the husband and the wife. In contrast, the degree of divergence from additivity is probably smaller when observing the free time activities of the husband and his best friend. It is still an open question how to measure the divergence from additivity of small systems like families or sports teams.

A. OUR CONTRIBUTION

The aim of this paper is to address the above challenge by proposing a simple and interpretable procedure for measuring divergence from additivity at different granularity levels of the system. The general idea is that we may approximate the information encoded in each micro-state of the system, such as two soccer players, by using the information provided by its lower level micro-states (e.g. the individual players) in a way that may range from trivial additivity of two independent and non-interacting sub-systems/players to their maximum divergence from additivity.

We develop an idea originally proposed by [6], and use it to develop the concept of a *Relative Entropy Lattice* (abbreviated as ReL) as a tool for modeling divergence from additivity in the behavior of small-scale systems. First, we introduce the basic idea of measuring "divergence from additivity" (ibid) through Relative Entropy. Next, we develop the idea of ReL as a descriptive tool for modeling this process and illustrate it using the data of soccer teams. It must be emphasized that we have no specific interest in measuring the behavior of sports teams and use the context of sports analytics only for illustrating our general approach. Finally, we show how the idea of ReL may be used both as a descriptive and predictive tool for better understanding the behavior of small systems. We illustrate this point by analyzing the behavior of soccer teams from the English premier league and conclude by

discussing the benefits of the proposed measure and methodology for the modeling and data analysis of small systems.

II. METHODOLOGY
A. MEASURING DIVERGENCE FROM ADDITIVITY

A small system may be defined as comprised of a small number of interacting components that exhibit a change of behavior as a function of the granularity level in which we examine the system. Human language epitomizes this unique behavior in the context of "semantic transparency" [7], that reflects the extent to which the meaning of a linguistic compound is a function of the meanings of its basic components [8]. For example, the meaning of the bi-gram "hotdog" cannot be trivially reduced to the meaning of its components "hot" and "dog".

We analyze a system by partitioning the set of its interacting components and by varying the granularity level of the set partition. The components are represented through their vector of interactions with each other. For example, the behavior of a soccer player A may be represented as the normalized vector of his ball-passes to the other players in his team. Along the same line, the behavior of players A and B, $(A + B)$, is represented as the vector of ball passes both have made to the rest of the players. Such a procedure of adding the vectors of the players' behavior is analogous to the graph operation of merging two vertices into a new node whose neighbor set is the union of each vertex's neighbor set.

Our motivation comes from the algebraic notion of co-product [9], [10]. The co-product, or sum of objects, is defined as follows: a co-product of objects A and B is a new object $A + B$ together with a pair of maps ($i_A : A \rightarrow A + B$, $i_B : B \rightarrow A + B$) such that for any pair of maps of the form $f : A \rightarrow C, g : B \rightarrow C$ there is precisely one arrow $[f, g] : A + B \rightarrow C$ that makes the diagram in Figure 1 commute in such a way that $[f, g] \circ i_A = f$ and $[f, g] \circ i_B = g$. $[f, g]$ is called the co-product arrow of f and g with respect to the injections i_A and i_B [9] [p.54].
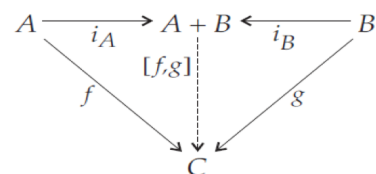

FIGURE 1. A schematic description of the co-product.

Informally, the co-product is a way of seeing the more general object from the perspective of the particular objects as a sum of its parts. In this sense, it is the "least specific" object to which the objects admit morphisms. We may measure the extent to which the ensemble is a co-product, or the simple sum of its constituents, by measuring the extent to which the co-product divergence from commuting in the above sense. In [6] it was proposed to use the Kullback-Leibler Divergence, also known as Relative Entropy to measure divergence

from additivity,

$$D_{KL}(P\|Q) = \sum_{x \in \Omega} P(x) \log \frac{P(x)}{Q(x)},$$

where P and Q are distributions.

For concreteness, let's assume that we would like to measure the divergence from additivity of two soccer players A and B who are members of a five players group. We represent the behaviour of each player as the vector of his ball-passes to the other players. We represent the behaviour of the dyad $A + B$ as the sum of their ball-passing vectors. The behaviour of $A + B$ can be analyzed from the perspective of its components, A and B, and the behaviour of some coarser partition of the players' set such as the triad of players: $A + B + C$. See Figure 2 where the numbers in the square brackets represent the values of the non-normalized vector of ball-passes to the other players:
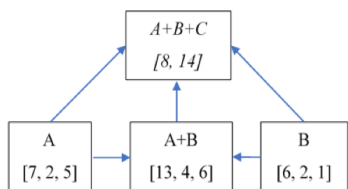

FIGURE 2. The ball-passes example.

In this context, the extent to which the partition (or micro-state) A+B is the co-product of the individual players A and B, can be measured by the extent to which the path from A to $A + B + C$ is somehow equal to the path from A to $A + B$ to $A + B + C$ and the same holds for the right triangle of the figure. For abbreviation we omit the '+' sign and let for example $A + B + C$ be ABC. We can model each path using the Kullback-Lebibler divergence and by averaging the absolute degree of:

$$D_{KL}(ABC\|AB) + D_{KL}(AB\|A) - D_{KL}(ABC\|A)$$

and

$$D_{KL}(ABC\|AB) + D_{KL}(AB\|B) - D_{KL}(ABC\|B).$$

Note that ABC and AB are vectors of different lengths. Therefore to compute $D_{KL}(ABC\|AB)$ we take their mutual support, namely, we ignore that ball passes of A and B to C in the vector AB.

In the next section we generalize this example and introduce our main measure, the $\mathfrak{w}$ measure.

B. THE $\mathfrak{w}$-MEASURE OF DIVERGENCE FROM ADDITIVITY AND REL

The formalism presented by [6] is somehow embryonic and therefore we chose to elaborate it for forming a more elegant and parsimonious measure. Let S be a set of n elements, where each element is represented as a vector of an interacting behaviour (e.g. ball-passes). Each subset of S is called a *micro-state* of the system. The *macro-state* of the system at

layer i is the set of all micro-states of size i. For example, we may consider a system of all pairs of soccer players as a macro-state of size 2, a set of all triads of players as a macro-state of size 3 and so on.

We use the Hebrew letter $\mathfrak{w}$ (Shin) for symbolizing divergence from additivity. $\mathfrak{w}$ stands for *Shi-tu-fi-yut* which means in Hebrew collaboration/cooperation. Somewhat abusing notation, we use the same letter, say V, to refer both to a micro-sate (i.e. the specific set of players, e.g. $V = A + B + C$ in the example we gave) and to the restricted vector of balls-passes from V to the rest of the team.

Before defining $\mathfrak{w}$, we introduce the auxiliary function d which is defined for every three micro-states U, V, W satisfying $U \subset V \subset W$,

$$d(U, V, W) = |D_{KL}(W\|V) + D_{KL}(V\|U) - D_{KL}(W\|U)|.$$

The function d measures the divergence from additivity of the triangle $U \rightarrow V \rightarrow W$, similarly to the triangle $A \rightarrow AB \rightarrow ABC$ in Figure 1.

The $\mathfrak{w}$-measure of micro-state V of size $|V| = i$, is given by:

$$\mathfrak{w}(V) = \frac{1}{i(N-i)} \sum_{\substack{U:U \subset V \\ |U|=i-1}} \sum_{\substack{W:V \subset W \\ |W|=i+1}} d(U, V, W) \tag{1}$$

The term $i(N-i)$ is a normalization factor which equals the total number of different pairs U, W that satisfy $U \subset V \subset W$, and $|U| = i - 1$, $|W| = i + 1$.

The $\mathfrak{w}$ measure quantifies the degree in which the partition V is the co-product of its U-components. Using it we can analyze a small system of N interacting components by computing the $\mathfrak{w}$ measure for each micro-state whose level is higher than 1 and lower than the entire set, N.

For forming the Relative Entropy Lattice (ReL), we stack the macro-states in layers, according to their size i, and compute for each micro-state its $\mathfrak{w}$ value. Figure 3 presents the ReL of the four players in Leicester who interact most with the others, and the accompanied $\mathfrak{w}$ score colored in red.
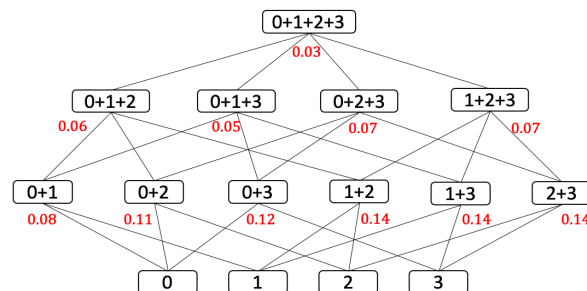

FIGURE 3. Relative Entropy lattice (ReL) for the leading 4 players in Leicester in terms of ball-passing statistics. Below each micro-state, in red, is its $\mathfrak{w}$ value computed via Eq. (1).

For illustrating the benefits of the divergence from additivity measure and ReL, we may turn to the data analysis section.

III. EVALUATION
As we scale up in a small system composed out of interacting units, constraints are accommodated to the extent that micro-states may become less and less informative about the coarser

macro-states in which they are included. If this hypothesis is grounded then we should observe a clear pattern of decay in the $\mathbb{w}$ measure as we scale up in the granularity level of the soccer team from dyads, to triads and so on. This trend is portrayed in the Relative Entropy Lattice in Figure 3 for $\mathbb{w}$ values of micro-states induced by the top four players in Leicester. In Analysis 1, Section III-B, we test this hypothesis in depth and detail.

A. THE DATA

The English Premier League[1] is the most watched sports league in the world. It involves 20 soccer clubs competing each season. During a season, each club/team plays against the others twice for a total of 38 games in a season. We have purchased from Perform group the data of the 2015/2016 season, describing ball passes between the players of each team. The data was provided for each game at the team's level of analysis. We have chosen this season, as a clear underdog, Leicester City, won the championship for the first time in their 132-year history, a victory that has been described by commentators as one of the greatest sports shocks in history.

B. ANALYSIS 1

Each team consists of about $N = 15$ players (including substitutes). For each team, we measure its $\mathbb{w}_i$-value, which is the average $\mathbb{w}$ value of macro-state i. In other words, it is the average of $\mathbb{w}(V)$ over all micro-states V of size i. Formally,

$$\mathbb{w}_i = \frac{1}{\binom{N}{i}} \sum_{V:|V|=i} \mathbb{w}(V). \tag{2}$$

Note that $\mathbb{w}_i$ is defined for $2 \leq i \leq N - 1$.
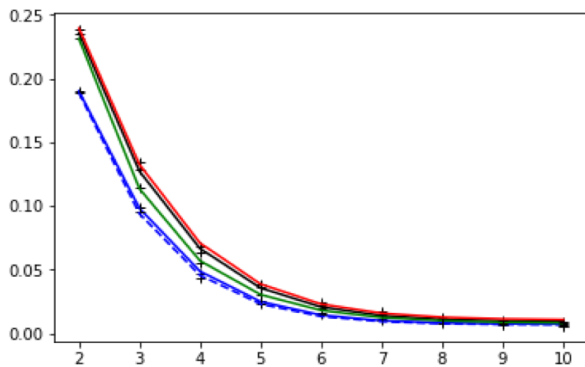
FIGURE 4. Stretched-exponential decay rate of $\mathbb{w}_i$ as a function of granularity level i. Red line – Leicester City (ranked 1), black - Crystal Palace (ranked 15), green - the average $\mathbb{w}_i$ over all 20 teams, solid blue - Tottenham Hotspur (ranked 3), dashed blue - Manchester United (ranked 4).

We have plotted the divergence from additivity score by granularity level and analyzed the extent to which the $\mathbb{w}_i$ measure decays as a function of the system's granularity level i. Figure 4 presents the overall results across all of the teams

[1] https://en.wikipedia.org/wiki/2015%E2%80%9316_Premier_League

used in the experiment plus illustrated results for four teams selected for representing different ranks in the final table.

What we see is that as i grows, namely the coarser is the level of analysis, the smaller the divergence from additivity is. In other words, when "zooming out", we see less and less divergence from additivity. Though eye goggling, we can easily identify a pattern of exponential decay. The decay pattern can be modeled using a stretched-exponential function. Concretely, for every team we fitted the set of points $\{(i, \mathbb{w}_i) : i = 2, 3, \ldots, 10\}$ using an stretch-exponential function of the following form:

$$f_{a,b,c,d}(i) = a \exp\{bx^c\} + d \tag{3}$$

and used least-squares minimization to find the parameters a, b, c, d. The fitted values of the average curve are $a = 0.9$, $b = -0.66$, $c = 1.06$, $d = 0.008$. While a, b, c govern the rate of decay, the parameter d is the limiting value of $\mathbb{w}_i$ as i goes to "infinity".

This finding is interesting, as the stretched exponential function, which has been identified several times in sociophysics, e.g. [11], [12], has been traditionally used in physics for describing the relaxation, or the returned to equilibrium, of a perturbed system. The relaxation property of the system will be discussed later.

Next we study the standard deviation between the $\mathbb{w}_i$-values computed across all teams, as a function of i, the granularity level. We have found that the behavior of the standard deviation as a function of i is also governed by a stretched-exponential law of the from given by Eq. (3). Concretely, we fitted a stretched-exponential function to the set of points $\{(i, std(\mathbb{w}_i)) : i = 2, 3, \ldots, 10\}$ using least-squares minimization and obtained the following fitted values $a = 0.03$, $b = -0.14$, $c = 1.65$, $d = 0.001$, see Figure 5.
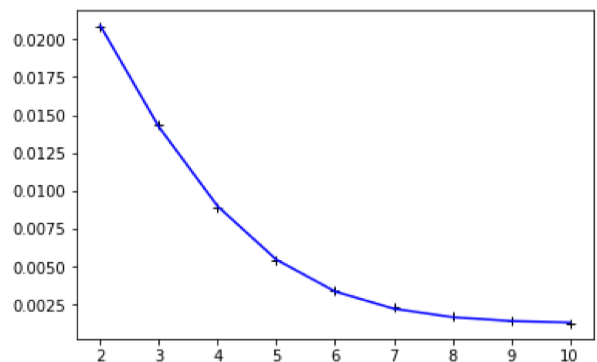
FIGURE 5. Stretch-exponential decay of variance of $\mathbb{w}_i$ across all 20 teams as a function of granularity level i.

The results presented so far invite two additional major analyses:

(a) First we ask whether the decay rate itself (the derivative of the team's exponential-stretch function, given by the parameters a, b, c in Eq.(3)) is a good predictor of the team's rank. From the immune system [13] to natural intelligence [14] to various combinatorial constraint satisfaction

problems [15], [16], real-world interactions are bounded in space and time, and constraints are formed on a local to a global base. The constraints coordinate the behavior of the system from perturbed and maximal degrees of freedom on the micro level to a more constrained and coordinated behavior as we zoom out globally. In this context, relaxation may be interpreted as the dynamics through which the system moves from local to coarser level interactions characterized by more constraints and therefore by a more ordered behavior. If this interpretation is reasonable then we should expect that the decay rate of the divergence from additivity measure for each group, should predict its rank at the end of the season. This analysis is given in Section III-C.

(b) Figure 5 suggests that there is variance in the ϖ_i-values across different teams, which decays as i grows. We check how well ϖ_i predicts the team's rank, for increasing values of i. As i grows, there is a trade-off between the growing number of micro-states V taking part in the computation of ϖ_i, Eq. (2) (there are $\binom{N}{i}$ states of size i, which increases with i up to $i = N/2$), but each particular micro-state V is perhaps less informative for the task of rank prediction (as the variance of ϖ_i decreases with i). This analysis is carried out in Section III-D.

C. ANALYSIS 2

In this section we test the hypothesis that the decay rate of ϖ_i, the divergence from additivity at level i, is a good proxy to the team's ability to coordinate its behavior across its levels, and as a result to its rank. The parameters that control the rate in which ϖ_i decays are b and c in the exponential stretch function Eq. (3). Regressing the rank on b and c yielded $R^2 = 0.46$ with p-value 0.007. This R^2 value is higher than the one gained by regressing on the team's position at the end of the previous season [3]. In line with our hypothesis, this finding shows that the faster the divergence from additivity decays, the better the team performs.

One can approximate the decay rate, encoded in the parameters b and c, using a single parameter, δ, which computes the decay in ϖ_i value averaged over $i = 2, 3, 4$. Formally, for every team we computed the following statistic

$$\delta = \frac{1}{3}\left(\frac{\varpi_3}{\varpi_2} + \frac{\varpi_4}{\varpi_3} + \frac{\varpi_5}{\varpi_4}\right).$$

Note that we average only over the first three levels of ϖ_i as we have found that adding more levels harms the predictive power. This makes sense since, as Figure 5 suggests, the derivative approaches 0 and becomes less informative as i grows.

Figure 6 plots the rank of a team against its δ value, with a linear regression line achieving $R^2 = 0.4219$ with p-value 0.002. An even better fit was obtained using a cubic regression function, achieving $R^2 = 0.48$ with the same p-value of 0.002.
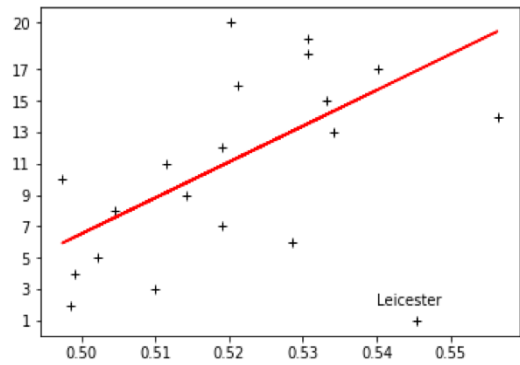
FIGURE 6. Linearly regressing the rank, y-axis, against decay rate δ, x-axis. The slower the divergence from additivity decays the worse the team is doing (further position in the table). $R^2 = 0.4219$, p-value 0.002.

TABLE 1. R^2 values when linearly regressing the rank of the team against ϖ_i-values. All p-values are below 0.05. In all regressions, the outlier, Leicester, was removed before regressing.

i	2	3-7	8	9	10	11	12
R^2	0.36	0.41	0.441	0.42	0.41	0.36	0.22

D. ANALYSIS 3

The previous sections point to the divergence from additivity measure as an indicator of the team's ability to move from local to global interactions and to coordinate their activity across different granularity levels. If this is the case, then the ϖ_i measure of the different macro-states should also have a predictive value of the team's rank.

Table 1 presents the predictive power of ϖ_i for different macros-states i. We use the linear regression's R^2 value as a proxy for the predictive power of the macro-state. The following trend is evident: the R^2 value increases as i increases up to $i = 8$, then decreases. This trend may be explained by a trade-off between the number of states at level i and the quality of information at that level. The predictive power of ϖ_i peaks at $i = 8$, which is roughly $N/2$ for most teams ($N \approx 15$ for most teams), and this is when the binomial coefficient $\binom{N}{i}$ peaks. For $i \geq 9$ we see a rapid decline in R^2 which is attributed to both a decrease in the number of states and a decrease in the variance of ϖ_i (Figure 5), making it rather useless for prediction.

Figure 7 plots the regression of the rank of a team against its ϖ_2 value. Leicester, the outlier, follows exactly the opposite trend, it has the highest ϖ_2 value and the lowest (meaning best) rank, #1.

Training various regression models with different subsets of the 12 macro level scores have gained the best model for the combinations of levels 4, 6 and 8 ($R^2 = 0.54$, p-value of 0.007).

E. ANALYSIS 4

In this last analysis we ask whether a smaller subset of the players may provide the same ability to predict the
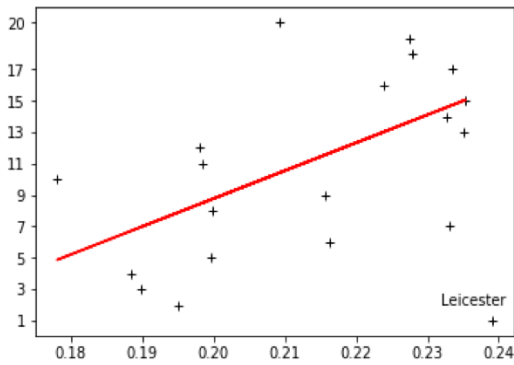
FIGURE 7. Linearly regressing the rank, *y*-axis against $\mathfrak{w}_2$, *x*-axis. The lower the $\mathfrak{w}_2$ value the lower (better) the rank. $R^2 = 0.36$, *p*-value 0.007.

team's rank using divergence from additivity. Similar "low-dimensionality" or sparsity phenomena were observed in other interactive systems such as online-social networks e.g. [17]–[19].

An affirmative answer bares important practical implications, as it means that the coach can focus on a much smaller set of states/players when trying to understand and manipulate his team.

A natural choice of a subset of players is the top *r* players in terms of ball passes (we will count the number of balls that a given player passed to others). Table 2 describes the R^2 values when computing $\mathfrak{w}_i$ but this time restricted to the top $r = 5$ players. Note that the highest we can climb in the $\mathfrak{w}_i$-hierarchy is $i = r - 1$.

TABLE 2. R^2 values when linearly regressing the rank of the team against $\mathfrak{w}_i$-values restricted to states contained in the top 5 players. All *p*-values are below 0.05. In all regressions, the outlier, Leicester, was removed before regressing.

i	2	3	4
R^2	0.25	0.409	0.413

The results show that we may gain the same prediction we have previously gained while using only a subset of the players. Specifically, the predictive power of triads ($\mathfrak{w}_3$) and tetrads ($\mathfrak{w}_4$) of the top 5 players is the same (even slightly larger) as taking all players into account.

Furthermore, smaller values of *r* gave weaker results. For example, looking at dyads of the top 4 players gave no significant prediction, and looking at triads gave $R^2 = 0.3$ with significance 0.01.

IV. DISCUSSION

The Latin phrases "Pars pro toto" (i.e. parts confused for the whole), and its complement "Totum pro parte" (i.e. the whole confused for its parts), have been used in the context of informal logic [20] to point to fallacious reasoning from the whole to its parts and vice versa. For example, the behavior of a whole soccer team cannot be trivially reduced the simple

aggregate of its players, and in statistics the "ecological fallacy" is a logical fallacy in the interpretation of statistical data where inferences about the nature of individuals are deduced from inference for the group to which those individuals belong. In this paper, we have presented a simple and interpretable way of measuring the extent to which wholes are composed from their sub-units and have shown the benefits of using these ideas in a context of analyzing the data of soccer teams. The measure of $\mathfrak{w}$ and the ReL may be further developed, for instance by including/developing a measure of redundancy between the sub-units of the systems [21]. The variety of applications for using the proposed methodology is wide. It may be used in the field of scientometrics for studying the real joint contribution of authors to academic publications, to study the way better combinations of working teams can be formed, and in the context of NLP for improving the automatic micro-reading of a text [22]. We therefore conclude by presenting a new tool for data analysis that should be further developed and tested in various contexts of engineering data-analytic methodologies for small systems.

REFERENCES

[1] C. Emmeche, S. Køppe, and F. Stjernfelt, "Explaining emergence: Towards an ontology of levels," *J. Gen. Philos. Sci.*, vol. 28, no. 1, pp. 83–117, Jan. 1997.

[2] H. Haken, *Some Aspects of Synergetics. From Laser Light to Cognition*. Cham, Switzerland: Springer, 2018, pp. 9–24.

[3] Y. Neuman, N. Israeli, D. Vilenchik, and Y. Cohen, "The adaptive behavior of a soccer team: An entropy-based analysis," *Entropy*, vol. 20, no. 10, p. 758, Oct. 2018.

[4] C. Tsallis, *Encyclopedia of Complexity and Systems Science*, R. A. Meyers, Ed. New York, NY, USA: Springer-Verlag, 2009.

[5] C. TsallisEmail, "Nonadditive entropy: The concept and its use," *Eur. Phys. J. A, Hadrons Nuclei*, vol. 40, no. 3, pp. 257–266, Jun. 2009.

[6] Y. Neuman, D. Noble, and Y. Cohen, "Is the whole different from the sum of its parts? A proposed procedure for measuring divergence from additivity," *Int. J. Gen. Syst.*, vol. 47, no. 7, pp. 665–678, Sep. 2018.

[7] Y. Neuman, Y. Neuman, and Y. Cohen, "A novel procedure for measuring semantic synergy," *Complexity*, vol. 2017, Jan. 2017, Art. no. 5785617.

[8] B. J. Juhasz, Y.-H. Lai, and M. L. Woodcock, "A database of 629 english compound words: Ratings of familiarity, lexeme meaning dominance, semantic transparency, age of acquisition, imageability, and sensory experience," *Behav. Res. Methods*, vol. 47, no. 4, pp. 1004–1019, Dec. 2015.

[9] R. Goldblatt, *Topoi: The Categorial Analysis of Logic*. New York, NY, USA: Elseviers, 1979.

[10] F. W. Lawvere and S. H. Schanuel, *Conceptual Mathematics: A First Introduction to Categories*. Cambridge, MA, USA: Cambridge Univ. Press, 1991.

[11] J. A. Davies, "The individual success of musicians, like that of physicists, follows a stretched exponential distribution," *Eur. Phys. J. B, Condens. Matter Complex Syst.*, vol. 27, no. 4, pp. 445–447, Jun. 2002.

[12] R. K. Pan and S. Sinha, "The statistical laws of popularity: Universal properties of the box-office dynamics of motion pictures," *New J. Phys.*, vol. 12, no. 11, Nov. 2010, Art. no. 115004.

[13] I. R. Cohen and S. Efroni, "The immune system computes the state of the body: Crowd wisdom, machine learning, and immune cell reference repertoires help manage inflammation," *Frontiers Immunol.*, vol. 10, p. 10, Jan. 2019.

[14] Y. Neuman, *Mathematical Structures of Natural Intelligence*. New York, NY, USA: Springer, 2017.

[15] E. Friedgut, and J. Bourgain, "Sharp thresholds of graph properties, and the *k*-sat problem," *J. Amer. Math. Soc.*, vol. 12, no. 4, pp. 1017–1054, May 1999.

[16] A. Coja-Oghlan and D. Vilenchik, "Chasing the k-colorability threshold," in *Proc. IEEE 54th Annu. Symp. Found. Comput. Sci.*, Oct. 2013, pp. 380–389.

[17] D. Vilenchik, "Simple statistics are sometime too simple: A case study in social media data," *IEEE Trans. Knowl. Data Eng.*, to be published.

[18] D. Vilenchik, "The million tweets fallacy: Activity and feedback are uncorrelated," in *Proc. 12th Int. AAAI Conf. Web Soc. Media*, Jun. 2018, pp. 688–691.

[19] C. Canali and R. Lancellotti, "A quantitative methodology based on component analysis to identify key users in social networks," *Int. J. Soc. Netw. Mining*, vol. 1, no. 1, pp. 27–50, Jan. 2012.

[20] D. N. Walton, *Informal Logic a Handbook for Critical Argumentation*. Cambridge, U.K.: Cambridge Univ. Press, 1989.

[21] Y. Neuman, Y. Cohen, N. Israeli, and B. Tamir, "A proposed methodology for studying the historical trajectory of words' meaning through tsallis entropy," *Phys. A, Statist. Mech. Appl.*, vol. 492, pp. 804–813, Feb. 2018.

[22] J. Berant *et al.*, "Modeling biological processes for reading comprehension," in *Proc. EMNLP*, 2014, pp. 1499–1510.

YAIR NEUMAN was born in 1968. He received the Ph.D. degree in cognition from Hebrew University. He is currently a Full Professor with the Department of Cognitive and Brain Sciences, BGU, Israel. He is interested in developing novel computational models for various interdisciplinary problems of data analytics and has developed cutting edge algorithms for various contexts and agencies ranging from IARPA to CitiBank. He is the author of numerous academic papers and six books, the most recent is *Mathematical Structures of Natural Intelligence* (Springer).

DAN VILENCHIK was born in 1977. He received the Ph.D. degree in computer science from Tel-Aviv University, Israel, in 2009. He is currently a Lecturer with the Department of Communication Systems Engineering, Ben Gurion University of the Negev, Israel. His research interests include data mining, optimization, and computational statistics.

• • •